Aula 00

BACEN (Analista - Área 2 - Economia e Finanças) Estatística e Econometria

Autor:

Amanda Aires, Equipe Exatas Estratégia Concursos

08 de Janeiro de 2025

Índice

APRESENTAÇÃO DO CURSO

Olá, pessoal! Tudo bem?

É com grande satisfação damos início ao nosso curso de **Estatística**!

Os professores **Luana Brandão** e **Djefferson Maranhão** ficarão responsáveis pela elaboração do **Livro Digital**.

Antes de continuarmos, vamos apresentar os professores do material escrito:

Luana Brandão: Oi, pessoal! O meu nome é Luana Brandão e sou professora de Estatística do Estratégia Concursos. Sou Graduada, Mestre e Doutora em Engenharia de Produção, pela Universidade Federal Fluminense. Passei nos concursos de Auditor Fiscal (2009/2010) e Analista Tributário (2009) da Receita Federal e de Auditor Fiscal do Estado do Rio de Janeiro (2010). Sou Auditora Fiscal do Estado do RJ desde 2010. Vamos juntos nesse caminho até a aprovação? **@professoraluanabrandao**

Djefferson Maranhão: Olá, caros alunos! Meu nome é Djefferson Maranhão, sou professor de Estatística do Estratégia Concursos. Graduado e Mestre em Ciência da Computação pela Universidade Federal do Maranhão (UFMA). Desde 2015, sou Auditor da Controladoria Geral do Estado do Maranhão (2015 - 5º lugar). Também exerci os cargos de Analista de Sistemas na UFMA (2010 - 1º lugar) e no TJ-MA (2011 - 1º lugar). Sinto-me honrado em fazer parte de sua jornada rumo à aprovação.

O material escrito em **PDF** está sendo construído para ser sua fonte **autossuficiente** de estudos. Isso significa que o livro digital será **completo** e **voltado para o seu edital**, justamente para que você não perca o seu precioso tempo procurando o conteúdo que será cobrado na sua prova. Ademais, sempre que necessário, você poderá fazer perguntas sobre as aulas no **fórum de dúvidas**. **Bons estudos!**

AVISO IMPORTANTE!

Olá, Alunos (as)!

Passando para informá-los a respeito da **disposição das questões** dentro do nosso material didático. Informamos que a escolha das bancas, dentro dos nossos Livros Digitais, é feita de maneira estratégica e pedagógica pelos nossos professores a fim de proporcionar a melhor didática e o melhor direcionamento daquilo que mais se aproxima do formato de cobrança da banca do seu concurso.

Assim, o formato de questões divididas por tópico facilitará o seu processo de estudo, deixando mais alinhado às disposições constantes no edital.

No mais, continuaremos à disposição de todos no Fórum de dúvidas!

Atenciosamente,

Equipe Exatas

Estratégia Concursos

Olá, amigo(a)!

Essa aula reúne assuntos mais **específicos** envolvendo variáveis aleatórias: alguns simples e outros que exigem mais atenção. Ademais, quando tratarmos de variáveis contínuas, utilizaremos ferramentas de Cálculo, como fizemos na aula de Variáveis Contínuas.

Novamente, você verá as principais operações, para resolver a grande maioria das questões. Mas, se você tem pouco tempo de estudo, acompanhe esta aula sem se preocupar com os pontos que envolvem Cálculo. Esse tipo de questão está se tornando mais comum nas provas, mas ainda são pouco frequentes.

Vamos lá?!

Luana Brandão

Doutora em Engenharia de Produção (UFF)

Auditora Fiscal da SEFAZ-RJ

Se tiver alguma dúvida, entre em **contato** comigo!

professoraluanabrandao@gmail.com

@professoraluanabrandao

"A alegria está na luta, na tentativa, no sofrimento envolvido.

Não na vitória propriamente dita."

Mahatma Gandhi

MOMENTOS E FUNÇÃO GERATRIZ

Momentos de uma Variável Aleatória

Os **momentos** de variáveis aleatórias **caracterizam** as suas distribuições de probabilidade, tanto para variáveis discretas quanto para variáveis contínuas.

Em particular, o **primeiro momento** caracteriza a **tendência central**; o **segundo** caracteriza a **dispersão**; o **terceiro** momento, a **assimetria**; e o **quarto**, a **curtose** (que representa o achatamento da distribuição).

É possível que distribuições **distintas** apresentem os **mesmos momentos**.

Por definição, o ***k*-ésimo momento** (ou **momento de ordem *k***) de uma variável aleatória X, que podemos indicar como $\boldsymbol{\mu_k}$, é a esperança de X^k:

$$\mu_k = E(X^k)$$

Essa definição vale tanto para variáveis discretas, quanto para variáveis contínuas. Especificamente, para o caso **discreto**, a esperança de X^k equivale ao somatório de $(x_i)^k$, multiplicado pela respectiva probabilidade:

$$\mu_k = E(X^k) = \sum_i (x_i)^k . P(X = x_i)$$

Para ilustrar essa fórmula, vamos supor uma distribuição binomial com $n = 2$ e $p = 0{,}5$ (logo, $q = 1 - p = 0{,}5$). A função de probabilidade para esse exemplo é:

- X = 0: $P(X = 0) = C_{2,0}.p^0.q^2 = 1 \times 1 \times 0{,}5^2 = 0{,}25$
- X = 1: $P(X = 1) = C_{2,1}.p^1.q^1 = 2 \times 0{,}5 \times 0{,}5 = 0{,}5$
- X = 2: $P(X = 2) = C_{2,2}.p^2.q^0 = 1 \times 0{,}5^2 \times 1 = 0{,}25$

Assim, o k-ésimo momento dessa variável é dado por:

$$\mu_k = E(X^k) = \sum_i (x_i)^k . P(X = x_i)$$

$$\mu_k = 0^k \times 0,25 + 1^k \times 0,5 + 2^k \times 0,25 = 0,5 + 2^k \times 0,25$$

- para k = 1: $\mu_1 = 0,5 + 2^1 \times 0,25 = 1$
- para k = 2: $\mu_2 = 0,5 + 2^2 \times 0,25 = 1,5$
- para k = 3: $\mu_3 = 0,5 + 2^3 \times 0,25 = 2,5$

...

Para o caso **contínuo**, a situação é análoga, substituindo o somatório pela integral e a probabilidade pela função densidade de probabilidade:

$$\mu_k = E(X^k) = \int_{x_I}^{x_S} (x)^k . f(x). dx$$

Em que x_I é o menor valor possível para a variável e x_S é o maior valor possível. Se a f.d.p. apresentar valor em toda a reta real, substituímos x_I por $-\infty$ e x_S por ∞.

EXEMPLIFICANDO

Vamos supor uma distribuição contínua **uniforme**, no intervalo $0 \leq x \leq 1$. A f.d.p. dessa distribuição é:

$$f(x) = \frac{1}{b-a} = \frac{1}{1-0} = 1$$

Assim, o k-ésimo momento dessa variável é dado por:

$$\mu_k = E(X^k) = \int_{x_I}^{x_S} (x)^k . f(x). dx$$

$$\mu_k = \int_0^1 (x)^k . 1. dx$$

A integral, sem aplicar os limites, é:

$$\int (x)^k . dx = \frac{x^{k+1}}{k+1}$$

Aplicando os limites, temos:

$$F(1) - F(0) = \frac{1^{k+1}}{k+1} - \frac{0^{k+1}}{k+1} = \frac{1}{k+1}$$

$$\mu_k = \frac{1}{k+1}$$

- para k = 1: $\mu_1 = \frac{1}{1+1} = \frac{1}{2}$
- para k = 2: $\mu_2 = \frac{1}{2+1} = \frac{1}{3}$
- para k = 3: $\mu_3 = \frac{1}{3+1} = \frac{1}{4}$

...

Em qualquer caso, seja para variáveis discretas, seja para variáveis contínuas, o **primeiro** momento (ou momento de **primeira ordem**), ou seja, para $\boldsymbol{k = 1}$, é:

$$\mu_1 = E(X^1) = E(X)$$

Ou seja, o **primeiro** momento (ou momento de **primeira ordem**) é igual à **esperança**.

De fato, podemos verificar isso a partir dos nossos exemplos. Em relação à variável binomial que vimos anteriormente, com $n = 2$ e $p = 0{,}5$, sabemos que a esperança é dada por:

$$E(X) = n \times p = 2 \times 0{,}5 = 1$$

Que é o mesmo resultado que obtivemos para o nosso primeiro exemplo.

Já para uma variável contínua uniforme no intervalo (0,1), sabemos que a esperança corresponde à média aritmética nos extremos do intervalo:

$$E(X) = \frac{0+1}{2} = \frac{1}{2}$$

Que é o mesmo resultado que obtivemos para o nosso segundo exemplo.

E quanto ao **segundo** momento (ou momento de **segunda ordem**), ou seja, para $k = 2$, temos:

$$\mu_2 = E(X^2)$$

Essa é a primeira expressão do **cálculo** da **variância**: $V(X) = E(X^2) - [E(X)]^2$.

Ou seja, podemos calcular a variância pela **diferença** entre o **segundo** momento e o **quadrado** do **primeiro** momento:

$$\boldsymbol{V(X) = E(X^2) - [E(X)]^2 = \mu_2 - (\mu_1)^2}$$

Vamos verificar isso para os nossos exemplos. Para o nosso exemplo da **distribuição binomial**, vimos que o primeiro momento (k = 1) $\mu_1 = 1$ e que o segundo momento (k = 2) é $\mu_2 = 1{,}5$. Então, a variância é a diferença:

$$V(X) = \mu_2 - (\mu_1)^2 = 1{,}5 - 1^2 = 0{,}5$$

Por outro lado, sabemos que a variância da distribuição binomial é dada por $V(X) = n \times p \times q$. Sabendo que $n = 2$ e $p = q = 0{,}5$, então a variância é:

$$V(X) = n \times p \times q = 2 \times 0{,}5 \times 0{,}5 = 0{,}5$$

Que é o mesmo resultado que obtivemos pela diferença entre o segundo momento e o quadrado do primeiro momento, $\mu_2 - (\mu_1)^2$!

Para o nosso exemplo da **distribuição contínua uniforme**, vimos que $\mu_1 = \frac{1}{2}$ e $\mu_2 = \frac{1}{3}$. Então, a variância é a diferença:

$$V(X) = \mu_2 - (\mu_1)^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{4-3}{12} = \frac{1}{12}$$

Por outro lado, sabemos que a variância da distribuição contínua uniforme é $V(X) = \frac{(b-a)^2}{12}$. Sabendo que a=0 e b=1, temos:

$$V(X) = \frac{(1-0)^2}{12} = \frac{1}{12}$$

Que é o mesmo resultado!

ESQUEMATIZANDO

$\boldsymbol{k}$**-ésimo momento**: $\boldsymbol{\mu_k = E(X^k)}$
Esperança = 1º momento: $\boldsymbol{\mu_1 = E(X)}$
Variância = 2º momento – [1º momento]²: $\boldsymbol{V(X) = \mu_2 - [\mu_1]^2}$

(2017 – TRF – 2ª Região – *Adaptada*) Considerando os Momentos de uma variável aleatória, julgue a afirmativa a seguir.

Duas variáveis aleatórias com funções densidade de probabilidade distintas podem ter os mesmos momentos.

Comentários:

Vimos que distribuições de probabilidade **distintas** podem ter os **mesmos** momentos.

Resposta: Certo.

(CESPE/2015 – Especialista em Gestão de Telecomunicações Telebras) Um analista da TELEBRAS, a fim de verificar o tempo durante o qual um grupo de consumidores ficou sem o serviço de Internet do qual eram usuários, selecionou uma amostra de 10 consumidores críticos. Os dados coletados, em minutos, referentes a esses consumidores foram listados na tabela seguinte.

consumidor	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
tempo (em minutos)	8	2	3	5	7	7	10	9	4	5

Com base nessa situação hipotética, julgue o item subsequente.

O segundo momento dos dados é menor do que o primeiro momento ao quadrado.

Comentários:

Para o caso discreto, o k-ésimo momento é definido como:

$$\mu_k = E(X^k) = \sum_i (x_i)^k \,.\, P(x_i)$$

O primeiro momento é a esperança matemática $\mu = E(X)$ e o segundo momento corresponde a $E(X^2)$.

A diferença entre o segundo momento e o primeiro momento ao quadrado é a variância:

$$\boldsymbol{\sigma^2 = E(X^2) - \mu^2}$$

Sabemos que a variância é sempre maior ou igual a zero, portanto:

$$\boldsymbol{\sigma^2 = E(X^2) - \mu^2 \geq 0}$$

$$E(X^2) \geq \mu^2$$

Assim, o segundo momento é sempre **maior ou igual** ao primeiro momento ao quadrado.

Mas vamos fazer os cálculos. O primeiro momento é:

$$\mu_1 = E(X^1) = \frac{8+2+3+5+7+7+10+9+4+5}{10} = \frac{60}{10} = 6$$

O primeiro momento ao quadrado é ${\mu_1}^2 = 36$. O segundo momento é:

$$\mu_2 = E(X^2) = \frac{64+4+9+25+49+49+100+81+16+25}{10} = \frac{422}{10} = 42{,}2$$

Confirmamos, portanto, que o segundo momento é **maior** que o primeiro momento ao quadrado.

Gabarito: Errado.

(CESPE/2013 – Analista Judiciário da STM) Supondo que o custo unitário X de um processo de execução fiscal na justiça federal seja descrito por uma distribuição exponencial com média igual a R$ 5.000, julgue os próximos itens.

Para cada r = 1, 2, 3, ... , o momento de ordem r da variável X, E[X'] é tal que E[X'] = (R$ 5.000)^r.

Comentários:

O momento de ordem r de uma variável aleatória é:

$$\mu_r = E(X^r)$$

Sabendo que a média da distribuição é $E(X) = 5000$, então a expressão (R$ 5.000)^r que o item forneceu corresponde a $[E(X)]^r$, que é **diferente** de $E(X^r)$.

Inclusive, a variância é justamente a diferença entre $E(X^2)$ e $[E(X)]^2$. Logo, o item está errado.

Gabarito: Errado.

Momentos Centrais

Também existe a definição de momento **central**. O ***k*-ésimo momento central** (ou **momento central de ordem *k***) de uma variável aleatória X, que podemos denotar por $\boldsymbol{\mu_{k_C}}$, é definido como:

$$\boxed{\mu_{k_C} = E((X-\mu)^k)}$$

Ou seja, o momento central trabalha com as **diferenças** em relação à **média** da distribuição.

Essa definição também vale tanto para variáveis discretas, quanto para variáveis contínuas. Especificamente, para o caso discreto, temos:

$$\mu_{k_C} = E((X-\mu)^k) = \sum_i (x_i-\mu)^k . P(X = x_i)$$

EXEMPLIFICANDO

Para ilustrar essa fórmula, vamos retornar ao nosso exemplo da distribuição binomial com $n = 2$ e $p = q = 0{,}5$. Já calculamos a função de probabilidade:

- X = 0: $P(X = 0) = 0{,}25$
- X = 1: $P(X = 1) = 0{,}5$
- X = 2: $P(X = 2) = 0{,}25$

Conhecendo a média $\mu = E(X) = 1$, podemos calcular o k-ésimo momento central dessa variável:

$$\mu_{k_C} = E((X - \mu)^k) = \sum_i (x_i - \mu)^k . P(X = x_i)$$

$$\mu_{k_C} = (0 - 1)^k \times 0{,}25 + (1 - 1)^k \times 0{,}5 + (2 - 1)^k \times 0{,}25$$

$$\mu_{k_C} = (-1)^k \times 0{,}25 + 1^k \times 0{,}25$$

- para k = 1: $\mu_{1_C} = (-1)^1 \times 0{,}25 + 1^1 \times 0{,}25 = 0$
- para k = 2: $\mu_{2_C} = (-1)^2 \times 0{,}25 + 1^2 \times 0{,}25 = 0{,}5$
- para k = 3: $\mu_{3_C} = (-1)^3 \times 0{,}25 + 1^3 \times 0{,}25 = 0$
- para k = 4: $\mu_{4_C} = (-1)^4 \times 0{,}25 + 1^4 \times 0{,}25 = 0{,}5$

...

Analogamente, para o caso contínuo, temos:

$$\mu_{k_C} = E((X - \mu)^k) = \int_{x_I}^{x_S} (x - \mu)^k . f(x). dx$$

Para $k = 1$ (**primeiro** momento central), temos:

$$\mu_{1_C} = E(X - \mu)$$

Ou seja, é o valor esperado da diferença entre os elementos e a média. Pela definição de média, esse valor é igual a zero (como acabamos de ver, para o nosso exemplo da distribuição binomial). Portanto, o **primeiro** momento central é sempre **nulo**.

$$\boldsymbol{\mu_{1_C} = 0}$$

Para $k = 2$ (**segundo momento central**), temos:

$$\mu_{2_C} = E((X - \mu)^2)$$

Essa é a definição de **variância**. Por isso, o **segundo** momento central, ou momento central de segunda ordem, é igual à **variância**!

De fato, já calculamos que a variância para o nosso exemplo de distribuição binomial é V(X) = 0,5, que foi justamente o resultado que encontramos pelo segundo momento central no exemplo acima.

ESQUEMATIZANDO

$\boldsymbol{k}$**-ésimo momento central**: $\boldsymbol{\mu_{k_C} = E((X - \mu)^k)}$

1º momento central é sempre **nulo**: $\boldsymbol{\mu_{1_C} = 0}$

2º momento central = **Variância**: $\boldsymbol{\mu_{2_C} = V(X)}$

Já, o **coeficiente de assimetria** está associado ao **terceiro momento central** ($k = 3$):

$$\mu_{3_C} = E((X - \mu)^3)$$

Mais precisamente, o **coeficiente de assimetria** é calculado pela fórmula abaixo, em que σ é o desvio padrão:

$$b_1 = E\left(\frac{X - \mu}{\sigma}\right)^3$$

Para variáveis discretas, temos:

$$b_1 = E\left(\frac{\boldsymbol{X - \mu}}{\boldsymbol{\sigma}}\right)^3 = \sum_i \left(\frac{X - \mu}{\sigma}\right)^3 \times P(X = x_i)$$

A assimetria é dita **positiva** (ou à direita) quando $\boldsymbol{b_1 > 0}$. Como o desvio padrão é necessariamente positivo, a assimetria é positiva quando o terceiro momento central for positivo: $\boldsymbol{\mu_{3_C} > 0}$. Analogamente, a assimetria é **negativa** quando o terceiro momento central for negativo: $\boldsymbol{\mu_{3_C} < 0}$.

Quando a distribuição é **simétrica**, todos os **momentos centrais** de ordem **ímpar** são iguais a **zero**, incluindo o 3º momento central (como vimos para o exemplo da distribuição binomial com $p = 0{,}5$).

No entanto, é possível ter $\mu_{3_C} = 0$ e a distribuição **não** ser **simétrica**.

EXEMPLIFICANDO

Sabemos que a distribuição binomial tem assimetria positiva quando p < 0,5. Vamos, então, efetuar os cálculos para p = 0,2 (logo, q = 0,8), mantendo n = 2:

- X = 0: $P(X = 0) = C_{2,0}.p^0.q^2 = 1 \times 1 \times 0{,}8^2 = 0{,}64$
- X = 1: $P(X = 1) = C_{2,1}.p^1.q^1 = 2 \times 0{,}2 \times 0{,}8 = 0{,}32$
- X = 2: $P(X = 2) = C_{2,2}.p^2.q^0 = 1 \times 0{,}2^2 \times 1 = 0{,}04$

Para calcular os momentos centrais, primeiro calculamos a média:

$$\mu = E(X) = n \times p = 2 \times 0{,}2 = 0{,}4$$

E para calcular o coeficiente de assimetria, precisamos do desvio padrão. Antes, vamos calcular a variância:

$$V(X) = n \times p \times q = 2 \times 0{,}2 \times 0{,}8 = 0{,}32$$

Logo, o desvio padrão é a raiz quadrada: $\sigma = \sqrt{0{,}32} \cong 0{,}566$

O coeficiente de assimetria é dado por:

$$b_1 = E\left(\frac{X-\mu}{\sigma}\right)^3 = \left(\frac{0-0{,}4}{0{,}566}\right)^3 \times 0{,}64 + \left(\frac{1-0{,}4}{0{,}566}\right)^3 \times 0{,}32 + \left(\frac{2-0{,}4}{0{,}566}\right)^3 \times 0{,}04$$

$$b_1 \cong -0{,}35 + 1{,}19 + 22{,}6 = 23{,}44$$

E o **coeficiente de curtose** está associado ao **quarto momento central** ($k = 4$):

$$\mu_{4_C} = E((X-\mu)^4)$$

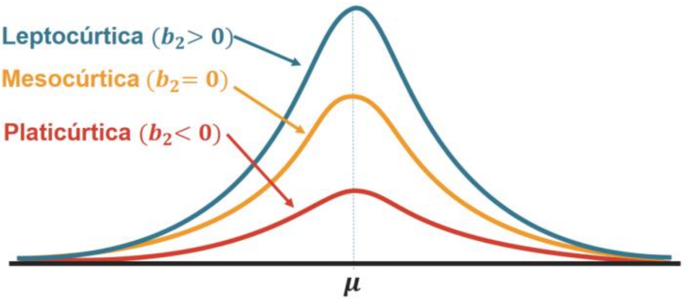
Mais precisamente, o **coeficiente de curtose** é calculado pela fórmula abaixo, em que σ é o desvio padrão:

$$b_2 = E\left(\frac{\boldsymbol{X}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}\right)^4 - 3$$

Para variáveis discretas, temos:

$$b_2 = E\left(\frac{\boldsymbol{X}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}\right)^4 - 3 = \sum_i \left(\frac{X-\mu}{\sigma}\right)^4 \times P(X = x_i) - 3$$

Distribuições **platicúrticas** ("achatadas") apresentam $\boldsymbol{b_2 < 0}$; distribuições **mesocúrticas** ("normais") apresentam $\boldsymbol{b_2 = 0}$; e distribuições **leptocúrticas** ("alongadas") apresentam $\boldsymbol{b_2 > 0}$, como ilustrado abaixo[1]:



EXEMPLIFICANDO

Para o mesmo exemplo, em que $\mu = 0{,}4$ e $\sigma \cong 0{,}566$, o coeficiente de curtose é:

$$b_2 = E\left(\frac{X-\mu}{\sigma}\right)^4 = \left(\frac{0-0{,}4}{0{,}566}\right)^4 \times 0{,}64 + \left(\frac{1-0{,}4}{0{,}566}\right)^4 \times 0{,}32 + \left(\frac{2-0{,}4}{0{,}566}\right)^4 \times 0{,}04 - 3$$

$$b_2 \cong 0{,}25 + 1{,}26 + 63{,}86 - 3 = 62{,}37$$

Pontue-se que é possível definir o coeficiente de curtose **sem a subtração**, isto é:

$$b_{2*} = E\left(\frac{\boldsymbol{X-\mu}}{\boldsymbol{\sigma}}\right)^{\mathbf{4}}$$

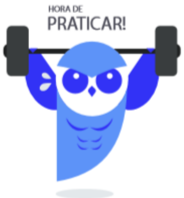
Porém, esse coeficiente apresenta sempre valores positivos, devendo ser comparado ao valor 3.

Distribuições **platicúrticas** ("achatadas") apresentam $\boldsymbol{b_{2*} < 3}$; distribuições **mesocúrticas** ("normais") apresentam $\boldsymbol{b_{2*} = 3}$; e distribuições **leptocúrticas** ("alongadas") apresentam $\boldsymbol{b_{2*} > 3}$.

[1] Gráfico obtido no site http://www.portalaction.com.br/estatistica-basica/26-curtose

Os momentos das variáveis aleatórias **nem sempre existem**.

Por exemplo, a distribuição de Cauchy (definida como a razão entre variáveis com distribuição normal) não possui esperança (isto é, primeiro momento) ou variância (isto é, segundo momento central). Na verdade, a distribuição de Cauchy não possui momento algum.



(CESPE/2011 – TER/ES)

quantidade de eleitores	quantidade de municípios
0 ├ 2.000	364
2.000 ├ 4.000	1.000
4.000 ├ 6.000	3.000
6.000 ├ 8.000	1.000
8.000 ├ 10.000	200
total	5.564

A tabela acima apresenta uma distribuição hipotética das quantidades de eleitores que não votaram no segundo turno da eleição para presidente da República bem como os números de municípios em que essas quantidades ocorreram. Com base nessa tabela, julgue os itens seguintes, relativos à análise exploratória de dados.

A curtose da distribuição em questão pode ser avaliada com base na estimativa do quarto momento central, a qual deve ser comparada com o valor de referência 3, visto que todas as distribuições simétricas possuem quarto momento central igual a 3.

Comentários:

A curtose é, de fato, avaliada com base no quarto momento, sendo comparada ao valor de referência 3. Porém, a simetria da distribuição não tem relação com o quarto momento.

Gabarito: Errado.

Momentos Amostrais

Os momentos também podem ser calculados a partir de **amostras**. O **k-ésimo momento amostral** (ou **momento amostral de ordem k**), que podemos denotar por m_k é dado por:

$$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

Em que n é o tamanho da amostra. Essa fórmula é similar ao k-ésimo momento populacional, porém, em vez de multiplicarmos cada valor pela respectiva probabilidade, basta **dividir** o somatório pelo **total de elementos**.

EXEMPLIFICANDO

Vamos supor o seguinte conjunto de dados X = {0, 1, 2, 3, 4}, de n = 5 elementos. A tabela abaixo apresenta os valores de X_i^k para k = 1, 2 e 3.

X	X^1	X^2	X^3
0	$0^1 = 0$	$0^2 = 0$	$0^3 = 0$
1	$1^1 = 1$	$1^2 = 1$	$1^3 = 1$
2	$2^1 = 2$	$2^2 = 4$	$2^3 = 8$
3	$3^1 = 3$	$3^2 = 9$	$3^3 = 27$
4	$4^1 = 4$	$4^2 = 16$	$4^3 = 64$
$\sum X^k$	10	30	100
$\frac{1}{n}\sum X^k$	$\frac{10}{5} = 2$	$\frac{30}{5} = 6$	$\frac{100}{5} = 20$

A linha $\sum X^k$ contém a soma dos valores X_i^k de cada coluna e a linha $\frac{1}{n}\sum X^k$ divide os valores da linha anterior por n = 5 e, assim, apresenta os valores dos **momentos amostrais** para k = 1, 2 e 3.

Note que o **primeiro momento amostral** (ou **momento amostral de primeira ordem**), isto é, para $k = 1$, é a própria **média amostral**:

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$

Assim como definimos momentos **centrais**, também é possível definir os **momentos amostrais em torno da média $\bar{X}$**. O k-ésimo momento amostral de ordem k em torno da média é definido como:

$$m_{k_C} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k$$

EXEMPLIFICANDO

Vamos supor o mesmo conjunto de dados X = {0, 1, 2, 3, 4}, de n = 5 elementos. A tabela abaixo apresenta os valores de $(X - \bar{X})^k$ para k = 1, 2 e 3.

$X - \bar{X}$	$(X - \bar{X})^1$	$(X - \bar{X})^2$	$(X - \bar{X})^3$
0 – 2 = -2	$(-2)^1 = -2$	$(-2)^2 = 4$	$(-2)^3 = -8$
1 – 2 = -1	$(-1)^1 = -1$	$(-1)^2 = 1$	$(-1)^3 = -1$
2 – 2 = 0	$0^1 = 0$	$0^2 = 0$	$0^3 = 0$
3 – 2 = 1	$1^1 = 1$	$1^2 = 1$	$1^3 = 1$
4 – 2 = 2	$2^1 = 2$	$2^2 = 4$	$2^3 = 8$
$\sum(X - \bar{X})^k$	0	10	0
$\frac{1}{n}\sum(X - \bar{X})^k$	$\frac{0}{5} = 0$	$\frac{10}{5} = 2$	$\frac{0}{5} = 0$

A linha $\sum(X - \bar{X})^k$ contém a soma dos valores $(X - \bar{X})^k$ de cada coluna e a linha $\frac{1}{n}\sum(X - \bar{X})^k$ divide os valores da linha anterior por n = 5 e, assim, apresenta os valores dos **momentos amostrais em torno de $\bar{X}$** para k = 1, 2 e 3.

Pela definição de média, o **primeiro momento amostral em torno de $\bar{X}$** será sempre **nulo** (assim como vimos em relação ao primeiro momento central):

$$\boldsymbol{m_{1_C} = 0}$$

FIQUE ATENTO!

Diferentemente dos momentos populacionais, o **segundo** momento amostral em torno de $\bar{X}$ **não** é igual à **variância amostral**.

Como vimos acima, no cálculo do **segundo momento amostral em torno de $\bar{X}$**, m_{2_C}, dividimos o somatório de $(X_i - \bar{X})^2$ pelo **número de elementos n**, enquanto, no **cálculo da variância amostral**, s^2, dividimos esse somatório por $n - 1$:

$$s^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n - 1}$$

Para o nosso exemplo, a variância amostral é dada por:

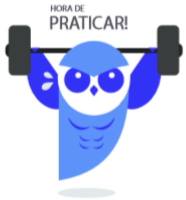
$$s^2 = \frac{10}{4} = 2{,}5$$

Os coeficientes de **assimetria** e de **curtose** também podem ser calculados a partir dos dados amostrais, dividindo o terceiro e quarto momentos em torno de $\bar{X}$, respectivamente, pelo desvio padrão amostral, s:

$$b_1 = E\left(\frac{X - \bar{X}}{s}\right)^3$$

$$b_2 = E\left(\frac{X - \bar{X}}{s}\right)^4 - 3$$

Para o nosso exemplo, o terceiro momento em torno de $\bar{X}$ foi igual a zero, o que implica no coeficiente de assimetria igual a zero. Isso ocorre porque o conjunto de dados é simétrico em torno da média $\bar{X} = 2$.



(2017 – TRF – 2ª Região – *Adaptada*) Considerando os Momentos de uma variável aleatória, julgue a afirmativa a seguir.

O k-ésimo momento amostral é dado por $m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$ e o k-ésimo momento populacional é dado por $E(X^k)$.

Comentários:

Vimos que, de fato, o k-ésimo momento amostral é definido como:

$$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

Anteriormente, também vimos que o k-ésimo momento populacional é definido, tanto para variáveis discretas, quanto para variáveis contínuas, como:

$$\mu_k = E(X^k)$$

Resposta: Certo.

(2019 – IBGE) A palavra momento é usada com frequência em contextos estatísticos e refere-se à soma dos desvios de uma variável X em relação à média relacionada ao tamanho da amostra. A fórmula geral para o cálculo dos momentos amostrais em torno da média é dada por:

$M_k = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^k}{n}$ ou $M_k = \frac{\sum_{i=1}^{c} f_i (x_i - \bar{x})^k}{\sum_{i=1}^{c} f_i}$, para dados agrupados, onde "c" é o número de classes

Sendo M_k representante do k-ésimo momento em torno da média amostral, considerando uma variável X, é correto afirmar que:

a) o primeiro momento em torno da média é a variância.

b) o primeiro momento em torno da média é a média e o segundo a variância.

c) o primeiro momento em torno da média é a moda.

d) o primeiro momento em torno da média é sempre igual a zero.

e) o primeiro momento em torno da média é o desvio padrão e o segundo a variância.

Comentários:

Vimos que o **primeiro momento amostral em torno da média** é sempre igual a **zero**. Logo, a alternativa correta é a letra D. Ele não é igual à variância (como descrito na alternativa A), à média (como descrito na alternativa B), à moda (como descrito na C) e nem ao desvio padrão (como descrito na E).

Em relação à alternativa B, vale acrescentar que a média amostral corresponde ao **primeiro momento amostral**, e não ao primeiro momento **em torno da média**.

Além disso, o segundo momento em torno da média amostral **não** corresponde à variância amostral (como descrito nas alternativas B e E), pois na fórmula do segundo momento, dividimos por n, enquanto na fórmula da variância amostral, dividimos por n – 1.

Gabarito: D.

Função Geradora de Momentos (f.g.m.)

Todos os momentos de uma variável aleatória (quando existirem) podem ser calculados a partir da **função geradora de momentos** (ou **função geratriz de momentos** ou, simplesmente, **f.g.m.**).

Ela é definida, tanto para variáveis discretas, quanto para variáveis contínuas, como a seguinte função:

$$\boxed{\boldsymbol{M_X(t) = E(e^{t.X})}}$$

Especificamente para o **caso discreto**, a f.g.m. é dada por:

$$M_X(t) = E(e^{t.X}) = \sum_i e^{t.x_i} \times P(X = x_i)$$

EXEMPLIFICANDO

Para o nosso exemplo da distribuição binomial com $n = 2$ e $p = 0{,}2$, calculamos as seguintes probabilidades para os valores de X:

- X = 0: $P(X = 0) = 0{,}64$
- X = 1: $P(X = 1) = 0{,}32$
- X = 2: $P(X = 2) = 0{,}04$

Assim, a f.g.m. dessa distribuição é dada por:

$$M_X(t) = \sum_i e^{t.x_i} \times P(X = x_i)$$

$$M_X(t) = e^{t\times 0} \times 0{,}64 + e^{t\times 1} \times 0{,}32 + e^{t\times 2} \times 0{,}04$$

Sabendo que $e^{t\times 0} = e^0 = 1$, temos:

$$M_X(t) = 0{,}64 + 0{,}32.e^t + 0{,}04.e^{2.t}$$

Para o **caso contínuo**, substituímos o somatório pela integral e a função de probabilidade, pela função densidade de probabilidade:

$$M_X(t) = E(e^{t.X}) = \int_{x_I}^{x_S} e^{t.X}.f(x).dx$$

Em que x_I e x_S são, respectivamente, o menor e o maior valor da variável. Caso a f.d.p. não tenha limite, isto é, apresente valor para toda a reta real, substituímos x_I por $-\infty$ e x_S por ∞.

Para o exemplo da distribuição contínua uniforme, no intervalo $0 \leq x \leq 1$, em que a f.d.p. é $f(x) = 1$, a f.g.m. dessa distribuição é:

$$M_X(t) = \int_0^1 e^{t.X}.dx$$

INDO MAIS FUNDO!

Sabemos que a derivada e a integral de e^x é ela mesma. Mas agora temos uma **constante** t no expoente ($e^{t.x}$). Como resolver?

Para isso, precisamos da **regra da cadeia**, que trata da derivada de uma função **composta**, isto é, uma função g de uma função $f(x)$, que representamos por $\boldsymbol{g[f(x)]}$ ou $\boldsymbol{g \circ f(x)}$.

A regra afirma que a derivada da função composta é a derivada da função externa, multiplicada pela derivada da função interna:

$$\frac{d(g[f(x)])}{dx} = \frac{d(g[f(x)])}{d[f(x)]} \times \frac{d[f(x)]}{dx}$$

A função $e^{t.x}$ tem $f(x) = t.x$ como função interna e $e^{f(x)}$ como função externa:

$$g[f(x)] = e^{f(x)}$$

A derivada de e^x é ela mesma, então a derivada da função externa é ela mesma:

$$\frac{d(g[f(x)])}{d[f(x)]} = e^{f(x)} = e^{t.x}$$

Agora, derivamos a função interna:

$$\frac{df(x)}{dx} = \frac{d(tx)}{dx} = t$$

Então, a derivada de $e^{t.x}$ é o produto:

$$\frac{d(e^{t.x})}{dx} = t.e^{t.x}$$

Mas precisamos integrar a f.g.m., não derivar!

A integral fará a operação **contrária**: em vez de multiplicarmos por t dividimos por t:

$$F(x) = \int e^{t.x}\,dx = \frac{e^{t.x}}{t}$$

Integrando a nossa f.g.m., $M_X(t) = \int_0^1 e^{t.X}.dx$, obtemos a seguinte função[2]:

$$M_X(t) = \left[\frac{e^{t.x}}{t}\right]_0^1 = \frac{e^{t.1}}{t} - \frac{e^{t.0}}{t} = \frac{e^t - 1}{t}$$

E como calculamos os momentos a partir da f.g.m.?

O k-ésimo momento da variável X corresponde à **k-ésima derivada da f.g.m.**, no ponto **t = 0**:

$$\boxed{\boldsymbol{M_X^{(k)}(0) = E(X^k)}}$$

A expressão da k-ésima derivada no ponto t = 0 também pode ser representada por:

$$\left.\frac{d^k M_X}{dx^k}\right]_{t=0} = E(X^k)$$

Ou seja, para calcular o primeiro momento da variável, derivamos a f.g.m. e aplicamos o ponto $t = 0$; para calcular o segundo momento, derivamos novamente a f.g.m. e aplicamos o ponto $t = 0$,...

EXEMPLIFICANDO

Anteriormente, calculamos a f.g.m. da distribuição binomial com $n = 2$ e $p = 0{,}2$:

$$M_X(t) = 0{,}64 + 0{,}32.e^t + 0{,}04.e^{2.t}$$

A **derivada** dessa função é (para o último termo, aplicamos a regra da cadeia):

$$M'_X(t) = 0{,}32.e^t + 2 \times 0{,}04.e^{2.t} = 0{,}32.e^t + 0{,}08.e^{2.t}$$

Em seguida, aplicamos **t = 0** para obter o primeiro momento:

$$M'_X(0) = 0{,}32.e^0 + 0{,}08.e^{2.0} = 0{,}32 + 0{,}08 = 0{,}40$$

De fato, a esperança dessa distribuição é $E(X) = n \times p = 2 \times 0{,}2 = 0{,}4$!

[2] Quando integramos uma função $f(x)$ em um intervalo (a, b), isto é, quando resolvemos $\int_a^b f(x).dx$, calculamos o integrando (isto é, o resultado da integral), $F(x)$, sem nos preocuparmos com os limites do intervalo nesse momento. Essa etapa pode ser representada como $[F(x)]_a^b$.

Em seguida, aplicamos os limites, calculando $F(b) - F(a)$.

E, para obter o segundo momento, **derivamos** a f.g.m. novamente, isto é, derivamos a função $M_X'(t) = 0{,}32.e^{t} + 0{,}08.e^{2.t}$:

$$M_X''(t) = 0{,}32.e^{t} + 2 \times 0{,}08.e^{2.t} = 0{,}32.e^{t} + 0{,}16.e^{2.t}$$

Em seguida, aplicamos **t = 0**:

$$M_X''(0) = 0{,}32.e^{0} + 0{,}16.e^{2.0} = 0{,}32 + 0{,}16 = 0{,}48$$

Sabemos que a variância é a diferença entre o 2º momento e o quadrado do 1º momento:

$$V(X) = M_X''(0) - [M_X'(0)]^2 = 0{,}48 - [0{,}4]^2 = 0{,}48 - 0{,}16 = 0{,}32$$

De fato, a variância da distribuição binomial é $V(X) = n \times p \times q = 2 \times 0{,}2 \times 0{,}8 = 0{,}32$!

INDO MAIS FUNDO!

Depois de aprender a regra da cadeia, podemos entender por que o k-ésimo momento da variável corresponde à k-ésima derivada da f.g.m., no ponto t = 0.

Sabemos que a f.g.m. é definida como:

$$M_X(t) = E(e^{t.X})$$

A sua primeira derivada em relação a t é (lembre-se da regra da cadeia):

$$M_X{}'(t) = \frac{d}{dt}E(e^{t.X}) = E(X.e^{t.X})$$

No ponto t = 0, temos:

$$M_X'(0) = E(X.e^{0.X}) = E(X)$$

Que é justamente o primeiro momento! Derivando a f.g.m. novamente (segunda derivada):

$$M_X''(t) = \frac{d}{dt}E(X.e^{t.X}) = E(X.X.e^{t.X}) = E(X^2.e^{t.X})$$

Ou seja, derivando a f.g.m. k vezes, temos:

$$M_X^{(k)}(t) = E(X^k.e^{t.X})$$

No ponto t = 0, temos:

$$M_X^{(k)}(0) = E(X^k.e^{0.X}) = E(X^k)$$

(CESPE/2012 – TJ-RO – *Adaptada*) X é uma variável aleatória cuja função geradora de momentos é $\psi(t)$. Considerando que {X_1, X_2, ..., X_{10}} é uma sequência de cópias independentes de X, julgue o item a seguir.

A covariância entre duas variáveis distintas X_i e X_k, em que i ≠ k , é dada por

$$cov(X_i, X_k) = \left.\frac{d^2\psi(t)}{dt^2}\right]_{t=0} - \left(\left.\frac{d\psi(t)}{dt}\right]_{t=0}\right)^2$$

Comentários:

Sendo $\psi(t)$ a f.g.m. de uma variável, a **segunda** derivada de $\psi(t)$ aplicada no ponto $t = 0$ fornece o **segundo** momento da variável:

$$\left.\frac{d^2\psi(t)}{dt^2}\right]_{t=0} = E(X^2)$$

E a **primeira** derivada de $\psi(t)$ aplicada no ponto $t = 0$ fornece o **primeiro** momento (esperança):

$$\left.\frac{d\psi(t)}{dt}\right]_{t=0} = E(X)$$

Logo, a expressão que o item forneceu corresponde a:

$$E(X^2) - [E(X)]^2$$

Essa expressão corresponde à **variância** da variável, que é igual à covariância da mesma variável:

$$Cov(X, X) = Var(X)$$

Como o enunciado deixou claro que as variáveis são distintas, i ≠ k, então o item está errado.

Resposta: Errado.

(FCC/2012 – TRT/PE) Se a função geratriz de momentos da variável aleatória X é dada por $M(t) = \left(\frac{1}{1-2t}\right)^6, t < \frac{1}{2}$, então a média da variável aleatória Y = 0,5X – 6 é igual a

a) 2.

b) 1

c) 0,5

d) 0

e) -1

Comentários:

A f.g.m. pode ser reescrita como:

$$M(t) = (1 - 2t)^{-6}$$

Para calcular a esperança (primeiro momento) dessa variável, derivamos a f.g.m. (considerando a regra da cadeia):

$$M'(t) = -6 \times (1-2t)^{-7} \times (-2) = 12 \times (1-2t)^{-7}$$

No ponto t = 0, temos:

$$E(X) = M'(0) = 12 \times (1-2.0)^{-7} = 12$$

Sendo Y = 0,5X – 6 e considerando as propriedades da esperança, temos:

$$E(Y) = E(0{,}5.X - 6) = 0{,}5.E(X) - 6 = 0{,}5 \times 12 - 6 = 6 - 6 = 0$$

Gabarito: D

Propriedades da F.G.M

As funções geradoras de momento seguem as seguintes propriedades, em que X e Y são variáveis aleatórias e a e b são constantes reais quaisquer:

i. Sendo $Y = aX + b$, a f.g.m. de Y é $\boldsymbol{M_Y(t) = e^{b.t} \times M_X(a.t)}$

Ou seja, na f.g.m. de X, substituímos t por $a.t$ e multiplicamos o resultado por $e^{b.t}$.

Vamos supor que X siga uma distribuição uniforme no intervalo $0 \leq X \leq 1$. Já calculamos a f.g.m. de X:

$$M_X(t) = \frac{e^t - 1}{t}$$

Supondo $Y = 2.X + 3$, então a f.g.m. de Y será:

$$M_Y(t) = e^{3.t} \times M_X(2.t) = e^{3.t} \times \frac{e^{(2.t)} - 1}{(2.t)}$$

Podemos obter esse resultado a partir da definição da f.g.m.:

$$M_X(t) = E(e^{t.X})$$

Sendo $Y = aX + b$, temos:

$$M_Y(t) = E(e^{t.Y}) = E\left(e^{t.(aX+b)}\right) = E(e^{a.t.X+b.t})$$

Pela propriedade da função exponencial, temos $a^{x+y} = a^x \times a^y$, logo:

$$M_Y(t) = E(e^{a.t.X+b.t}) = E(e^{a.t.X} \times e^{b.t})$$

Pela propriedade da esperança, temos $E(k.X) = k.E(X)$, logo:

$$M_Y(t) = E(e^{a.t.X} \times e^{b.t}) = e^{b.t} \times E(e^{a.t.X})$$

Aplicando, novamente, a definição da f.g.m., temos $E(e^{a.t.X}) = M_X(a.t)$, logo:

$$M_Y(t) = e^{b.t} \times M_X(a.t)$$

ii. A f.g.m. da soma $\boldsymbol{X+Y}$, sendo X e Y **independentes**, é $\boldsymbol{M_{X+Y}(t) = M_X(t) \times M_Y(t)}$

A f.g.m. da **soma** de duas variáveis independentes é o **produto** das funções geradoras.

Além da f.g.m. que acabamos de ver, $M_X(t) = \frac{e^t - 1}{t}$, vamos supor que Y seja uma variável com distribuição contínua uniforme no intervalo (0,2), cuja f.g.m. é[3]:

$$M_Y(t) = \frac{e^{2.t} - 1}{2.t}$$

Assim, a f.g.m. da soma $X + Y$, sendo X e Y independentes, é:

$$M_{X+Y}(t) = \left(\frac{e^t}{t} - \frac{1}{t}\right) \times \left(\frac{e^{2.t}}{2.t} - \frac{1}{2.t}\right)$$

Porém, podemos aplicar essa propriedade somente para variáveis **independentes**!

[3] Calculamos essa f.g.m. da mesma forma que vimos anteriormente. A função densidade é $f(y) = \frac{1}{2}$, logo:

$$M_Y(t) = E(e^{t.Y}) = \int_0^2 e^{t.Y}.f(y).dy = \int_0^2 e^{t.Y}.\frac{1}{2}.dy = \frac{1}{2} \times \frac{e^{2.t} - e^{2.0}}{t} = \frac{e^{2.t} - 1}{2.t}$$

INDO MAIS FUNDO!

Podemos obter esse resultado da propriedade a partir da definição da f.g.m.:

$$M_X(t) = E(e^{t.X})$$

Sendo $S = X + Y$, temos:

$$M_S(t) = E(e^{t.S}) = E\left(e^{t.(X+Y)}\right) = E(e^{t.X+t.Y})$$

Pela propriedade da função exponencial, temos $a^{x+y} = a^x \times a^y$, logo:

$$M_S(t) = E(e^{t.X+t.Y}) = E(e^{t.X} \times e^{t.Y})$$

Pela propriedade multiplicativa da esperança, válida para variáveis **independentes**, temos $E(X \times Y) = E(X) \times E(Y)$, logo:

$$M_S(t) = E(e^{t.X} \times e^{t.Y}) = E(e^{t.X}) \times E(e^{t.Y})$$

Aplicando, novamente, a definição da f.g.m., temos:

$$M_S(t) = M_X(t) \times M_Y(t)$$

Essa propriedade pode ser aplicada para qualquer número de variáveis, desde que todas sejam independentes. Para $S = X_1 + X_2 + X_3 + ... + X_n$, sendo todas as variáveis **independentes**, temos:

$$M_S(t) = M_{X_i}(t) \times M_{X_2}(t) \times ... \times M_{X_n}(t)$$

Esse produto pode ser indicado por $\prod_{i=1}^{n} M_{X_i}(t)$.

iii. A f.g.m. determina **completamente** a distribuição de probabilidades.

Ou seja, se as f.g.m´s. de duas variáveis forem iguais, isto é, se $\boldsymbol{M_X(t) = M_Y(t)}, \forall \boldsymbol{t}$ (leia-se, para todo t), então podemos concluir que X e Y têm a **mesma distribuição de probabilidade**.

ESQUEMATIZANDO

Propriedades da F.G.M.

i) $M_{a.X+b}(t) = e^{b.t} \times M_X(a.t)$

ii) Para X e Y **independentes**: $M_{X+Y}(t) = M_X(t) \times M_Y(t)$

iii) **Se $M_X(t) = M_Y(t)\ \forall t$ então X = Y**

HORA DE PRATICAR!

(FGV/2019 – DPE-RJ – *Adaptada*) Se X é uma variável aleatória contínua com função densidade de probabilidade representada por $f_x(x)$, considere a função dada por: $M_X(t) = \int_{-\infty}^{\infty} e^{xt}.f_X(x)dx$.

Então julgue os seguintes itens:

I – Se Y = a.X + b, então $M_Y(t) = e^{aX}.M_X(t)$;

II – Se $Z = \sum X_i$ é uma soma das variáveis aleatórias, então:

$$M_Z(t) = M_{X1}(t).M_{X2}(t).\dots.M_{Xn}(t) = \prod M_{Xi}(t)$$

III – Se $M_X(t) = M_Y(t), \forall t$, então X e Y são independentes.

Comentários:

Em relação ao item I, vimos que, sendo Y = a.X + b, a f.g.m. de Y é:

$$M_Y(t) = e^{b.t} \times M_X(a.t)$$

Que é diferente do que consta no item I, logo, ele está errado.

Em relação ao item II, sendo Z a **soma** de variáveis **independentes**, então a f.g.m. de Z é o **produto** das funções geradoras.

Porém, o enunciado **não** informou que as variáveis são independentes, então não podemos concluir que $M_Z(t) = M_{X1}(t).M_{X2}(t).\dots.M_{Xn}(t) = \prod M_{Xi}(t)$.

Por isso, o item II está errado.

Em relação ao item III, se $M_X(t) = M_Y(t), \forall t$, então concluímos que X e Y têm a **mesma** distribuição de probabilidade, mas nada podemos concluir a respeito da sua **independência**.

Logo, o item III está errado.

Resposta: todos os itens errados.

Funções geradoras de distribuições teóricas

As distribuições teóricas ou especiais, tanto discretas, quanto contínuas, possuem funções geradoras de momentos **conhecidas**.

De fato, elas podem ser calculadas a partir das definições que acabamos de ver, mas vale conhecer algumas delas para que seja possível resolver questões dessa matéria nas provas, **sem calcular integrais e derivadas**.

Distribuição Binomial

A f.g.m. de uma distribuição binomial com parâmetros $\boldsymbol{n}$ e $\boldsymbol{p}$, sendo $\boldsymbol{q} = 1 - p$, é dada por:

$$\boldsymbol{M_{X_{Bi}}(t) = (p.e^{t} + q)^{n}}$$

Vamos verificar essa fórmula. Anteriormente, havíamos calculado a f.g.m. da distribuição binomial com $n = 2$ e $p = 0{,}2$ (logo, $q = 0{,}8$) como $M_X(t) = 0{,}64 + 0{,}32.e^{t} + 0{,}04.e^{2.t}$. Agora, vamos aplicar a fórmula acima para comparar os resultados:

$$M_X(t) = (p.e^{t} + q)^{n} = (0{,}2.e^{t} + 0{,}8)^{2}$$

Sabendo que $(a + b)^2 = a^2 + b^2 + 2.a.b$, temos[4]:

$$M_X(t) = (0{,}2.e^{t})^2 + (0{,}8)^2 + 2 \times (0{,}2.e^{t}) \times (0{,}8)$$

$$M_X(t) = 0{,}04.e^{t^2} + 0{,}64 + 0{,}32.e^{t}$$

Pela propriedade da função exponencial, temos $e^{t^2} = e^{2.t}$, logo:

$$M_X(t) = 0{,}04.e^{2.t} + 0{,}64 + 0{,}32.e^{t}$$

Que é exatamente o resultado que obtivemos anteriormente com a aplicação da fórmula!

---

[4] Para obter esse resultado (que chamamos de produto notável), basta aplicar a distributiva da multiplicação:

$$(a + b)^2 = (a + b) \times (a + b) = a \times a + a \times b + b \times a + b \times b$$

$$(a + b)^2 = a^2 + 2 \times a \times b + b^2$$

INDO MAIS FUNDO!

Podemos obter esse resultado para uma variável **binomial** qualquer, utilizando a definição da f.g.m. para uma variável discreta:

$$M_X(t) = E(e^{t.X}) = \sum_i e^{t.x_i}.P(X = x_i)$$

Para uma variável binomial, a função de probabilidade é, para $k = 0,1, \dots, n$:

$$P(X = k) = \binom{n}{k}.p^k.q^{n-k}$$

Então, a f.g.m. é:

$$M_X(t) = \sum_{k=0}^{n} e^{t.k}.\binom{n}{k}.p^k.q^{n-k} = \sum_{k=0}^{n} \binom{n}{k}.(e^t.p)^k.q^{n-k}$$

Agora, utilizamos o resultado do binômio de Newton:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k}.a^k.b^{n-k}$$

Então, fazendo $a = e^t.p$ e $b = q$, temos:

$$\sum_{k=0}^{n} \binom{n}{k}.(e^t.p)^k.q^{n-k} = (e^t.p + q)^n$$

Que é a f.g.m. que vimos para uma distribuição binomial.

Observe que, na f.g.m. da variável binomial, podemos localizar os parâmetros $\boldsymbol{n}$ e $\boldsymbol{p}$ da distribuição. Ou seja, podemos percorrer o caminho inverso, localizando os parâmetros da distribuição a partir da f.g.m. e, em seguida, calcular qualquer informação a respeito da distribuição.

Suponha, por exemplo, a seguinte f.g.m.:

$$M_X(t) = (0{,}4.e^t + 0{,}6)^5$$

Com base nessa expressão, podemos concluir que se trata de uma variável com distribuição binomial, com parâmetros $\boldsymbol{n = 5}$ e $\boldsymbol{p = 0,4}$, logo $\boldsymbol{q = 0,6}$. Para isso, é importante notar que é o parâmetro $\boldsymbol{p}$ que aparece **multiplicado por $\boldsymbol{e^t}$**, enquanto $\boldsymbol{q}$ aparece **sem** a multiplicação.

A partir dos parâmetros da distribuição, podemos calcular a média, a variância e a probabilidade $P(X = 2)$, por exemplo:

$$E(X) = n \times p = 5 \times 0{,}4 = 2$$

$$V(X) = n \times p \times q = 5 \times 0{,}4 \times 0{,}6 = 1{,}2$$

$$P(X = k) = \binom{n}{k} \times p^k \times q^{n-k}$$

$$P(X = 2) = \binom{5}{2} \times 0{,}4^2 \times 0{,}6^3 = \frac{5!}{(5-2)! \times 2!} \times 0{,}16 \times 0{,}216 = \frac{5 \times 4}{2} \times 0{,}03456 = 0{,}3456$$

Pontue-se que, para $\boldsymbol{n = 1}$, temos a f.g.m. da **Distribuição de Bernoulli**, com parâmetro $\boldsymbol{p}$:

$$\boldsymbol{M_{X_{Ber}}(t) = p.e^t + q}$$

(CESPE/2013 – ANTT) Considere que a função geradora de momentos de uma variável aleatória discreta X seja dada pela relação $M_X(q) = (0{,}8 + 0{,}2e^q)^2$, em que $q \in R$. Com base nessas informações, julgue o item a seguir.

A variância de X é igual a 0,32.

Comentários:

Sabendo que a f.g.m. de uma variável binomial é:

$$M_X(t) = (p.e^t + q)^n$$

Então, comparando com a f.g.m. fornecida, $M_X(q) = (0{,}8 + 0{,}2e^q)^2$, podemos observar que:

$$n = 2$$

$$p = 0{,}2$$

$$q = 1 - p = 0{,}8$$

A variância é, portanto:

$$V(X) = n \times p \times q = 2 \times 0{,}2 \times 0{,}8 = 0{,}32$$

Gabarito: Certo.

(FCC/2015 – TER/RR) Sabe-se que a função geratriz de momentos da variável aleatória X é dada por $[0{,}1e^t + 0{,}9]^{12}$. Nestas condições, a variância da variável aleatória Y = −2X + 3 é igual a

a) 4,32

b) 7,24

c) 2,16

d) 5,10

e) 4,56

Comentários:

Sabendo que a f.g.m. de uma variável binomial é:

$$M_X(t) = (p.e^t + q)^n$$

Então, comparando com a f.g.m. fornecida, $[0{,}1e^t + 0{,}9]^{12}$, podemos observar que:

$$n = 12$$

$$p = 0{,}1$$

$$q = 1 - p = 0{,}9$$

A variância de X é, portanto:

$$V(X) = n \times p \times q = 12 \times 0{,}1 \times 0{,}9 = 1{,}08$$

Pelas propriedades da variância, podemos calcular a variância de Y:

$$V(Y) = V(-2X + 3) = 4.V(X) = 4 \times 1{,}08 = 4{,}32$$

Gabarito: A

Distribuição Geométrica

A **f.g.m. da distribuição geométrica**, com parâmetro $\boldsymbol{p}$, sendo $\boldsymbol{q} = 1 - p$, é:

$$\boldsymbol{M_{X_{Geo}}(t) = \frac{p.e^t}{1 - q.e^t}}$$

Por exemplo, sendo X uma variável geométrica, com probabilidade de sucesso **p = 0,2**, logo a probabilidade de fracasso é **q = 0,8**, a f.g.m. é dada por:

$$M_X(t) = \frac{0{,}2.e^t}{1 - 0{,}8.e^t}$$

Assim, a partir da f.g.m. de uma distribuição geométrica, também podemos localizar o seu parâmetro.

Suponha, por exemplo, a seguinte f.g.m.:

$$M_X(t) = \frac{0{,}3.e^t}{1 - 0{,}7.e^t}$$

A partir dela, sabemos que o parâmetro da distribuição é $\boldsymbol{p = 0,3}$ e $\boldsymbol{q = 0,7}$. Para essa distribuição, é importante notar que o parâmetro $\boldsymbol{p}$ consta no **numerador** e $\boldsymbol{q}$ no denominador.

Assim, a média de lançamentos realizados até o 1º sucesso, a variância da distribuição e a probabilidade de obter o primeiro sucesso em 3 lançamentos, por exemplo, são:

$$E(X) = \frac{1}{p} = \frac{1}{0{,}3} \cong 3{,}33$$

$$V(X) = \frac{q}{p^2} = \frac{0{,}7}{(0{,}3)^2} \cong 7{,}78$$

$$P(X = k) = q^{k-1} \times p$$

$$P(X = 3) = q^2 \times p = (0{,}7)^2 \times 0{,}3 = 0{,}49 \times 0{,}3 = 0{,}147$$

(FGV/2022 – TJDFT) Suponha um processo de Bernoulli com probabilidade de sucesso de cada prova p, sendo p > 0. Seja X o número de tentativas realizadas até o primeiro sucesso (inclusive). Se 0 ≤ p ≤ 1, a função geradora de momentos de X é:

a) $1 + e^t p$

b) $e^t(1 - p)$

c) $(1 - p) + e^t p$

d) $e^{t^2/2}$

e) $\frac{pe^t}{1-(1-p).e^t}$

Comentários:

Se X representa o número de tentativas até o primeiro sucesso, então X segue distribuição geométrica, cuja f.g.m. é dada por:

$$M_{X_{Geo}}(t) = \frac{p.e^t}{1 - q.e^t}$$

Sabendo que $q = 1 - p$, então:

$$M_{X_{Geo}}(t) = \frac{p.e^t}{1 - (1 - p).e^t}$$

Gabarito: E

(FCC/2014 – TRT/PB) A função geratriz de momentos da variável aleatória X é dada por $M_x(t) = pe^t / (1 - qe^t)$, onde p é o parâmetro do modelo, p + q = 1 e 0 < qe^t < 1. Seja a variável aleatória Y = X - 1. A esperança de Y é igual a

a) p/q

b) q

c) 1 – p

d) q/p

e) 1 – p/q

Comentários:

Pela f.g.m. fornecida, sabemos que se trata de uma distribuição geométrica. A esperança dessa distribuição é:

$$E(X) = \frac{1}{p}$$

Sendo Y = X – 1, a esperança de Y pode ser obtida pelas propriedades da esperança:

$$E(Y) = E(X - 1) = E(X) - 1 = \frac{1}{p} - 1 = \frac{1 - p}{p}$$

Sabendo que 1 – p = q, então:

$$E(Y) = \frac{q}{p}$$

Gabarito: D

Distribuição de Poisson

A **f.g.m. da distribuição de Poisson**, com parâmetro $\lambda > 0$:

$$\boldsymbol{M_{X_{Po}}(t) = e^{\lambda(e^t - 1)}}$$

A expressão do expoente na base e pode ser representada entre colchetes, após o termo "exp":

$$M_X(t) = \exp\{\lambda(e^t - 1\}$$

Por exemplo, sendo X uma distribuição de Poisson, com parâmetro $\lambda = 2$, a f.g.m. é:

$$M_X(t) = e^{2(e^t - 1)}$$

Assim, a partir da f.g.m. de uma distribuição de Poisson podemos localizar o seu parâmetro $\boldsymbol{\lambda}$ e, com base nele, calcular qualquer informação da distribuição.

Suponha, por exemplo, a seguinte f.g.m.:

$$M_X(t) = e^{5(e^t - 1)}$$

Com base nessa f.g.m., sabemos que o parâmetro da distribuição é $\boldsymbol{\lambda = 5}$. Com ele, podemos calcular a média, a variância da distribuição e a probabilidade $P(X = 1)$, por exemplo:

$$E(X) = V(X) = \lambda = 5$$

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

$$P(X = 1) = \frac{e^{-5}5^1}{1!} = 5.e^{-5}$$

Distribuição Exponencial

A **f.g.m. da distribuição exponencial**, com parâmetro λ:

$$\boldsymbol{M_{X_{Exp}}(t) = \frac{\lambda}{\lambda - t}}$$

Por exemplo, sendo X uma variável exponencial, com parâmetro $\boldsymbol{\lambda = \frac{1}{100}}$, a f.g.m. é:

$$M_X(t) = \frac{\frac{1}{100}}{\frac{1}{100} - t} = \frac{\frac{1}{100}}{\frac{1 - 100.t}{100}} = \frac{1}{1 - 100.t}$$

Essa expressão vale para $t < \lambda$, para que a expressão seja positiva.

Podemos obter esse resultado, pela definição da f.g.m. para uma variável contínua:

$$M_X(t) = E(e^{t.X}) = \int e^{t.X}.f(x).dt$$

A f.d.p. da variável exponencial é $f(x) = \lambda.e^{-\lambda.x}$, com $x > 0$:

$$M_X(t) = \int_0^{\infty} e^{t.x}.\lambda.e^{-\lambda.x}.dx = \lambda \int_0^{\infty} e^{-(\lambda-t).x}.dx$$

Pela regra da cadeia, vimos que:

$$\int e^{-(\lambda-t).x}\,dx = \frac{e^{-(\lambda-t).x}}{-(\lambda-t)}$$

Quando $x \to \infty$, $e^{-x} \to 0$ e $e^0 = 1$, então, aplicando os limites da integral, temos:

$$M_X(t) = F(x \to \infty) - F(x = 0) = \lambda\left(\frac{0}{-(\lambda-t)} - \frac{1}{-(\lambda-t)}\right) = \lambda\left(\frac{1}{(\lambda-t)}\right) = \frac{\lambda}{\lambda-t}$$

Também podemos percorrer o caminho inverso, calculando a distribuição de probabilidade, a partir da f.g.m., como fizemos com as demais variáveis. Suponha a seguinte f.g.m.:

$$M_X(t) = \frac{1}{1 - 50.t}$$

Conhecendo a f.g.m. da distribuição exponencial, vamos igualar as expressões:

$$M_X(t) = \frac{\lambda}{\lambda - t} = \frac{1}{1 - 50.t}$$

$$\lambda \times (1 - 50.t) = 1 \times (\lambda - t)$$

$$\lambda - 50.t.\lambda = \lambda - t$$

Subtraindo λ de ambos os lados da equação, temos:

$$-50.t.\lambda = -t$$

Dividindo ambos os lados por $-t$:

$$50.\lambda = 1$$

$$\lambda = \frac{1}{50} = 0{,}02$$

A partir do parâmetro da distribuição, podemos calcular a média, a variância e a probabilidade $P(X < 30)$, por exemplo:

$$E(X) = \frac{1}{\lambda} = \frac{1}{0{,}02} = 50$$

$$V(X) = \frac{1}{\lambda^2} = \frac{1}{(0{,}02)^2} = \frac{1}{0{,}0004} = 2500$$

$$P(X < x) = 1 - e^{-\lambda x}$$

$$P(X < 30) = 1 - e^{-0{,}02 \times 30} = 1 - e^{0{,}6}$$

(FGV/2019 – DPE-RJ – *Adaptada*) Se X é uma variável aleatória contínua com função densidade de probabilidade representada por $f_X(x)$, considere a função dada por:

$$M_X(t) = \int_{-\infty}^{\infty} e^{xt}.f_X(x)dx$$

Então julgue o seguinte item:

Se $M_X(t) = \frac{\lambda}{\lambda - t}$; $\lambda > 0$, então $E(X) = 2\lambda$

Comentários:

Sabemos que essa f.g.m. corresponde à distribuição exponencial. A média dessa distribuição é:

$$E(X) = \frac{1}{\lambda}$$

Resposta: Errado.

(2012 – DPE-PR) Uma variável aleatória X tem função geradora de momentos dada por:

$m_X(t) = \frac{\lambda}{\lambda - t}$, para $t < \lambda$;

O valor de $E(X^2)$ é:

a) $^1/_\lambda$

b) λ

c) $^2/_{\lambda^2}$

d) 2λ

e) λ^2

Comentários:

O segundo momento da variável pode ser calculado a partir da fórmula da variância:

$$V(X) = E(X^2) - [E(X)]^2$$
$$E(X^2) = V(X) + [E(X)]^2$$

A f.g.m. $\frac{\lambda}{\lambda - t}$ corresponde à distribuição exponencial. A média e variância dessa distribuição são:

$$E(X) = \frac{1}{\lambda}$$
$$V(X) = \frac{1}{\lambda^2}$$

Logo, o segundo momento é:

$$E(X^2) = \frac{1}{\lambda^2} + \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} + \frac{1}{\lambda^2} = \frac{2}{\lambda^2}$$

Gabarito: C.

Distribuição Normal

A **f.g.m. da distribuição normal**, com média μ e variância σ^2 é:

$$M_{X_N}(t) = e^{\frac{\sigma^2 . t^2}{2}} \times e^{\mu . t} = e^{\frac{\sigma^2 . t^2}{2} + \mu . t}$$

Essa expressão também pode ser representada como:

$$\boldsymbol{M_X(t) = \exp\left\{\frac{\sigma^2 . t^2}{2} + \mu . t\right\}}$$

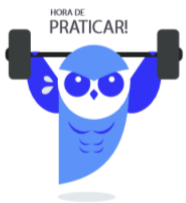
Por exemplo, para $\boldsymbol{\mu = 2}$ e $\boldsymbol{\sigma^2 = 9}$, a f.g.m. é:

$$\boldsymbol{M_X(t) = e^{\frac{9.t^2}{2} + 2.t}}$$

Para encontrar os parâmetros dessa distribuição, a partir da f.g.m., note que a média $\boldsymbol{\mu}$ multiplica $\boldsymbol{t}$ e a variância $\boldsymbol{\sigma^2}$ multiplica $\boldsymbol{t^2}$.

Vale pontuar que a f.g.m. da **Normal Padrão**, em que $\mu = 0$ e $\sigma^2 = 1$, é:

$$\boldsymbol{M_X(t) = e^{\frac{t^2}{2}}}$$



(FGV/2019 – DPE-RJ – *Adaptada*) Se X é uma variável aleatória contínua com função densidade de probabilidade representada por $f_x(x)$, considere a função dada por:

$$M_X(t) = \int_{-\infty}^{\infty} e^{xt} . f_X(x)dx$$

Então julgue o seguinte item:

Se $M_X(t) = e^{\mu t} . e^{(\sigma t)^2/2}$, então $Var(X) = \sigma^2$;

Comentários:

Sabemos que essa f.g.m. corresponde à distribuição **normal**, com média μ e variância σ^2.

Resposta: Certo.

(CESPE/2012 – TJ-RO – *Adaptada*) Considerando que X é uma variável aleatória cuja função geradora de momentos é dada por $\psi(t) = exp\left[\mu t + \frac{(\sigma t)^2}{2}\right]$, em que μ e σ são, respectivamente, a média e o desvio padrão de X, julgue o seguinte item.

A distribuição de X é assimétrica.

Comentários:

Podemos observar que a f.g.m. fornecida no enunciado se refere a uma distribuição normal. Sabemos que distribuições normais são **simétricas**.

Resposta: Errado.

(FCC/2015 – Analista CNMP – *Adaptada*) Sejam X_1, X_2, ..., X_n variáveis aleatórias independentes e identicamente distribuídas com distribuição normal padrão e $Y = X_1 + X_2 + \cdots + X_n$. Com base nessas informações, julgue a seguinte afirmativa:

A função geratriz de momentos de Y, quando n = 2, é m(t) = e^{2t}.

Comentários:

O enunciado definiu Y como a soma de n variáveis independentes com distribuição **normal padrão**, logo, Y também apresenta distribuição normal padrão. Para n = 2, temos $Y = X_1 + X_2$ com a seguinte média e variância:

$$\mu = \mu_1 + \mu_2 = 0$$

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 = 2$$

Vimos que a f.g.m. da distribuição normal é dada por:

$$M_X(t) = e^{\sigma^2 t^2/2 + \mu . t}$$

Para $\mu = 0$ e $\sigma^2 = 2$, temos:

$$M_X(t) = e^{t^2}$$

Logo, a afirmativa está incorreta.

Resposta: Errado.

Distribuição Qui-quadrado

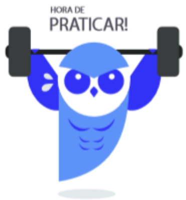
A **f.g.m. da distribuição qui-quadrado**, com $\boldsymbol{k}$ graus de liberdade é:

$$\boldsymbol{M_{X_{Qui}}(t) = \left(\frac{1}{1 - 2.t}\right)^{\frac{k}{2}}}$$

Por exemplo, para $\boldsymbol{k = 8}$, isto é, sendo X uma variável com distribuição qui-quadrado equivalente à soma de 8 variáveis normais padrão, elevadas ao quadrado, $X = \sum_{i=1}^{8} Z_i^2$, então a sua f.g.m. é:

$$M_X(t) = \left(\frac{1}{1 - 2.t}\right)^4$$

Também podemos fazer o caminho inverso, ou seja, a partir da f.g.m., definir o número de graus de liberdade, calculando o **dobro do expoente**.



(CESPE/2012 – TJ-RO – *Adaptada*) X é uma variável aleatória cuja função geradora de momentos é dada por $\psi(t) = exp\left[\mu t + \frac{(\sigma t)^2}{2}\right]$, em que μ e σ são, respectivamente, a média e o desvio padrão de X.

Considerando que $\{X_1, X_2, ..., X_{10}\}$ é uma sequência de cópias independentes de X, julgue o seguinte item.

A função geradora de momentos da soma de quadrados $\sum_{i=1}^{10} X_i^2$ é dada pela expressão $\psi_Z(t) = \left(\frac{1}{1-2t}\right)^5$

Comentários:

A f.g.m. fornecida no item, $\psi_Z(t) = \left(\frac{1}{1-2t}\right)^5$, realmente corresponde à distribuição qui-quadrado com o seguinte número de graus de liberdade:

$$\boldsymbol{k = 2 \times 5 = 10}$$

Ou seja, essa distribuição qui-quadrado corresponde à soma de $\boldsymbol{k = 10}$ variáveis com distribuição **normal padrão**, elevadas ao quadrado.

Pela f.g.m. fornecida no enunciado $\psi(t) = exp\left[\mu t + \frac{(\sigma t)^2}{2}\right]$, sabemos que se trata de variável normal com **média $\boldsymbol{\mu}$** e **variância $\boldsymbol{\sigma^2}$**, ou seja, uma variável normal **diferente** da normal padrão.

Assim, para formar uma distribuição qui-quadrado precisamos primeiro transformá-la em uma normal padrão, para depois elevá-la ao quadrado:

$$\sum_{i=1}^{k} Z_i^2 = \sum_{i=1}^{k} \left(\frac{X_i - \mu}{\sigma}\right)^2$$

Logo, a soma de quadrados $\sum_{i=1}^{10} X_i^2$, sem a transformação para a normal padrão, **não** segue uma distribuição qui-quadrado e, portanto, não apresenta a f.g.m. dessa distribuição $\left(\frac{1}{1-2t}\right)^5$.

Resposta: Errado.

F.G.Ms. para Distribuições Teóricas

Distribuição Binomial: $M_{X_{Bi}}(t) = (p.e^t + q)^n$

Distribuição Geométrica: $M_{X_{Geo}}(t) = \frac{p.e^t}{1-q.e^t}$

Distribuição de Poisson: $M_{X_{Po}}(t) = e^{\lambda(e^t-1)}$

Distribuição Exponencial: $M_{X_{Exp}}(t) = \frac{\lambda}{\lambda-t}$

Distribuição Normal: $M_{X_N}(t) = e^{\frac{\sigma^2.t^2}{2}+\mu.t}$

Distribuição Qui-Quadrado: $M_{X_{Qui}}(t) = \left(\frac{1}{1-2.t}\right)^{\frac{k}{2}}$

DISTRIBUIÇÕES CONJUNTAS – VARIÁVEIS DISCRETAS

As distribuições de probabilidade **conjuntas** apresentam os valores de probabilidade para duas (ou mais) variáveis **simultaneamente**.

Quando estudamos conjuntamente 2 variáveis, dizemos que se trata de uma **distribuição bidimensional** ou de uma **variável bidimensional**. Quando houver mais variáveis, utilizamos o termo **multidimensional**.

No caso **discreto**, essas distribuições são denotadas por $\boldsymbol{P(X = x, Y = y)}$ ou $\boldsymbol{p(x, y)}$, que correspondem à probabilidade de $X = x$ **e** de $Y = y$ (**interseção**):

$$\boxed{p(x,y) = P(X = x, Y = y) = P(X = x \cap Y = y)}$$

Essas probabilidades são normalmente representadas por uma tabela.

Vamos construir uma tabela de distribuição conjunta.

Considerando 10 bolas numeradas de 1 a 10, vamos supor que o **sucesso** da variável **X** (isto é, X = 1) corresponda à seleção de bolas com números **múltiplos de 3** (e que as demais bolas estejam associadas ao fracasso de X, em que X = 0); e que o **sucesso de Y** (Y = 1) corresponda à seleção de bolas com números **primos** (e as demais, ao fracasso de Y, em que Y = 0).

Nesse caso, temos os seguintes valores para X e Y, para cada bola numerada:

Bola	X	Y
1	0	0
2	0	1
3	1	1
4	0	0
5	0	1
6	1	0
7	0	1
8	0	0
9	1	0
10	0	0

Considerando que a probabilidade de selecionar uma bola é a mesma para todas as bolas, $p = \frac{1}{10}$, então a distribuição conjunta de X e Y para esse exemplo é:

$$P(X = 0, Y = 0) = P[Bola\text{: } 1, 4, 8 \; ou \; 10] = \frac{4}{10} = 40\%$$

$$P(X=0,Y=1)=P[Bola\!:2,5\ ou\ 7]=\frac{3}{10}=30\%$$

$$P(X=1,Y=0)=P[Bola\!:6\ ou\ 9]=\frac{2}{10}=20\%$$

$$P(X=1,Y=1)=P[Bola\!:3]=\frac{1}{10}=10\%$$

Assim, a **tabela de distribuição conjunta** para esse exemplo é:

		Y	
		0	**1**
X	**0**	40%	30%
	1	20%	10%

De maneira geral, a probabilidade conjunta é calculada pela mesma fórmula da **interseção**:

$$P(X=x,Y=y)=P(X=x|Y=y)\times P(Y=y)$$

Vale ressaltar que $P(X=x|Y=y)$ é a probabilidade **condicional** de $X=x$ (evento a **posteriori**), **dado** que ocorreu $Y=y$ (evento a **priori**).

Por exemplo, supondo que a probabilidade de um evento $Y=2$ seja $P(Y=2)=0{,}5$ e que a probabilidade de um evento $X=1$, dado que o evento $Y=2$ ocorreu seja $P(X=1|Y=2)=0{,}6$, então a probabilidade de **ambos** os eventos ocorrerem, $P(X=1,Y=2)$, é:

$$P(X=1,Y=2)=0{,}6\times 0{,}5=0{,}3$$

Se o evento a priori for a variável X, a probabilidade conjunta $P(X=x,Y=y)$ é calculada como:

$$P(X=x,Y=y)=P(Y=y|X=x)\times P(X=x)$$

Para esse exemplo que acabamos de ver, se soubermos que a probabilidade (não condicionada) de $X=1$ for $P(X=1)=0{,}9$, podemos calcular a probabilidade de $Y=2$ condicionada a $X=1$, $P(Y=2|X=1)$:

$$P(X=x,Y=y)=P(Y=y|X=x)\times P(X=x)$$

$$0{,}3=P(Y=2|X=1)\times 0{,}9$$

$$P(Y=2|X=1)=\frac{0{,}3}{0{,}9}=\frac{1}{3}$$

Se os eventos $X = x$ e $Y = y$ forem **independentes**, então a probabilidade da **interseção** é o **produto**:

$$P(X = x, Y = y) = P(X = x) \times P(Y = y)$$

Para distribuições conjuntas, a probabilidade associada a todo o Espaço Amostral também é de 100%, ou seja, a **soma** de **todos os campos** é necessariamente **100% = 1**. Vale lembrar, ainda, que todos os valores de probabilidades são **não negativos**: $0 \leq p \leq 1$.

Por exemplo, se não soubéssemos o valor de $P(X = 1, Y = 1)$ na tabela referente às bolas numeradas, como indicado abaixo, poderíamos calcular essa probabilidade, pela soma de todos os campos:

		Y	
		0	**1**
X	**0**	40%	30%
	1	20%	k

$$P(U) = 40\% + 30\% + 20\% + k = 100\%$$

$$90\% + k = 100\%$$

$$k = 10\%$$

Mesmo quando a distribuição não estiver representada por uma tabela, a soma de todas as probabilidades conjuntas $p(x, y)$, para **todos os valores possíveis** de x e **todos os valores possíveis** de y, será igual a **1**.

Podemos definir, ainda, a **Função de Distribuição Acumulada Conjunta** para as variáveis X e Y:

$$\boxed{F(x, y) = P(X \leq x, Y \leq y)}$$

Assim, a função de distribuição acumulada conjunta no ponto (x, y) apresenta a probabilidade de $X \leq x$ **e** de $Y \leq y$ (interseção).

Para o exemplo das bolas numeradas, temos os seguintes valores da função de distribuição conjunta:

$$F(0,0) = P(X \leq 0, Y \leq 0) = P(X = 0, Y = 0) = 40\%$$

$$F(1,0) = P(X \leq 1, Y \leq 0) = P(X = 0, Y = 0) + P(X = 1, Y = 0) = 40\% + 20\% = 60\%$$

$$F(0,1) = P(X \leq 0, Y \leq 1) = P(X = 0, Y = 0) + P(X = 0, Y = 1) = 40\% + 30\% = 70\%$$

$$F(1,1) = P(X \leq 0, Y \leq 1) = P(X = 0, Y = 0) + P(X = 1, Y = 0) + P(X = 0, Y = 1) + P(X = 1, Y = 1)$$

$$F(1,1) = 40\% + 20\% + 30\% + 10\% = 100\%$$

A função de distribuição acumulada pode ser calculada para todos os valores reais, mesmo que a variável aleatória não apresente um valor de probabilidade para o ponto. Logo, na fórmula acima, x e y podem ser **quaisquer** números reais.

Para o nosso exemplo podemos calcular, por exemplo, o valor da função nos seguintes pontos:

$$F(-1; -0{,}5) = 0$$

$$F(1{,}5; 0{,}5) = F(1; 0) = 60\%$$

$$F(1{,}5; 3) = F(1; 1) = 100\%$$

Pontue-se que toda f.d.a. conjunta apresenta as mesmas características da f.d.a. para uma variável, isto é, $F(x, y) = 0$ para $x \to -\infty$ e $y \to -\infty$ e $F(x, y) = 1$ para $x \to +\infty$ e $y \to +\infty$.

Probabilidades Marginais e Independência

A partir da distribuição conjunta, podemos calcular as probabilidades das variáveis **individualmente**, chamadas de **probabilidades marginais**.

Ou seja, a partir da tabela do nosso exemplo, podemos calcular as probabilidades P(X = 0), P(X = 1), P(Y = 0) e P(Y = 1), isoladamente. Para isso, devemos **somar** as probabilidades de **cada linha** e de **cada coluna**.

Abaixo, replicamos a tabela anterior, acrescentando os campos com os respectivos somatórios:

		Y		**P(X=x)**
		0	**1**	
X	**0**	40%	30%	70%
	1	20%	10%	30%
P(Y=y)		60%	40%	100%

Ou seja, as probabilidades marginais são P(X = 0) = 70%, P(X = 1) = 30%, P(Y = 0) = 60% e P(Y = 1) = 40%.

E quando não houver tabela?

De maneira geral, as probabilidades marginais são calculadas pelos seguintes somatórios:

$$P(X = x) = \sum_j p(x, y_j) \qquad e \qquad P(Y = y) = \sum_i p(x_i, y)$$

Ou seja, para obtermos a probabilidade marginal P(**X = 0**), por exemplo, precisamos **somar** todas probabilidades conjuntas da forma P(**X = 0**, Y = y), para **todos os possíveis valores de Y**, no caso:

P(**X = 0**) = P(**X = 0**, Y = 0) + P(**X = 0**, Y = 1)

P(**X = 0**) = 40% + 30% = 70%

Similarmente, para obtermos a probabilidade marginal P(**Y = 1**), precisamos **somar** todas as probabilidades conjuntas da forma P(X = x, **Y = 1**), para todos os possíveis valores de X, no caso:

P(**Y = 1**) = P(X = 0, **Y = 1**) + P(X = 1, **Y = 1**)

P(**Y = 1**) = 30% + 10% = 40%

As probabilidades marginais são essenciais para sabermos se duas variáveis são **independentes** ou não.

Para que duas variáveis X e Y sejam **independentes**, é necessário que as probabilidades conjuntas sejam iguais ao **produto** das probabilidades marginais, para **todos** os possíveis valores de X e Y:

$$P(X = x, Y = y) = P(X = x) \times P(Y = y), \ \forall (x, y)$$

Na tabela da distribuição conjunta, isso pode ser verificado observando-se as **proporções**. Para variáveis **independentes**, as **proporções** de todos os campos seguem as proporções **marginais** (totais) das variáveis. Em outras palavras, a probabilidade de campo corresponde ao **produto** das probabilidades marginais (totais) da linha e da coluna correspondentes.

Para o nosso exemplo das bolas numeradas, caso as variáveis fossem independentes, as probabilidades dos campos seriam:

		Y		**P(X=x)**
		0	**1**	
X	**0**	70% x 60% = 42%	70% x 40% = 28%	70%
	1	30% x 60% = 18%	30% x 40% = 12%	30%
P(Y=y)		60%	40%	100%

Porém, como a tabela de distribuição conjunta é **diferente** desta, concluímos que as variáveis **não** são independentes.

(2020 – UEPA – *Adaptada*) Considere o quadro abaixo, representando a distribuição conjunta de X e Y.

X↓ Y→	1	2	3	Total
1	0,08	0,12	0,2	0,4
2	0,06	0,09	0,15	0,3
3	0,06	0,09	0,15	0,3
Total	0,2	0,3	0,5	1

Julgue a seguinte afirmação:

X e Y são independentes.

Comentários:

Em atenção ao item I, X e Y são independentes se:

$$P(X = x, Y = y) = P(X = x) \times P(Y = y)$$

Para X = 1 e Y = 1, temos $P(X = 1, Y = 1) = 0{,}08$ e $P(X = 1) \times P(Y = 1) = 0{,}4 \times 0{,}2 = 0{,}08$

Para X = 2 e Y = 1, temos $P(X = 2, Y = 1) = 0{,}06$ e $P(X = 2) \times P(Y = 1) = 0{,}3 \times 0{,}2 = 0{,}06$

Para X = 3 e Y = 1, temos os mesmos valores correspondentes a X = 2 e Y = 1.

Para X = 1 e Y = 2, temos $P(X = 1, Y = 2) = 0{,}12$ e $P(X = 1) \times P(Y = 2) = 0{,}4 \times 0{,}3 = 0{,}12$

Para X = 2 e Y = 2, temos $P(X = 2, Y = 2) = 0{,}09$ e $P(X = 2) \times P(Y = 2) = 0{,}3 \times 0{,}3 = 0{,}09$

Para X = 3 e Y = 2, temos os mesmos valores correspondentes a X = 2 e Y = 2.

Como as proporções se mantiveram para Y = 1 e Y = 2, então elas serão mantidas para Y = 3.

Portanto, as variáveis X e Y são independentes.

Resposta: Certo.

(FGV/2019 – DPE/RJ – *Adaptada*) Seja a distribuição de probabilidade conjunta de variáveis aleatórias discretas conforme abaixo, onde k1 e k2 são probabilidades inicialmente desconhecidas.

(X , Y)		X			
		-1	0	1	2
Y	1	0.14	0.08	0.16	k1
	2	0.03	0.12	k2	0.25

Sendo assim, julgue os itens a seguir:

I – Para que os eventos X = 0 e Y = 2 sejam independentes, é necessário que k1 = 0,02.

II – Para que os eventos X = -1 e Y = 1 sejam independentes, é necessário que k2 = 0,15.

Comentários:

Em atenção ao item I, para que os eventos X = 0 e Y = 2 sejam independentes, então é necessário que:

$$P(X=0,Y=2)=P(X=0)\times P(Y=2)$$

Sabemos que a probabilidade marginal $P(X=0)$ é a soma das linhas da 2ª coluna:

$$P(X=0)=0{,}08+0{,}12=0{,}20$$

E que a probabilidade marginal $P(Y=2)$ é a soma das colunas da 2ª linha:

$$P(Y=2)=0{,}03+0{,}12+k_2+0{,}25=0{,}4+k_2$$

Considerando que a probabilidade conjunta é $P(X=0,Y=2)=0{,}12$, conforme tabela, então para que esses eventos sejam independentes precisamos que:

$$0{,}12=0{,}20\times(0{,}4+k_2)$$

$$0{,}12=0{,}08+0{,}2k_2$$

$$0{,}2k_2=0{,}04$$

$$k_2=0{,}2$$

Sabendo que a soma de todos os campos é igual a 100%, então:

$$\boldsymbol{P(U)}=0{,}14+0{,}03+0{,}08+0{,}12+0{,}16+k_2+k_1+0{,}25=\mathbf{1}$$

$$0{,}78+k_2+k_1=1$$

Considerando que $k_2=0{,}2$, logo:

$$P(U)=0{,}78+0{,}2+k_1=1$$

$$k_1=0{,}02$$

Resposta: Item I correto.

Em atenção ao item II, para que os eventos X = -1 e Y = 1 sejam independentes, então é necessário que:

$$P(X=-1,Y=1)=P(X=-1)\times P(Y=1)$$

Sabemos que a probabilidade marginal $P(X=-1)$ é a soma das linhas da 1ª coluna:

$$P(X=-1)=0{,}14+0{,}03=0{,}17$$

E que a probabilidade marginal $P(Y=1)$ é a soma das colunas da 1ª linha:

$$P(Y=1)=0{,}14+0{,}08+0{,}16+k_1=0{,}38+k_1$$

Considerando que a probabilidade conjunta é $P(X=-1,Y=1)=0{,}14$, conforme tabela, então para que esses eventos sejam independentes precisamos que:

$$0{,}14=0{,}17\times(0{,}38+k_1)$$

$$0{,}14=0{,}0646+0{,}17k_1$$

$$0{,}0754=0{,}17k_1$$

$$k_1\cong 0{,}444$$

Sabendo que a soma de todos os campos é igual a 100%, então:

$$P(U) = 0{,}78 + k_2 + k_1 = 1$$

$$k_2 + k_1 = 0{,}22$$

Ou seja, se tivéssemos $k_1 \cong 0{,}444$, teríamos um valor negativo para k_2. Como todas as probabilidades devem ser não negativos, então não podemos ter $k_1 \cong 0{,}444$. Logo, os eventos P(X=-1) e P(Y=1) não podem ser independentes.

Resposta: Item II errado.

Cálculo das Medidas de Distribuição

A partir das probabilidades marginais, é possível calcular as diversas medidas relacionadas a distribuições de probabilidades (esperança, variância, covariância, coeficiente de correlação).

Sabemos que, para calcular a **esperança**, multiplicamos os valores da variável pelas respectivas probabilidades (marginais):

$$\mu = E(X) = \sum_i x_i . p(x_i)$$

E vimos que as probabilidades marginais são calculadas pela **soma** das probabilidades conjuntas:

$$p(x_i) = \sum_j p(x_i, y_j)$$

Vamos considerar o nosso exemplo para calcular a esperança das variáveis.

		Y		**P(X=x)**
		0	**1**	
X	**0**	40%	30%	70%
	1	20%	10%	30%
P(Y=y)		60%	40%	100%

A esperança de X é dada por:

$$E(X) = \sum_i x_i . p(x_i)$$

$$E(X) = 0 \times P(X = 0) + 1 \times P(X = 1) = 0 \times 70\% + 1 \times 30\% = 0{,}3$$

Analogamente, a esperança de Y é dada por:

$$\mu_Y = E(Y) = \sum_j y_j . p(y_j)$$

$$E(Y) = 0 \times P(Y = 0) + 1 \times P(Y = 1) = 0 \times 60\% + 1 \times 40\% = 0{,}4$$

Também é possível calcular a esperança matemática, **sem** calcular previamente as **probabilidades marginais**, por meio das seguintes fórmulas:

$$E(X) = \sum_i \sum_j x_i . p(x_i, y_j) \qquad e \qquad E(Y) = \sum_j \sum_i y_j . p(x_i, y_j)$$

Ou seja, para encontrar $E(X)$, podemos multiplicar os valores de x_i pelas probabilidades **conjuntas** e, em seguida, **somar** os produtos, como indicado abaixo, para o nosso exemplo:

$$E(X) = x_1 \times p(x_1, y_1) + x_1 \times p(x_1, y_2) + x_2 \times p(x_2, y_1) + x_2 \times p(x_2, y_2)$$

$$E(X) = 0 \times 40\% + 0 \times 30\% + 1 \times 20\% + 1 \times 10\% = 0{,}3$$

Em relação à **variância**, temos:

$$\sigma^2 = V(X) = E(X^2) - [E(X)]^2$$

Vamos calcular o segundo momento $E(X^2)$ para esse exemplo, a partir das **probabilidades marginais**:

$$E(X^2) = \sum_i (x_i)^2 . p(x_i)$$

$$E(X^2) = 0^2 \times P(X = 0) + 1^2 \times P(X = 1) = 0^2 \times 70\% + 1^2 \times 30\% = 0{,}3$$

Logo, a variância é:

$$V(X) = E(X^2) - [E(X)]^2 = 0{,}3 - (0{,}3)^2 = 0{,}3 - 0{,}09 = 0{,}21$$

Analogamente, para Y, temos:

$$E(Y^2) = \sum_j (y_j)^2 . p(y_j)$$

$$E(Y^2) = 0^2 \times P(Y = 0) + 1^2 \times P(Y = 1) = 0^2 \times 60\% + 1^2 \times 40\% = 0{,}4$$

Assim, a variância de Y é dada por:

$$V(Y) = E(Y^2) - [E(y)]^2 = 0{,}4 - (0{,}4)^2 = 0{,}4 - 0{,}16 = 0{,}24$$

Também podemos calcular o segundo momento **sem** calcular previamente as **probabilidades marginais**, por meio das seguintes fórmulas:

$$E(X^2) = \sum_i \sum_j (x_i)^2 . p(x_i, y_j) \qquad e \qquad E(Y^2) = \sum_j \sum_i (y_j)^2 . p(x_i, y_j)$$

Ou seja, para encontrar $E(X^2)$, podemos multiplicar os valores $X = x_i$ elevados ao quadrado, pelas probabilidades **conjuntas** e, em seguida, **somar** os produtos:

$$E(X^2) = (x_1)^2 \times p(x_1, y_1) + (x_1)^2 \times p(x_1, y_2) + (x_2)^2 \times p(x_2, y_1) + (x_2)^2 \times p(x_2, y_2)$$

$$E(X^2) = 0^2 \times 40\% + 0^2 \times 30\% + 1^2 \times 20\% + 1^2 \times 10\% = 0{,}3$$

Ademais, sabemos que a **covariância** entre as variáveis é dada por:

$$Cov(X,Y) = E(X.Y) - E(X).E(Y)$$

Em que:

$$E(X.Y) = \sum_i \sum_j x_i.y_j.p(x_i,y_j)$$

Essa fórmula **não** utiliza as probabilidades **marginais**, mas sim as **conjuntas**. Para esse exemplo, temos:

$$E(X.Y) = x_1 \times y_1 \times p(x_1,y_1) + x_1 \times y_2 \times p(x_1,y_2) + x_2 \times y_1 \times p(x_2,y_1) + x_2 \times y_2 \times p(x_2,y_2)$$

$$E(X.Y) = 0 \times 0 \times P(X=0,Y=0) + 0 \times 1 \times P(X=0,Y=1) + 1 \times 0 \times P(X=1,Y=0) + 1 \times 1 \times P(X=1,Y=1)$$

$$E(X.Y) = 0 \times 0 \times 40\% + 0 \times 1 \times 30\% + 1 \times 0 \times 20\% + 1 \times 1 \times 10\% = 0{,}1$$

Existe outra forma de calcular $E(X.Y)$, construindo uma tabela com os valores de **X.Y** e as respectivas **probabilidades**, obtidas por meio das probabilidades **marginais**.

Para esse exemplo, os possíveis valores de X.Y são 0 ou 1. A probabilidade de X.Y = 0 equivale à probabilidade X = 0 ou Y = 0:

$$P(X.Y=0) = P(X=0 \cup Y=0) = P(X=0) + P(Y=0) - P(X=0,Y=0)$$

$$P(XY=0) = 70\% + 60\% - 40\% = 90\%$$

E a probabilidade de XY = 1 é dada por:

$$P(XY=1) = P(X=1,Y=1) = 10\%$$

Esses valores estão representados na tabela abaixo:

XY	0	1
P(XY)	90%	10%

O valor esperado de XY é, portanto, o produto de XY pela respectiva probabilidade:

$$E(XY) = \sum x.y.p(x.y)$$

$$E(XY) = 0 \times 90\% + 1 \times 10\% = 0{,}1$$

Assim, a covariância é dada por:

$$Cov(X,Y) = E(X.Y) - E(X).E(Y) = 0{,}1 - 0{,}3 \times 0{,}4 = 0{,}1 - 0{,}12 = -0{,}02$$

Podemos indicar as variâncias e as covariâncias entre as variáveis, em uma distribuição multidimensional, utilizando uma matriz, chamada matriz de covariâncias (Σ).

Para uma n variáveis (ou uma variável n-dimensional), a matriz de covariância terá dimensão $n \times n$. Cada campo a_{ii} da diagonal principal representa a variância da variável i; e os demais campos a_{ij} representam as covariâncias entre a variável i e a variável j. Portanto, a matriz de covariância é simétrica, isto é, os valores acima da diagonal principal são simétricos aos valores abaixo da diagonal principal.

Para o nosso exemplo, temos $V(X) = 0{,}21$, $V(Y) = 0{,}24$ e $Cov(X,Y) = -0{,}02$. Assim, a nossa matriz de covariâncias seria:

$$\Sigma = \begin{bmatrix} 0{,}21 & -0{,}02 \\ -0{,}02 & 0{,}4 \end{bmatrix}$$

Além da esperança do produto das variáveis, $E(XY)$, podemos calcular a variância do produto, dada por:

$$Var(XY) = E[(XY)^2] - [E(XY)]^2$$

Para variáveis independentes, temos $E[(XY)^2] = E(X^2).E(Y^2)$ e $E(XY) = E(X).E(Y)$.

Por fim, podemos calcular o **coeficiente de correlação**, dado por:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X.\sigma_Y}$$

Sendo σ_X o desvio padrão de X: $\sigma_X = \sqrt{V(X)}$; e σ_Y o desvio padrão de Y: $\sigma_Y = \sqrt{V(Y)}$.

Para esse exemplo, temos $Cov(X,Y) = -0{,}02$, $\sigma_X = \sqrt{V(X)} = \sqrt{0{,}21}$ e $\sigma_Y = \sqrt{V(Y)} = \sqrt{0{,}24}$, logo:

$$\rho(X,Y) = \frac{-0{,}02}{\sqrt{0{,}21}.\sqrt{0{,}24}} \cong -0{,}089$$

Vale ressaltar as seguintes propriedades para essas medidas.

Propriedades da Esperança

i) $\boldsymbol{E(k.X) = k.E(X)}$
ii) $\boldsymbol{E(X+k) = E(X)+k}$
iii) $\boldsymbol{E(X+Y) = E(X)+E(Y)}$
iv) $\boldsymbol{E(k) = k}$
v) **Se X e Y são independentes, então $\boldsymbol{E(X.Y) = E(X).E(Y)}$**

Propriedades da Variância

i) $\boldsymbol{V(X+k) = V(X)}$
ii) $\boldsymbol{V(k.X) = k^2.V(X)}$
iii) **Se X e Y são independentes, então $\boldsymbol{V(X+Y) = V(X)+V(Y)}$**

Propriedades da Covariância e da Correlação

i) $\boldsymbol{Cov(X,Y) = Cov(Y,X),\ \rho(X,Y) = \rho(Y,X)}$
ii) $\boldsymbol{Cov(X,X) = Var(X),\ \rho(X,X) = 1}$
iii) $\boldsymbol{Cov(k,X) = 0,\ \rho(k,X) = 0}$
iv) $\boldsymbol{Cov(X+Y,Z) = Cov(X,Z)+Cov(Y,Z)}$
v) $\boldsymbol{Cov(k.X,l.Y) = Cov(l.X,k.Y) = k.l.Cov(X,Y)}$

(2020 – UEPA – *Adaptada*) Considere o quadro abaixo, representando a distribuição conjunta de X e Y.

X↓ Y→	1	2	3	Total
1	0,08	0,12	0,2	0,4
2	0,06	0,09	0,15	0,3
3	0,06	0,09	0,15	0,3
Total	0,2	0,3	0,5	1

Julgue a seguinte afirmação:

E(X)=1,9 e E(Y)=2,3.

Comentários:

Sabemos que a esperança é calculada como:

$$E(X) = \sum_i x_i.p(x_i)$$

Sabendo que os possíveis valores de X são X = 1, X = 2 e X = 3, a esperança de X é dada por:

$$E(X) = 1 \times P(X = 1) + 2 \times P(X = 2) + 3 \times P(X = 3)$$

As probabilidades $P(X = 1)$, $P(X = 2)$ e $P(X = 3)$ são as probabilidades marginais (coluna Total), logo

$$E(X) = 1 \times 0{,}4 + 2 \times 0{,}3 + 3 \times 0{,}3 = 0{,}4 + 0{,}6 + 0{,}9 = 1{,}9$$

Sabendo que os possíveis valores de Y são Y = 1, Y = 2 e Y = 3, a esperança de Y é dada por:

$$E(Y) = 1 \times P(Y = 1) + 2 \times P(Y = 2) + 3 \times P(Y = 3)$$

As probabilidades $P(Y = 1)$, $P(Y = 2)$ e $P(Y = 3)$ são as probabilidades marginais (linha Total), logo:

$$E(Y) = 1 \times 0{,}2 + 2 \times 0{,}3 + 3 \times 0{,}5 = 0{,}2 + 0{,}6 + 1{,}5 = 2{,}3$$

Resposta: Certo.

(FGV/2019 – DPE-RJ – *Adaptada*) Seja a distribuição de probabilidade conjunta de variáveis aleatórias discretas conforme abaixo, onde k1 e k2 são probabilidades inicialmente desconhecidas.

(X , Y)		X			
		-1	0	1	2
Y	1	0.14	0.08	0.16	k1
	2	0.03	0.12	k2	0.25

Sendo assim, julgue os itens a seguir:

I – Para que a média de X seja igual a 0,75, é necessário que k1 = 0,12 e k2 = 0,10.

II – Para que a média de Y seja igual a 1,40, é necessário que k1 = 0,15 e k2 = 0,07.

Comentários:

Em relação ao item I, podemos calcular a média de X, sem antes obter as probabilidades marginais, fazendo:

$$E(X) = \sum_i \sum_j x_i.p(x_i, y_j)$$

Sabendo que os possíveis valores de X são X = -1, X = 0, X = 1 e X = 2, devemos multiplicar esses valores pelas respectivas probabilidades conjuntas para Y = 1 e para Y = 2:

$$E(X) = -1 \times P(X = -1, Y = 1) - 1 \times P(X = -1, Y = 2) + 0 \times P(X = 0, Y = 1) + 0 \times P(X = 0, Y = 2) + 1 \times P(X = 1, Y = 1) + 1 \times P(X = 1, Y = 2) + 2 \times P(X = 2, Y = 1) + 2 \times P(X = 2, Y = 2)$$

$$E(X) = -1 \times 0{,}14 - 1 \times 0{,}03 + 0 \times 0{,}08 + 0 \times 0{,}12 + 1 \times 0{,}16 + 1 \times k_2 + 2 \times k_1 + 2 \times 0{,}25$$

$$E(X) = -0{,}17 + 0{,}66 + k_2 + 2.k_1 = 0{,}49 + k_2 + 2k_1$$

Para que a média de X seja igual a 0,75, então:

$$E(X) = 0{,}49 + k_2 + 2k_1 = 0{,}75$$

$$\boldsymbol{k_2 + 2k_1 = 0,26}$$

Sabendo que a soma de todos os campos é igual a 100%, então:

$$P(U) = 0{,}14 + 0{,}03 + 0{,}08 + 0{,}12 + 0{,}16 + k_2 + k_1 + 0{,}25 = 1$$

$$P(U) = 0{,}78 + k_2 + k_1 = 1$$

$$\boldsymbol{k_2 + k_1 = 0,22}$$

Agora resolvemos o sistema de equações. Subtraindo a segunda equação da primeira, temos:

$$k_2 + 2k_1 - k_2 - k_1 = 0{,}26 - 0{,}22$$

$$k_1 = 0{,}04$$

Substituindo esse resultado na segunda equação, $k_2 + k_1 = 0{,}22$:

$$k_2 = 0{,}22 - 0{,}04 = 0{,}18$$

Resposta: Item I errado.

Em relação ao item II, podemos calcular a média de Y por:

$$E(Y) = \sum_j \sum_i y_j.p(x_i, y_j)$$

Sabendo que os possíveis valores de Y são Y = 1 e Y = 2, devemos multiplicar esses valores pelas respectivas probabilidades conjuntas para X = -1, X = 0, X = 1 e X = 2:

$$E(Y) = 1 \times P(X = -1, Y = 1) + 1 \times P(X = 0, Y = 1) + 1 \times P(X = 1, Y = 1) + 1 \times P(X = 2, Y = 1) + 2 \times P(X = -1, Y = 2) + 2 \times P(X = 0, Y = 2) + 2 \times P(X = 1, Y = 2) + 2 \times P(X = 2, Y = 2)$$

$$E(Y) = 1 \times 0{,}14 + 1 \times 0{,}08 + 1 \times 0{,}16 + 1 \times k_1 + 2 \times 0{,}03 + 2 \times 0{,}12 + 2 \times k_2 + 2 \times 0{,}25$$

$$E(Y) = 0{,}38 + 0{,}8 + k_1 + 2k_2 = 1{,}18 + k_1 + 2k_2$$

Para que a média de Y seja igual a 1,4, então:

$$E(Y) = 1{,}18 + k_1 + 2k_2 = 1{,}4$$

$$\boldsymbol{k_1 + 2k_2 = 0,22}$$

Pela soma do Espaço Amostral, calculamos no item I a seguinte equação:

$$\boldsymbol{k_2 + k_1 = 0,22}$$

Agora resolvemos o sistema de equações. Subtraindo a segunda equação da primeira, temos:

$$k_1 + 2k_2 - k_1 - k_2 = 0{,}22 - 0{,}22$$

$$k_1 = 0$$

Substituindo esse resultado na segunda equação, $k_2 + k_1 = 0{,}22$:

$$k_2 = 0{,}22$$

Resposta: Item II errado.

(2014 – DETRAN/RO) Considere as informações para responder à questão. A tabela apresenta a distribuição conjunta de duas variáveis aleatórias X e Y.

	Y	
X	**1**	**2**
−1	0,4	0,2
2	0,1	0,3

É correto afirmar que a covariância entre X e Y é

a) 0,2.

b) 0,3.

c) 0,4.

d) 0,5.

e) 0,6.

Comentários:

A covariância é dada por $Cov(X,Y) = E(X.Y) - E(X).E(Y)$. Podemos calcular as expectâncias, por meio das probabilidades marginais:

$$P(X = -1) = 0{,}4 + 0{,}2 = 0{,}6$$

$$P(X = 2) = 0{,}1 + 0{,}3 = 0{,}4$$

$$P(Y = 1) = 0{,}4 + 0{,}1 = 0{,}5$$

$$P(Y = 2) = 0{,}2 + 0{,}3 = 0{,}5$$

Assim, temos:

$$E(X) = -1 \times P(X = -1) + 2 \times P(X = 2) = -0{,}6 + 0{,}8 = 0{,}2$$

$$E(Y) = 1 \times P(Y = 1) + 2 \times P(Y = 2) = 0{,}5 + 1 = 1{,}5$$

E a esperança $E(XY)$ pode ser calculada a partir das probabilidades conjuntas:

$$E(XY) = \sum\sum x.y.p(x,y)$$

$$E(XY) = (-1 \times 1) \times 0{,}4 + (-1 \times 2) \times 0{,}2 + (1 \times 2) \times 0{,}1 + (2 \times 2) \times 0{,}3$$

$$E(XY) = -0{,}4 - 0{,}4 + 0{,}2 + 1{,}2 = 0{,}6$$

A covariância é, então:

$$Cov(X,Y) = 0{,}6 - 0{,}2 \times 1{,}5 = 0{,}6 - 0{,}3 = 0{,}3$$

Gabarito: B

Distribuições Condicionais

A **probabilidade condicional** corresponde às chances de ocorrer um evento $Y = y$ (a posteriori), **dado** que ocorreu outro evento $X = x$ (a priori), conforme fórmula indicada a seguir.

$$P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)}$$

Sabemos que a tabela de probabilidade conjunta apresenta as probabilidades das **interseções**, $P(X = x, Y = y)$, e que, a partir dela, podemos calcular a probabilidade marginal $P(X = x)$. Ou seja, podemos calcular as probabilidades **condicionais**, a partir da tabela de distribuição **conjunta**.

Para ilustrar, vamos retomar o exemplo das bolas numeradas de 1 a 10, em que o sucesso de X representa a seleção de um múltiplo de 3, enquanto o sucesso de Y representa a seleção um número primo:

		Y		P(X=x)
		0	**1**	
X	**0**	40%	30%	70%
	1	20%	10%	30%
P(Y=y)		60%	40%	100%

Vamos calcular, por exemplo, a probabilidade $P(Y = 1|X = 0)$. Observamos que a probabilidade **conjunta** de Y = 1 e X = 0 é $P(Y = 1, X = 0) = 30\%$; e que a probabilidade **marginal** de X = 0 é $P(X = 0) = 70\%$:

$$P(Y = 1|X = 0) = \frac{P(Y = 1, X = 0)}{P(X = 0)} = \frac{30\%}{70\%} = \frac{3}{7}$$

Podemos calcular também a **função de distribuição condicional** de uma variável, dado que a outra assume determinado valor.

Com base nessa mesma tabela, podemos calcular, por exemplo, $P(Y|X = 0)$, que corresponde às probabilidades que cada valor de Y assume, dado X = 0.

Considerando que Y pode assumir os valores 0 ou 1 e que já calculamos $P(Y = 1|X = 0) = \frac{3}{7}$, resta calcular $P(Y = 0|X = 0)$:

$$P(Y = 0|X = 0) = \frac{P(Y = 0, X = 0)}{P(X = 0)} = \frac{40\%}{70\%} = \frac{4}{7}$$

Assim, a função de distribuição condicional de Y, dado X = 0 pode ser representada como:

$$P(Y|X=0) = \begin{cases} 4/7, & se\ y=0 \\ 3/7, & se\ y=1 \end{cases}$$

Pontue-se que, se as variáveis X e Y forem **independentes**, então a probabilidade **condicionada** é **igual** à probabilidade **não condicionada**:

$$P(Y=y|X=x) = P(Y=y)$$

Utilizando a definição da probabilidade condicional, podemos calcular a **distribuição conjunta**, multiplicando a **distribuição condicional** pela distribuição da **variável a priori**:

$$P(X=x, Y=y) = P(Y=y|X=x) \times P(X=x)$$

E para calcular a **distribuição marginal** de Y (a posteriori), somamos as probabilidades conjuntas para todos os valores de X, ou seja:

$$P(Y=y) = \sum_i P(X=x_i, Y=y) = \sum_i P(Y=y|X=x_i) \times P(X=x_i)$$

(2014 – DETRAN/RO) A tabela apresenta a distribuição conjunta de duas variáveis aleatórias X e Y.

	Y	
X	**1**	**2**
–1	0,4	0,2
2	0,1	0,3

É correto afirmar que a probabilidade de X = 2, dado Y = 1, é

a) 0,1

b) 0,2

c) 0,3

d) 0,4

e) 0,5

Comentários:

A probabilidade condicional de X = 2 dado Y = 1 é calculada por:

$$P(X = 2|Y = 1) = \frac{P(X = 2 \cap Y = 1)}{P(Y = 1)}$$

A probabilidade marginal de Y = 1 é a soma dos valores da primeira coluna:

$$P(Y = 1) = 0{,}4 + 0{,}1 = 0{,}5$$

Sabendo que $P(X = 2, Y = 1) = 0{,}1$, conforme tabela, então a probabilidade condicional é:

$$P(X = 2|Y = 1) = \frac{0{,}1}{0{,}5} = 0{,}2$$

Gabarito: B.

(FGV/2022 – TJDFT) A probabilidade de um determinado time ser classificado entre os 4 primeiros colocados na primeira fase de um campeonato é de 40%. É sabido que, se for classificado entre os 4 primeiros na primeira fase, o time tem 50% de chance de vencer o campeonato. O time não venceu o campeonato, seja esse evento representado por Y = 0.

Seja também X uma variável aleatória que assume valor 0, se o time não se classificou entre os 4 primeiros na primeira fase, e que assume valor 1, caso tenha se classificado entre os 4 primeiros.

A função de probabilidade da variável aleatória X|Y = 0 é:

a) $P(X|Y = 0) = \begin{Bmatrix} 0{,}80, & se\ x = 0 \\ 0{,}20, & se\ x = 1 \end{Bmatrix}$

b) $P(X|Y = 0) = \begin{Bmatrix} 0{,}75, & se\ x = 0 \\ 0{,}25, & se\ x = 1 \end{Bmatrix}$

c) $P(X|Y = 0) = \begin{Bmatrix} 0{,}50, & se\ x = 0 \\ 0{,}50, & se\ x = 1 \end{Bmatrix}$

d) $P(X|Y = 0) = \begin{Bmatrix} 0{,}25, & se\ x = 0 \\ 0{,}75, & se\ x = 1 \end{Bmatrix}$

e) $P(X|Y = 0) = \begin{Bmatrix} 0{,}20, & se\ x = 0 \\ 0{,}80, & se\ x = 1 \end{Bmatrix}$

Comentários:

Segundo o enunciado X representa a classificação na primeira fase, sendo X = 0 se o time **não** se classifica e X = 1 se o time se **classifica**. O enunciado informa que a probabilidade de um time se classificar entre os 4 primeiros é de 40%. Assim, temos:

- P(X = 1) = 40%; e
- P(X = 0) = 100% - P(X = 1) = 60%

Já a variável Y representa o resultado do campeonato, sendo Y = 0 se o time **não** vence e Y = 1 se o time vence o campeonato.

Vamos, então, construir a tabela da distribuição conjunta dessas variáveis, sabendo que as probabilidades marginais de X são P(X = 0) = 60% e P(X = 1) = 40%:

X\Y	Y = 0	Y = 1	Total
X = 0			60%
X = 1			40%
Total			100%

O enunciado informa que o time tem 50% de vencer, se for classificado:

$$P(Y = 1|X = 1) = 50\%$$

Consequentemente, a probabilidade de o time **não** vencer, se for classificado, é complementar (também igual a 50%):

$$P(Y = 0|X = 1) = 50\%$$

Com base nesses dados e na fórmula da probabilidade condicional, podemos calcular as probabilidades conjuntas. Primeiro em relação a $P(Y = 1|X = 1)$, temos:

$$P(Y = 1|X = 1) = \frac{P(Y = 1, X = 1)}{P(X = 1)} = 50\%$$

Sabendo que $P(X = 1) = 40\%$:

$$P(Y = 1, X = 1) = 50\% \times 40\% = 20\%$$

Similarmente, para $P(Y = 0|X = 1)$, temos:

$$P(Y = 0|X = 1) = \frac{P(Y = 0, X = 1)}{P(X = 1)} = 50\%$$

$$P(Y = 0, X = 1) = 50\% \times 40\% = 20\%$$

Vamos preencher esses dados na tabela:

X\Y	Y = 0	Y = 1	Total
X = 0			60%
X = 1	20%	20%	40%
Total			100%

Em relação aos campos de X = 0, se o time não se classificar, ele não irá vencer o campeonato. Em outras palavras, dado que o time não se classificou, a probabilidade de ele vencer o campeonato é nula:

$$P(Y = 1|X = 0) = 0$$

E a probabilidade de o time não vencer o campeonato, dado que não se classificou é igual a 100%:

$$P(Y = 0|X = 0) = 1$$

Assim como fizemos anteriormente, utilizamos essas fórmulas para calcular as probabilidades conjuntas:

$$P(Y = 1|X = 0) = \frac{P(Y = 1, X = 0)}{P(X = 0)} = 0$$

Para qualquer valor de P(X = 0), temos:

$$P(Y = 1, X = 0) = 0$$

$$P(Y = 0|X = 0) = \frac{P(Y = 0, X = 0)}{P(X = 0)} = 1$$

Sabendo que P(X = 0) = 60%, temos:

$$P(Y = 0, X = 0) = 60\%$$

X\Y	Y = 0	Y = 1	Total
X = 0	60%	0%	60%
X = 1	20%	20%	40%
Total	80%	20%	100%

Já a distribuição de probabilidade de X dado Y = 0, que a questão pede, é composta das probabilidades $P(X = 0|Y = 0)$ e $P(X = 1|Y = 0)$, que podem ser calculadas com base na tabela. Em relação a X=0, temos:

$$P(X = 0|Y = 0) = \frac{P(Y = 0, X = 0)}{P(Y = 0)} = \frac{60\%}{80\%} = 0{,}75$$

Em relação a X = 1, temos:

$$P(X = 1|Y = 0) = \frac{P(Y = 1, X = 0)}{P(Y = 0)} = \frac{20\%}{80\%} = 0{,}25$$

Gabarito: B

Esperança e Variância Condicionais

É possível calcular a esperança e as medidas de dispersão que conhecemos para a **distribuição condicional**.

Para calcular a **esperança condicional** $E(Y|X = x)$, multiplicamos os valores de y pelas **probabilidades condicionais** $P(Y|X = x)$:

$$E(Y|X = x) = \sum_j y_j . P(Y = y_j|X = x)$$

Também podemos encontrar a **variância condicional**, como indicado a seguir:

$$V(Y|X = x) = E(Y^2|X = x) - [E(Y|X = x)]^2$$

Para isso, precisamos do cálculo do **segundo momento condicional**, $E(Y^2|X = x)$, em que multiplicamos os **quadrados** dos valores de y pelas probabilidades **condicionais** correspondentes $P(Y|X = x)$:

$$E(Y^2|X = x) = \sum_j (y_j)^2 . P(Y = y_j|X = x)$$

EXEMPLIFICANDO

Com base no mesmo exemplo das bolas numeradas, vamos primeiro calcular a **esperança condicional** $E(Y|X=0)$:

$$E(Y|X=0) = 0 \times P(Y=0|X=0) + 1 \times P(Y=1|X=0)$$

$$E(Y|X=0) = P(Y=1|X=0)$$

Como vimos antes, $P(Y=1|X=0) = \frac{3}{7}$, logo, $E(Y|X=0) = \frac{3}{7}$.

Para o cálculo da **variância condicional** $V(Y|X=0) = E(Y^2|X=0) - [E(Y|X=0)]^2$, precisamos do valor de $E(Y^2|X=0)$:

$$E(Y^2|X=0) = 0^2 \times P(Y=0|X=0) + 1^2 \times P(Y=1|X=0)$$

$$E(Y^2|X=0) = P(Y=1|X=0) = \frac{3}{7}$$

Logo, a variância condicional é:

$$V(Y|X=0) = \frac{3}{7} - \left(\frac{3}{7}\right)^2 = \frac{21-9}{49} = \frac{12}{49}$$

SE LIGA!

Pontue-se que, se as variáveis forem **independentes**, então as probabilidades **condicionadas** serão iguais às probabilidades **marginais**:

$$P\left(Y^* = y_j | X^* = x\right) = P(Y^* = y_j)$$

Consequentemente, a **esperança condicional** será igual à **esperança marginal**; e a **variância condicional** será igual à **variância marginal**:

$$E(Y^*|X^* = x) = E(Y^*)$$

$$V(Y^*|X^* = x) = V(Y^*)$$

Podemos, ainda, calcular a **esperança marginal** de Y, a partir da **esperança condicional**:

$$E(Y) = E_X[E(Y|X)] = \sum_i E(Y|X = x_i) \times P(X = x_i)$$

Ou seja, somamos os produtos da esperança condicional $E(Y|X = x_i)$ pelas respectivas probabilidades, para todos os possíveis valores de x_i, como se fosse a "esperança da esperança".

Invertendo os eventos a priori e a posteriori, teríamos o somatório para todos os valores de Y:

$$E(X) = E_Y[E(X|Y)] = \sum_j E(X|Y = y_j) \times P(Y = y_j)$$

Esse cálculo é baseado na **Lei das Expectativas Iteradas**, que afirma que o valor esperado das esperanças condicionais é igual à esperança não condicional. Ela também pode ser aplicada para $E(XY)$:

$$E(XY) = E_X[E(XY|X)], \qquad E(XY) = E_Y[E(XY|Y)]$$

Como o evento a priori é uma constante, podemos indicar essas esperanças como:

$$E(XY) = E_X[E(XY|X)] = E_X[X.E(Y|X)], \qquad E(XY) = E_Y[E(XY|Y)] = E_Y[Y.E(X|Y)]$$

Para calcular **variância marginal** $V(Y) = E(Y^2) - [E(Y)]^2$, precisamos do segundo momento marginal $E(Y^2)$, que também pode ser calculado pelo segundo momento condicional, calculando a "esperança da esperança":

$$E(Y^2) = E_X[E(Y^2|X)] = \sum_i E(Y^2|X = x_i) \times P(Y = x_i)$$

EXEMPLIFICANDO

Vamos exemplificar esses cálculos com base na mesma tabela, replicada a seguir:

		Y		P(X=x)
		0	**1**	
X	**0**	40%	30%	70%
	1	20%	10%	30%
P(Y=y)		60%	40%	100%

O valor da **esperança marginal** de Y é dado por:

$$E(Y) = E(Y|X = 0) \times P(X = 0) + E(Y|X = 1) \times P(X = 1)$$

Ou seja, precisamos de $E(Y|X = 0)$, que já calculamos, e de $E(Y|X = 1)$:

$$E(Y|X = 1) = 0 \times P(Y = 0|X = 1) + 1 \times P(Y = 1|X = 1)$$

$$E(Y|X = 1) = P(Y = 1|X = 1) = \frac{P(Y=1, X=1)}{P(X=1)} = \frac{10\%}{30\%} = \frac{1}{3}$$

Sabendo que $E(Y|X = 0) = \frac{3}{7}$, o valor de $E(Y)$ é:

$$E(Y) = \frac{3}{7} \times 0{,}7 + \frac{1}{3} \times 0{,}3 = 0{,}3 + 0{,}1 = 0{,}4$$

Para calcular o valor da **variância marginal** de Y, $V(Y) = E(Y^2) - [E(Y)]^2$, precisamos do valor de $E(Y^2)$, dado por:

$$E(Y^2) = E(Y^2|X = 0) \times P(X = 0) + E(Y^2|X = 1) \times P(X = 1)$$

Ou seja, precisamos de $E(Y^2|X = 0)$, que já calculamos, e de $E(Y^2|X = 1)$:

$$E(Y^2|X = 1) = 0^2 \times P(Y = 0|X = 1) + 1^2 \times P(Y = 1|X = 1)$$

$$E(Y^2|X = 1) = P(Y = 1|X = 0)$$

Já calculamos $P(Y = 1|X = 0) = \frac{1}{3}$, logo $E(Y^2|X = 1) = \frac{1}{3}$.

Sabendo que $E(Y^2|X = 0) = \frac{3}{7}$, o valor de $E(Y^2)$ é:

$$E(Y^2) = \frac{3}{7} \times 0{,}7 + \frac{1}{3} \times 0{,}3 = 0{,}3 + 0{,}1 = 0{,}4$$

Assim, o valor da variância marginal de Y é:

$$V(Y) = 0{,}4 - (0{,}4)^2 = 0{,}4 - 0{,}16 = 0{,}24$$

Também é possível calcular a **esperança**, **sem** calcular previamente as **probabilidades marginais**, por meio das seguintes fórmulas:

$$E(X) = \sum_i \sum_j x_i . p(x_i, y_j) \qquad e \qquad E(Y) = \sum_j \sum_i y_j . p(x_i, y_j)$$

Ou seja, para encontrar $E(Y)$, podemos multiplicar os valores y pelas **probabilidades conjuntas** e, em seguida, somar os produtos para **todos os possíveis valores y**, como indicado abaixo:

$$E(Y) = 0 \times 40\% + 0 \times 20\% + 1 \times 30\% + 1 \times 10\% = 0{,}4$$

E, para calcular a **variância**, também podemos calcular o segundo momento $E(Y^2)$ **sem** calcular previamente as **probabilidades marginais**, por meio das seguintes fórmulas:

$$E(X^2) = \sum_i \sum_j (x_i)^2 . p(x_i, y_j) \qquad e \qquad E(Y^2) = \sum_j \sum_i (y_j)^2 . p(x_i, y_j)$$

Ou seja, para encontrar $E(Y^2)$, podemos multiplicar os valores y elevados ao quadrado, pelas probabilidades conjuntas e, em seguida, somar os produtos:

$$E(Y^2) = 0^2 \times 40\% + 0^2 \times 20\% + 1^2 \times 30\% + 1^2 \times 10\% = 0{,}4$$

A variância marginal também pode ser calculada pela soma da **esperança** da **variância condicionada** X|Y com a **variância** da **esperança condicionada** X|Y.

$$V(X) = E[V(X|Y)] + V[E(X|Y)]$$

(FGV/2019 – DPE/RJ – *Adaptada*) Seja a distribuição de probabilidade conjunta de variáveis aleatórias discretas conforme abaixo, onde k1 e k2 são probabilidades inicialmente desconhecidas.

(X , Y)		X			
		-1	0	1	2
Y	1	0.14	0.08	0.16	k1
	2	0.03	0.12	k2	0.25

Sendo assim, julgue o item a seguir:

Se k1 = 0,08, então a esperança condicional de Y dado X = 1, E(Y|X=1) é superior a 1,5

Comentários:

Primeiro, vamos calcular o valor de k2, considerando que a soma de todos os campos é P(U) = 1:

$$P(U) = 0{,}14 + 0{,}03 + 0{,}08 + 0{,}12 + 0{,}16 + k_2 + k_1 + 0{,}25 = 1$$

$$0{,}78 + k_2 + k_1 = 1$$

$$k_2 + k_1 = 0{,}22$$

Se $k_1 = 0{,}08$, como descrito no item, então:

$$k_2 = 0{,}22 - 0{,}08 = 0{,}14$$

Para calcular a esperança condicional, fazemos:

$$E(Y|X = x) = \sum_y y.P(Y = y|X = x)$$

Então, precisamos das probabilidades condicionais $P(Y = y|X = 1)$ para todos os valores de y. A probabilidade condicional é dada por:

$$P(Y = y|X = 1) = \frac{P(Y = y, X = 1)}{P(X = 1)}$$

Ou seja, precisamos da probabilidade marginal de X = 1, que é a soma das probabilidades da terceira coluna:

$$P(X = 1) = 0{,}16 + k2 = 0{,}16 + 0{,}14 = 0{,}3$$

Então, a probabilidade condicional $P(Y = 1|X = 1)$ é:

$$P(Y = 1|X = 1) = \frac{P(Y = 1, X = 1)}{P(X = 1)} = \frac{0{,}16}{0{,}3} = \frac{16}{3}$$

Então, a probabilidade condicional $P(Y = 2|X = 1)$ é:

$$P(Y = 2|X = 1) = \frac{P(Y = 2, X = 1)}{P(X = 1)} = \frac{0{,}14}{0{,}3} = \frac{14}{3}$$

A esperança $E(Y|X = 1)$ é a soma dessas probabilidades condicionais, multiplicadas pelos respectivos valores de Y:

$$E(Y|X = 1) = 1 \times P(Y = 1|X = 1) + 2 \times P(Y = 2|X = 1)$$

$$E(Y|X = 1) = 1 \times \frac{16}{3} + 2 \times \frac{14}{3} = \frac{16 + 28}{3} = \frac{44}{3} \cong 1{,}467$$

Como 1,467 < 1,5, o item está errado.

Resposta: Errado.

(FGV/2008 – SEN) Seja X uma variável aleatória com distribuição uniforme no intervalo (0,1) e suponha que a distribuição condicional de Y dado X = x seja binomial (n, x).

Nesse caso, a esperança e a variância da variável aleatória Y são respectivamente:

a) nx e $nx(1-x)$

b) $\frac{n}{2}$ e $\frac{n^2+2n}{12}$

c) nx e $\frac{x(1-x)}{n}$

d) $\frac{n}{2}$ e $\frac{n^2}{12}$

e) $\frac{n}{2}$ e $\frac{2n}{3}$

Comentários:

Aqui, temos a variável X, com distribuição uniforme no intervalo (0, 1); e a variável Y, dependente de X, cuja distribuição condicional Y|X segue distribuição binomial com parâmetros (n, x), em que x é a probabilidade de sucesso.

Para calcular a **esperança** marginal de Y, vamos utilizar a seguinte fórmula:

$$E(Y) = E[E(Y|X)]$$

Sabendo que Y|X segue distribuição binomial com parâmetros n e x, a esperança condicional de Y|X é:

$$E(Y|X) = n.x$$

Pela fórmula, temos que E(Y) corresponde à esperança desse resultado:

$$E(Y) = E(n.x)$$

Pelas propriedades da esperança, quando multiplicamos a variável por uma constante, a sua esperança é multiplicada por essa constante:

$$E(Y) = n.E(x)$$

Sabendo que X segue distribuição uniforme no intervalo (0, 1), a esperança de X é:

$$E(x) = \frac{a+b}{2} = \frac{0+1}{2} = \frac{1}{2}$$

Então a esperança marginal de Y é:

$$E(Y) = n.\frac{1}{2} = \frac{n}{2}$$

Para calcular a **variância** marginal de Y, vamos utilizar a seguinte fórmula:

$$V(Y) = E[V(Y|X)] + V[E(Y|X)]$$

Sabendo que Y|X segue distribuição binomial com parâmetros n e x, a variância condicional de Y|X é:

$$V(Y|X) = n.x.(1-x) = n.x - n.x^2$$

Já sabemos que a esperança condicional de Y|X é $E(Y|X) = n.x$, logo:

$$V(Y) = E[n.x - n.x^2] + V[n.x] = E[n.x] - E[n.x^2] + V[n.x]$$

Utilizando a propriedade da esperança e sabendo que quando multiplicamos uma variável por uma constante, a sua variância é multiplicada pelo quadrado dessa constante, temos:

$$V(Y) = n.E[x] - n.E[x^2] + n^2.V[x]$$

A variância de uma variável com distribuição uniforme, no intervalo (0,1) é:

$$V(X) = \frac{(b-a)^2}{12} = \frac{(1-0)^2}{12} = \frac{1}{12}$$

Para encontrar $E[x^2]$, vamos considerar a fórmula geral da variância:

$$V(X) = E[X^2] - [E(X)]^2$$

$$E[X^2] = V(X) + [E(X)]^2$$

Sabendo que $V(X) = \frac{1}{12}$ e que $E(X) = \frac{1}{2}$, então o segundo momento é:

$$E[X^2] = \frac{1}{12} + \left[\frac{1}{2}\right]^2 = \frac{1}{12} + \frac{1}{4} = \frac{1+3}{12} = \frac{4}{12}$$

Agora, conhecemos todos as medidas para calcularmos a variância marginal de Y:

$$V(Y) = n.\frac{1}{2} - n.\frac{4}{12} + n^2.\frac{1}{12} = \frac{6n - 4n + n^2}{12} = \frac{2n + n^2}{12}$$

Gabarito: B

Distribuição Multinomial

A distribuição **multinomial** é uma **generalização** da distribuição **binomial**, para quando houver **mais de dois** resultados possíveis em cada ensaio, ou seja, quando os ensaios **não** forem necessariamente de **Bernoulli**.

Por exemplo, se eu quisesse estudar a cor dos olhos de uma população, para utilizar a distribuição binomial, eu teria que dividir a população em apenas duas categorias, olhos claros e olhos escuros. Já, utilizando a distribuição multinomial, posso dividir a população em quantas categorias interessar: olhos verdes, azuis, marrons, pretos, bicolores, ...

Generalizando, na distribuição multinomial, a população é dividida em $\boldsymbol{k}$ categorias, com probabilidades $p_1, p_2, \dots, p_k$, respectivamente, cuja soma é igual a 1:

$$\sum_{i=1}^{k} p_i = 1$$

Assim como a distribuição binomial, serão realizados $\boldsymbol{N}$ **ensaios independentes**, mas os resultados desses ensaios poderão ser qualquer um dos $\boldsymbol{k}$ valores possíveis da população. O número de vezes em que obtemos, como **resultado**, cada uma das **categorias** $1, 2, \dots, k$ é representado pela respectiva variável aleatória $\boldsymbol{X_1, X_2, \dots, X_k}$, chamada **variável multinomial**.

Dessa forma, o conjunto das variáveis $X_1, X_2, \dots, X_k$ segue uma distribuição multinomial, em que a probabilidade **conjunta** de obtermos n_1 vezes a categoria 1 (isto é, $X_1 = n_1$), n_2 vezes a categoria 2 (isto é, $X_2 = n_2$), ..., e n_k vezes a categoria k (isto é, $X_k = n_k$), é dada por:

$$\boldsymbol{p(n_1, n_2, \dots, n_k) = \frac{N!}{n_1!n_2!\dots n_k!} \times (p_1)^{n_1} \times (p_2)^{n_2} \times \dots \times (p_k)^{n_k}}$$

Em que o número total de ensaios N equivale ao somatório dos números de ensaios de cada variável:

$$\sum_{i=1}^{k} n_i = N$$

Vamos primeiro entender essa fórmula. Ela envolve a probabilidade de obter $X_1 = n_1, X_2 = n_2, \dots, X_k = n_k$, **nessa ordem**, dada por:

$$(p_1)^{n_1} \times (p_2)^{n_2} \times \dots \times (p_k)^{n_k}$$

Mas como a ordem não precisa ser essa, multiplicamos esse resultado pela **permutação** de N elementos, com **repetição** de $n_1, n_2, \dots, n_k$ elementos:

$$P_N^{n_1, n_2, \dots, n_k} = \frac{N!}{n_1!\, n_2! \dots n_k!}$$

EXEMPLIFICANDO

Suponha uma caixa com **10 bolas**, das quais 4 são pretas, 3 são brancas, 2 são azuis e 1 é verde, e que vamos selecionar uma bola ao acaso, observar a sua cor e devolvê-la à caixa. Vamos repetir esse procedimento **5 vezes** e calcular a probabilidade de obter 2 vezes uma bola preta e 1 vez as demais cores.

Assim, X_1 representará o número de vezes que obtemos uma bola **preta**, X_2 representará o número de vezes que obtemos uma bola **branca**, X_3 representará o número de vezes que obtemos uma bola **azul** e X_4, o número de vezes que obtemos a bola **verde**. As respectivas probabilidades são:

$$p_1 = \frac{4}{10} = \frac{2}{5}, \quad p_2 = \frac{3}{10}, \quad p_3 = \frac{2}{10} = \frac{1}{5}, \quad p_4 = \frac{1}{10}$$

Logo, para $n_1 = 2$ e $n_2 = n_3 = n_4 = 1$, temos a seguinte probabilidade:

$$p(n_1, n_2, n_3, n_4) = \frac{N!}{n_1!n_2!n_3!n_4!} \times (p_1)^{n_1} \times (p_2)^{n_2} \times (p_3)^{n_3} \times (p_4)^{n_4}$$

$$p(2,1,1,1) = \frac{5!}{2!1!1!1!} \times \left(\frac{2}{5}\right)^2 \times \left(\frac{3}{10}\right)^1 \times \left(\frac{1}{5}\right)^1 \times \left(\frac{1}{10}\right)^1$$

$$p(2,1,1,1) = 5 \times 4 \times 3 \times \frac{4}{25} \times \frac{3}{10} \times \frac{1}{5} \times \frac{1}{10} = \frac{144}{2500} = 5{,}76\%$$

Nesse exemplo, poderíamos utilizar a distribuição **binomial** caso quiséssemos calcular a probabilidade de obter **apenas um dos resultados**, por exemplo, **apenas 2** vezes a bola preta. Nesse caso, teríamos os parâmetros $n = 2$ e $p = {}^{4}/_{10}$.

Pontue-se que a distribuição multinomial é **reduzida** à binomial para $\boldsymbol{k = 2}$. Nesse caso, a fórmula da distribuição multinomial é dada por:

$$p(n_1, n_2) = \frac{N!}{n_1!\, n_2!} \times (p_1)^{n_1} \times (p_2)^{n_2}$$

Sabendo que $\sum_{i=1}^{k} p_i = 1$ e $\sum_{i=1}^{k} n_i = N$, então nesse caso, temos: $p_2 = 1 - p_1$ e $n_2 = N - n_1$:

$$p(n_1, N - n_1) = \frac{N!}{n_1!\,(N - n_1)!} \times (p_1)^{n_1} \times (1 - p_1)^{N - n_1}$$

Essa é a distribuição binomial, com os parâmetros N e $p = p_1$, em que normalmente denotamos n_1 (o número de resultados com o atributo sucesso) por k.

O **valor esperado** e a **variância** são calculados para **cada variável** $X_i, i = 1, 2, \dots, k$, como na distribuição **binomial**, considerando os parâmetros N e p_i:

$$E(X_i) = N.p_i$$

$$V(X_i) = N.p_i(1 - p_i)$$

A **covariância** entre duas variáveis X_u e X_v é dada por:

$$Cov(X_u, X_v) = -n.p_u.p_v$$

(FGV/2019 – DPE/RJ – *Adaptada*) O nível de escolaridade dos cidadãos que necessitam recorrer à Defensoria Pública do RJ segue uma distribuição multinomial com parâmetros p1 = 0,4, p2 = 0,3, p3 = 0,2 e p4 = 0,1, que são as probabilidades de que pertençam à classe menos instruída (Cp1) até a classe mais instruída (Cp4). Com base nessa distribuição, julgue o item a seguir.

Selecionando 10 pessoas ao acaso, com reposição, a probabilidade de obter as quantidades Cp1 = 4, Cp2= 3, Cp3 = 2 e Cp4 = 1 é menor que 3%.

Comentários:

A probabilidade de obter $n_1 = 4$ pessoas da classe p1; $n_2 = 3$ pessoas da Cp2; $n_3 = 2$ pessoas da Cp3; e $n_4 = 1$ pessoas da Cp4 é dada por

$$p(n_1, n_2, n_3, n_4) = \frac{N!}{n_1!\, n_2!\, n_3!\, n_4!} \times (p_1)^{n_1} \times (p_2)^{n_2} \times (p_3)^{n_3} \times (p_4)^{n_4}$$

Considerando que o número total de pessoas na amostra é $N = 10$; e que as probabilidades das classes são $p_1 = 0{,}4$, $p_2 = 0{,}3$, $p_3 = 0{,}2$ e $p_4 = 0{,}1$, podemos calcular a probabilidade desejada:

$$p(4,3,2,1) = \frac{10!}{4!\,3!\,2!\,1!} \times (0{,}4)^4 \times (0{,}3)^3 \times (0{,}2)^2 \times (0{,}1)^1$$

$$p(4,3,2,1) = 10 \times 9 \times 4 \times 7 \times 5 \times 0{,}0256 \times 0{,}027 \times 0{,}04 \times 0{,}1 \cong 0{,}0348\ (ou\ 3{,}48\%)$$

Resposta: Errado.

(2014 – HC-UFPE – *Adaptada*) Um geneticista acredita que a cor dos olhos, classificada em três categorias (preto, marrom, verde/azul), ocorre, respectivamente, com as seguintes porcentagens: 30%, 50% e 20%. Com base nessa crença, julgue o item a seguir.

A probabilidade de selecionar ao acaso, com reposição, 3 pessoas com olhos pretos, 4 pessoas com olhos marrons e 1 pessoa com olhos verdes/azuis é inferior a 10%.

Comentários:

A probabilidade de obter $n_1 = 3$ pessoas com olhos pretos, $n_2 = 4$ pessoas com olhos marrons e $n_3 = 1$ pessoa com olhos verdes/azuis é dada por:

$$p(n_1, n_2, n_3) = \frac{N!}{n_1!\, n_2!\, n_3!} \times (p_1)^{n_1} \times (p_2)^{n_2} \times (p_3)^{n_3}$$

Considerando que o número total de pessoas na amostra é $N = 10$; e que as probabilidades das categorias de olhos são $p_1 = 0{,}3$, $p_2 = 0{,}5$ e $p_3 = 0{,}2$, podemos calcular a probabilidade desejada:

$$p(3,4,1) = \frac{8!}{4!\,3!\,1!} \times (0{,}3)^3 \times (0{,}5)^4 \times (0{,}2)^1$$

$$p(3,4,1) = 8 \times 7 \times 5 \times 0{,}027 \times 0{,}0625 \times 0{,}2 = 0{,}0945$$

Resposta: Certo.

DENSIDADE CONJUNTA – VARIÁVEIS CONTÍNUAS

As **variáveis contínuas** também podem ser analisadas **simultaneamente**, por meio da **função densidade de probabilidade conjunta**. Essa função pode ser representada por $f_{X,Y}(x, y)$ ou simplesmente $f(x, y)$, para as variáveis aleatórias X e Y.

Analogamente ao caso discreto, a função densidade conjunta é **não negativa** para todos os possíveis valores de x e y:

i) $\boldsymbol{f(x, y) \geq 0}$, para todo (x, y)

E a probabilidade de todo o **Espaço Amostral** é igual a **1**.

Como não podemos somar os possíveis valores, pois variáveis contínuas são não enumeráveis, devemos utilizar a **integral**.

Para calcular a probabilidade de todo o Espaço amostral, devemos "somar" as probabilidades para todos os possíveis valores de X (isto é, integrar a função em relação a X) e para todos os possíveis valores de Y (isto é, integrar a função em relação a Y também). Para isso, utilizamos a integral dupla:

ii) $\boldsymbol{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y) \,.\, dx . dy = 1}$

Se a função densidade apresentar valores em um intervalo $x \in (a, b)$ e $y \in (c, d)$, podemos escrever essa segunda característica como:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y) \,.\, dx . dy = \int_{c}^{d}\int_{a}^{b} f(x, y) \,.\, dx . dy = 1$$

A **integral dupla** leva em consideração 2 variáveis sendo uma **interna** e a outra **externa**, no caso, a variável Y é externa e a variável X é a interna.

$$\int_{c}^{d}\int_{a}^{b} f(x, y) \,.\, dx \,.\, dy$$

Isso significa que os limites da **primeira** integral (no caso, “c” e “d”) se referem à variável **externa** (no caso, y); e os limites da **segunda** integral (no caso, “a” e “b”) se referem à variável **interna** (no caso, x).

Quando os limites da integral interna são **constantes**, e não uma função da outra variável, você pode **inverter** a ordem das variáveis. A integral dupla indicada a seguir é equivalente:

$$\int_a^b \int_c^d f(x,y).dy.dx$$

O cálculo é feito em duas etapas. Primeiro integramos em relação à variável **interna**, como se a variável externa fosse uma **constante** e aplicamos os respectivos limites. Em seguida, integramos em relação à variável externa e aplicamos os respectivos limites.

Por exemplo, vamos supor a função densidade conjunta $f(x,y) = x + y$, definida no intervalo $x \in (0,1)$ e $y \in (0,1)$.

$$\int_0^1 \int_0^1 (x+y).dx.dy$$

Integrando em relação a x (como se y fosse uma constante):

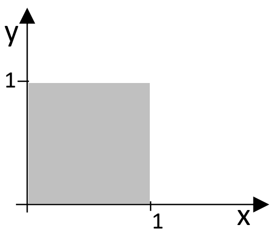
$$\int_0^1 (x+y).dx = \left[\frac{x^2}{2} + y.x\right]_0^1 = \frac{1^2-0^2}{2} + y.(1-0) = \frac{1}{2} + y$$

Agora, integramos em relação a y:

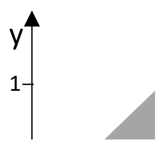
$$\int_0^1 \left(\frac{1}{2} + y\right).dy = \left[\frac{1}{2}y + \frac{y^2}{2}\right]_0^1 = \frac{1}{2}(1-0) + \frac{1^2-0^2}{2} = \frac{1}{2} + \frac{1}{2} = 1$$

Assim, confirmamos que essa função atende às duas características (ser não negativa no intervalo e ter probabilidade P(U) = 1).

Quando os limites das integrais são constantes, as variáveis estão definidas em um **retângulo** (ou **quadrado**). No exemplo anterior, as variáveis estavam definidas no espaço $x \in (0,1)$ e $y \in (0,1)$, representado a seguir:



No entanto, a região das variáveis pode assumir outros formatos, em que o espaço de uma variável **depende** da outra. Para $0 < x < y < 1$, por exemplo, temos a seguinte região:



Nessa situação, os limites da integral interna (primeira a ser resolvida) será uma **função** da outra variável, enquanto os limites da integral externa serão **fixos**.

Neste exemplo, podemos observar que x varia entre o eixo $X = 0$ e a reta $X = y$. Assim, se integrarmos primeiro em relação a x, os limites dessa variável serão $(0, y)$, enquanto os limites da variável y serão $(0, 1)$.

$$P(U) = \int_0^1 \int_0^y f(x, y) . dx . dy$$

EXEMPLIFICANDO

Vamos integrar a função $f(x, y) = 2(x + y)$ nesse espaço $0 < x < y < 1$:

$$P(U) = \int_0^1 \int_0^y 2. (x + y) . dx . dy$$

Integrando primeiro em relação a $\boldsymbol{x}$ (integral interna):

$$\int_0^y 2. (x + y) . dx = 2. \left[\frac{x^2}{2} + y. x\right]_0^y$$

Agora, substituímos x pelos limites:

$$= 2. \left[\frac{y^2 - 0^2}{2} + y. (y - 0)\right] = y^2 + 2y^2 = 3y^2$$

Integrando em relação a $\boldsymbol{y}$:

$$P(U) = \int_0^1 (3y^2) . dy = \left[3. \frac{y^3}{3}\right]_0^1 = 1^3 - 0^3 = 1$$

Considerando que $2(x + y) > 0$ para todos os valores no intervalo $0 < x < y < 1$ e que a probabilidade associada a todo o intervalo é $P(U) = 1$, confirmamos que essa é uma função **densidade** de probabilidade conjunta, por atender aos dois requisitos.

Para calcular a **probabilidade conjunta** de um **intervalo** $P(x_1 < X < x_2, y_1 < Y < y_2)$, utilizamos a integral dupla no referido intervalo:

$$P(x_1 < X < x_2, y_1 < Y < y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} f(x, y) . dx. dy$$

E a **função de distribuição acumulada conjunta**, $F(x_o, y_o) = P(X \leq x_o, Y \leq y_o)$, é dada por:

$$F(x_o, y_o) = P(X \leq x_o, Y \leq y_o) = \int_{-\infty}^{y_o} \int_{-\infty}^{x_o} f(x,y)\,.\,dx.\,dy$$

Se a função densidade assumir valores apenas em um intervalo $x \in (\boldsymbol{a}, b)$ e $y \in (\boldsymbol{c}, d)$, então a função de distribuição acumulada conjunta será:

$$F(x_o, y_o) = P(X \leq x_o, Y \leq y_o) = \int_{c}^{y_o} \int_{a}^{x_o} f(x,y)\,.\,dx.\,dy$$

EXEMPLIFICANDO

Vamos calcular a probabilidade de X estar no intervalo $\left(0, \frac{1}{2}\right)$ e Y, no intervalo $\left(0, \frac{1}{4}\right)$ para a função densidade conjunta $f(x,y) = x + y$

$$P\left(0 < X < \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2}\right) = \int_0^{1/4} \int_0^{1/2} (x+y).\,dx\,.\,dy$$

Integrando em relação a x (como se y fosse constante), temos:

$$\int_0^{1/2} (x+y).\,dx = \left[\frac{x^2}{2} + yx\right]_0^{1/2} = \frac{(1/2)^2 - 0^2}{2} + y.\left(\frac{1}{2} - 0\right) = \frac{1}{8} + \frac{1}{2}y$$

Integrando em relação a y, temos:

$$P = \int_0^{\frac{1}{4}} \left(\frac{1}{8} + \frac{1}{2}y\right).\,dy = \left[\frac{1}{8}y + \frac{1}{2} \times \frac{y^2}{2}\right]_0^{\frac{1}{4}}$$

$$P = \frac{1}{8}.\left(\frac{1}{4} - 0\right) + \frac{(1/4)^2 - 0^2}{4} = \frac{1}{32} + \frac{1}{16 \times 4} = \frac{2+1}{64} = \mathbf{\frac{3}{64}}$$

Dependendo da função densidade, é possível calcular a probabilidade **geometricamente**, em vez de utilizar a integral. Por exemplo, a função $x^2 + y^2 + z^2 \leq r^2$, sendo r uma constante, corresponde a uma **esfera** de raio r, com centro na origem (0, 0, 0). Nessa situação, as variáveis assumem valores entre $-r$ e r.

Considerando que a função densidade é necessariamente maior que zero, vamos supor que a função corresponda à parcela da esfera em que as variáveis são **positivas**, $x, y, z > 0$. Isso corresponde a $\frac{1}{8}$ da esfera, porque a restrição representa metade do intervalo para cada variável $\left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}\right)$.

Sabendo que o volume da esfera é dado por $V_{esf} = \frac{4}{3}.\pi.r^3$, então o volume dessa parcela da esfera, que corresponde à probabilidade conjunta da função $x^2 + y^2 + z^2 \leq r^2$, no intervalo $x, y, z \in (0, r)$, é igual a 1/8 do volume da esfera:

$$V_{p\_esf} = \frac{1}{8} \times \frac{4}{3}.\pi.r^3 = \frac{1}{6}\pi.r^3$$

Se essa função representar a função densidade conjunta das variáveis, esse volume precisa ser igual a 1. Nessa situação, o valor do raio pode ser calculado como:

$$P(U) = \frac{1}{6}\pi.r^3 = 1$$

$$r = \sqrt[3]{\frac{6}{\pi}}$$

(CESPE/2010 – DETRAN/ES – *Adaptada*) Com relação a distribuições conjuntas, julgue os itens que se seguem.

A função de distribuição acumulada conjunta é definida por $F(x, y) = P(X \leq x, Y \leq y)$.

Comentários:

De fato, a função de distribuição acumulada conjunta $F(x, y)$ corresponde à probabilidade de ambas as variáveis serem menores ou iguais aos respectivos valores, isto é, $X \leq x$ e $Y \leq y$.

Gabarito: Certo

Densidade Marginal e Independência

A partir da densidade conjunta, podemos calcular a **densidade marginal**.

No caso discreto, obtivemos a probabilidade marginal de **X** **somando** a probabilidade conjunta P(X = x, Y = y) para todos os valores de **Y**. Similarmente, no caso variável, **integramos** a f.d.p. conjunta em relação a **Y** para obter a f.d.p. marginal de **X**:

$$\boxed{f(x) = \int_{-\infty}^{\infty} f(x, y).dy}$$

Se Y estiver definida no intervalo $y \in (c, d)$, então a densidade marginal de X será dada por:

$$f(x) = \int_{c}^{d} f(x, y).dy$$

Para o nosso exemplo, temos $f(x, y) = x + y$, com $y \in (0,1)$, logo a função densidade marginal de X é:

$$f(x) = \int_0^1 (x+y).dy = \left[x.y + \frac{y^2}{2}\right]_0^1 = x(1-0) + \frac{1^2 - 0^2}{2} = x + \frac{1}{2}$$

Analogamente, a f.d.p. marginal de **Y** é calculada pela integral da f.d.p. conjunta em relação a **X**:

$$\boxed{f(y) = \int_{-\infty}^{\infty} f(x,y).dx}$$

Se X estiver definida no intervalo $x \in (a,b)$, então a densidade marginal de X será dada por:

$$f(y) = \int_a^b f(x,y).dx$$

Para o nosso exemplo, temos $f(x,y) = x + y$, com $x \in (0,1)$, logo a função densidade marginal de Y é:

$$f(y) = \int_0^1 (x+y).dx = \left[\frac{x^2}{2} + y.x\right]_0^1 = \frac{1^2 - 0^2}{2} + y(1-0) = \frac{1}{2} + y$$

Analogamente ao caso discreto, podemos concluir que duas variáveis são **independentes** se a função densidade **conjunta** for igual ao **produto** das densidades **marginais**:

$$\boxed{\boldsymbol{f(x,y) = f(x) \times f(y)}}$$

Para o nosso exemplo, o produto das funções marginais é:

$$f(x) \times f(y) = \left(x + \frac{1}{2}\right) \times \left(\frac{1}{2} + y\right) = \frac{1}{2}x + xy + \frac{1}{2}.\frac{1}{2} + \frac{1}{2}y = xy + \frac{1}{2}(x+y) + \frac{1}{4}$$

Como essa função é **diferente** da densidade conjunta, $f(x,y) = x + y$, podemos concluir que as variáveis **não** são **independentes**.

Podemos calcular, ainda, a função de distribuição **acumulada marginal**. A função de distribuição acumulada no ponto x_o da variável X equivale à integral da função marginal de X até o ponto x_o:

$$F(x_o) = P(X \leq x_o) = \int_{-\infty}^{x_o} f(x)\, dx$$

Acabamos de ver que a função densidade marginal de X é a integral da função densidade conjunta em relação a Y:

$$f(x) = \int_{-\infty}^{\infty} f(x,y).dy$$

Logo, para calcular a função de distribuição acumulada de X, a partir da função densidade conjunta, precisamos integrar em relação a Y e, depois, em relação a X (integral dupla):

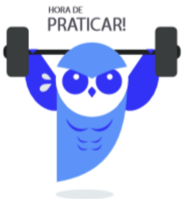
$$F(x_o) = P(X \leq x_o) = \int_{-\infty}^{x_o} \int_{-\infty}^{\infty} f(x,y).\,dy\,dx$$

$$f(x)$$

Analogamente, a função de distribuição acumulada marginal de Y pode ser calculada a partir da função densidade conjunta:

$$F(y_o) = P(Y \leq y_o) = \int_{-\infty}^{y_o} \underbrace{\int_{-\infty}^{\infty} f(x,y).\,dx}_{f(y)}.\,dy$$

Para o nosso exemplo, já calculamos as funções marginais, então precisamos integrá-las em função da mesma variável para obter as funções de distribuição acumulada.



(2019 – IBGE – *Adaptada*) Sendo X e Y duas variáveis aleatórias contínuas com função de densidade conjunta:

$$f_{X,Y}(x,y) = \begin{Bmatrix} \frac{1}{4}(2x+y), & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0, & caso\ contrário \end{Bmatrix}$$

e considerando as variáveis aleatórias X e Y e a função de densidade conjunta, julgue as afirmativas a seguir:

I - $f_X(x) = \frac{1}{2}(x+1), 0 \leq x \leq 1$

II - $f_Y(y) = \frac{1}{4}(2+y), 0 \leq y \leq 2$

III – As variáveis aleatórias X e Y são independentes.

Comentários:

Em relação à primeira alternativa, a densidade marginal de X é a integral da função densidade conjunta em relação a Y. Sabendo que $0 \leq y \leq 2$, temos:

$$f(x) = \int f(x,y).\,dy$$

$$f(x) = \int_0^2 \frac{1}{4}(2x+y).\,dy = \frac{1}{4}\left[2x.y + \frac{y^2}{2}\right]_0^2 = \frac{1}{4}\left(2x.(2-0) + \frac{2^2-0^2}{2}\right) = \frac{1}{4}\left(4x + \frac{4}{2}\right) = x + \frac{1}{2}$$

Resposta: Afirmativa I errada.

Em relação à segunda alternativa, a densidade marginal de Y é a integral da função densidade conjunta em relação a X. Sabendo que $0 \leq x \leq 1$, temos:

$$f(y) = \int f(x,y).\,dx$$

$$f(y) = \int_0^1 \frac{1}{4}(2x+y).dx = \frac{1}{4}\left[2.\frac{x^2}{2} + y.x\right]_0^1 = \frac{1}{4}[1^2 - 0^2 + y.(1-0)] = \frac{1}{4}(1+y)$$

Resposta: Afirmativa II errada.

Em relação à afirmativa III, para verificar a independência, devemos multiplicar as funções marginais:

$$f(x) \times f(y) = \left(x + \frac{1}{2}\right) \times \left(\frac{1}{4} + \frac{y}{4}\right) = \frac{x}{4} + \frac{x.y}{4} + \frac{1}{8} + \frac{y}{8} = \frac{1}{4}\left(x + x.y + \frac{y+1}{2}\right)$$

Esse resultado é diferente da função densidade conjunta, isto é:

$$f(x,y) \neq f(x) \times f(y)$$

Logo, podemos concluir que as variáveis **não** são independentes.

Resposta: Afirmativa III errada.

Cálculo das Medidas de Distribuição

Com base nos cálculos das densidades marginais, calculamos as diversas medidas relacionadas a distribuições de probabilidades que temos visto.

Sabemos que a **esperança** de uma variável contínua é:

$$E(X) = \int_{-\infty}^{\infty} x.f(x).dx$$

Também sabemos que a função **marginal** de X, $f(x)$, pode ser obtida integrando a função densidade conjunta $f(x,y)$ em relação a Y:

$$f(x) = \int_{-\infty}^{\infty} f(x,y).dy$$

Logo, a **esperança marginal** de X pode ser calculada a partir da **função densidade conjunta**:

$$\boxed{E(X) = \int_{-\infty}^{\infty} \underbrace{\int_{-\infty}^{\infty} x.f(x,y).dy}_{x.f(x)} .dx}$$

Para o nosso exemplo, em que a função densidade conjunta é $f(x,y) = x + y$, no intervalo $x \in (0,1)$ e $y \in (0,1)$, a esperança é:

$$E(X) = \int_0^1 \int_0^1 x.(x+y).dy.dx = \int_0^1 \int_0^1 (x^2 + x.y).dy.dx$$

Integrando primeiro em relação a Y (considerando x como constante), temos:

$$\int_0^1 (x^2 + x.y).dy = \left[x^2.y + x.\frac{y^2}{2}\right]_0^1 = x^2.(1-0) + x.\frac{1^2 - 0^2}{2} = x^2 + \frac{1}{2}x$$

Esse é o produto de x pela sua função marginal $f(x)$, que é uma função de x, somente. Para calcular a esperança marginal de X, integramos essa expressão em relação a X:

$$E(X) = \int_0^1 \left(x^2 + \frac{1}{2}x\right).dx = \left[\frac{x^3}{3} + \frac{1}{2}.\frac{x^2}{2}\right]_0^1 = \frac{1^3 - 0^3}{3} + \frac{1^2 - 0^2}{4} = \frac{1}{3} + \frac{1}{4} = \frac{4+3}{12} = \frac{7}{12}$$

Analogamente, a **esperança marginal** de Y pode ser calculada a partir da função densidade conjunta:

$$\boxed{E(Y) = \int_{-\infty}^{\infty} \underbrace{\int_{-\infty}^{\infty} y.f(x,y).dx}_{y.f(y)}.dy}$$

Já a **variância** é calculada como:

$$V(X) = E(X^2) - [E(X)]^2$$

A esperança $E(X)$ é calculada como vimos acima, mas precisamos do **segundo momento marginal** $E(X^2)$, calculado, a partir da função densidade marginal $f(x)$, como:

$$E(X^2) = \int_{-\infty}^{\infty} x^2.f(x).dx$$

Sabemos que a função densidade **marginal** $f(x)$ pode ser calculada a partir da densidade conjunta:

$$f(x) = \int_{-\infty}^{\infty} f(x,y).dy$$

Então, o **segundo momento marginal** é dado por:

$$\boxed{E(X^2) = \int_{-\infty}^{\infty} \underbrace{\int_{-\infty}^{\infty} x^2.f(x,y).dy}_{x^2.f(x)}.dx}$$

Para o nosso exemplo, temos:

$$E(X^2) = \int_0^1 \int_0^1 x^2.(x+y).dy.dx = \int_0^1 \int_0^1 (x^3 + x^2.y).dy.dx$$

Integrando primeiro em relação a Y (considerando x como constante), temos:

$$\int_0^1 (x^3 + x^2.y).dy = \left[x^3.y + x^2.\frac{y^2}{2}\right]_0^1 = x^3.(1-0) + x^2.\frac{1^2 - 0^2}{2} = x^3 + \frac{1}{2}x^2$$

Esse é o produto de x^2 pela sua função marginal $f(x)$, que é uma função de x, somente. Para calcular o segundo momento marginal de X, integramos essa expressão em relação a X:

$$E(X^2) = \int_0^1 \left(x^3 + \frac{1}{2}x^2\right).dx = \left[\frac{x^4}{4} + \frac{1}{2}.\frac{x^3}{3}\right]_0^1 = \frac{1^4 - 0^4}{4} + \frac{1^3 - 0^3}{6} = \frac{1}{4} + \frac{1}{6} = \frac{3+2}{12} = \frac{5}{12}$$

E a variância é:

$$V(X) = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{5}{12} - \frac{49}{144} = \frac{60-49}{144} = \frac{11}{144}$$

Em relação a Y, o **segundo momento marginal** é dado por:

$$\boxed{E(Y^2) = \int_{-\infty}^{\infty} \underbrace{\int_{-\infty}^{\infty} y^2.f(x,y).dx}_{y^2.f(y)}.dy}$$

Lembre-se de que o desvio padrão é, em qualquer situação, a **raiz quadrada** da variância.

▪

Podemos calcular, ainda, a **covariância** como:

$$Cov(X,Y) = E(X.Y) - E(X).E(Y)$$

Em que:

$$\boxed{E(X.Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x.y.f(x,y).dy.dx}$$

Para o nosso exemplo, temos:

$$E(X.Y) = \int_0^1 \int_0^1 x.y.(x+y).dy.dx = \int_0^1 \int_0^1 (x^2.y + x.y^2).dy.dx$$

Começando pela integral em relação a Y (como x fosse uma constante), temos:

$$\int_0^1 (x^2.y + x.y^2).dy = \left[x^2.\frac{y^2}{2} + x.\frac{y^3}{3}\right]_0^1 = x^2.\frac{1^2 - 0^2}{2} + x.\frac{1^3 - 0^3}{3} = \frac{1}{2}x^2 + \frac{1}{3}x$$

Agora, integrando em relação a X, temos:

$$E(X.Y) = \int_0^1 \left(\frac{1}{2}x^2 + \frac{1}{3}x\right).dx = \left[\frac{1}{2}.\frac{x^3}{3} + \frac{1}{3}.\frac{x^2}{2}\right]_0^1 = \frac{1^3 - 0^3}{6} + \frac{1^2 - 0^2}{6} = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

A covariância será calculada subtraindo-se o produto das esperanças desse resultado.

Por fim, o **coeficiente de correlação** é calculado dividindo-se a covariância pelo produto dos desvios padrão:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X.\sigma_Y}$$

(2019 – IBGE – *Adaptada*) Sendo X e Y duas variáveis aleatórias contínuas com função de densidade conjunta dada por:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{4}(2x+y), & 0 \le x \le 1, 0 \le y \le 2 \\ 0, & caso\ contrário \end{cases}$$

e considerando as variáveis aleatórias X e Y e a função de densidade conjunta, julgue a afirmativa a seguir.

O coeficiente de correlação entre X e Y é igual a 0 (zero).

Comentários:

O coeficiente de correlação é calculado como:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X.\sigma_Y}$$

E a covariância é dada por:

$$Cov(X,Y) = E(X.Y) - E(X).E(Y)$$

Para calcular as esperanças, vamos considerar as probabilidades marginais calculadas anteriormente em relação a esse mesmo enunciado $f(x) = x + \frac{1}{2}$ e $f(y) = \frac{1}{4}(1+y)$.

A esperança de X é dada por:

$$E(X) = \int_{-\infty}^{\infty} x.f(x).dx$$

$$E(X) = \int_0^1 x.\left(x + \frac{1}{2}\right).dx = \int_0^1 \left(x^2 + \frac{1}{2}x\right).dx$$

$$E(X) = \left[\frac{x^3}{3} + \frac{1}{2}.\frac{x^2}{2}\right]_0^1 = \frac{1^3 - 0^3}{3} + \frac{1^2 - 0^2}{4} = \frac{1}{3} + \frac{1}{4} = \frac{4+3}{12} = \frac{7}{12}$$

E a esperança de Y é:

$$E(Y) = \int_{-\infty}^{\infty} y.f(y).dy$$

$$E(Y) = \int_0^2 y.\frac{1}{4}(1+y).dy = \int_0^2 \left(\frac{1}{4}y + \frac{1}{4}.y^2\right).dy$$

$$E(Y) = \left[\frac{1}{4}.\frac{y^2}{2} + \frac{1}{4}.\frac{y^3}{3}\right]_0^2 = \frac{2^2-0^2}{8} + \frac{2^3-0^3}{12} = \frac{4}{8} + \frac{8}{12} = \frac{1}{2} + \frac{2}{3} = \frac{3+4}{6} = \frac{7}{6}$$

Para calcular a covariância, precisamos ainda de $E(X.Y)$:

$$E(X.Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x.y.f_{X,Y}(x,y).dy.dx$$

$$E(X.Y) = \int_0^1\int_0^2 x.y.\frac{1}{4}(2x+y)dy.dx = \int_0^1\int_0^2 \left(\frac{x^2}{2}.y + \frac{x}{4}.y^2\right).dy.dx$$

Integrando primeiro em relação a Y (como se x fosse constante), temos:

$$\int_0^2 \left(\frac{x^2}{2}.y + \frac{x}{4}.y^2\right).dy = \left[\frac{x^2}{2}.\frac{y^2}{2} + \frac{x}{4}.\frac{y^3}{3}\right]_0^2 = x^2.\frac{2^2-0^2}{4} + x.\frac{2^3-0^3}{12} = x^2.1 + x.\frac{8}{12} = x^2 + \frac{2}{3}x$$

Integrando em relação a X, temos:

$$E(X.Y) = \int_0^1 \left(x^2 + \frac{2}{3}x\right)dx = \left[\frac{x^3}{3} + \frac{2}{3}.\frac{x^2}{2}\right]_0^1 = \frac{1^3-0^3}{3} + \frac{1^2-0^2}{3} = \frac{2}{3}$$

A covariância é, então:

$$Cov(X,Y) = \frac{2}{3} - \frac{7}{12} \times \frac{7}{6} = \frac{48-49}{72} = -\frac{1}{72}$$

Como $Cov(X,Y) \neq 0$, o coeficiente de correlação não é zero. Portanto, a afirmativa está errada.

Resposta: Errado.

Nota: Embora soubéssemos que as variáveis não eram independentes, **não** poderíamos concluir que o coeficiente de correlação **não** seria nulo, pois é possível ter variáveis dependentes com covariância nula.

Por outro lado, se as variáveis fossem independentes, então necessariamente a covariância e, consequentemente, o coeficiente de correlação seriam nulos.

Densidade Condicional

Assim como fizemos para o caso discreto, podemos definir a densidade de probabilidade **condicional**, a partir da densidade conjunta:

$$f(x|Y=y) = \frac{f(x,y)}{f(Y=y)}$$

$$f(y|X = x) = \frac{f(x, y)}{f(X = x)}$$

Aqui, $Y = y$ e $X = x$ representam possíveis valores que as variáveis podem assumir.

Para o nosso exemplo, em que $f(x, y) = x + y$, calculamos $f(x) = x + \frac{1}{2}$, então $f(y|X = x)$ é:

$$f(y|X = x) = \frac{x + y}{x + \frac{1}{2}}$$

Para $X = 0$, por exemplo, temos:

$$f(y|X = 0) = \frac{0 + y}{0 + \frac{1}{2}} = 2. y$$

Lembre-se de que **se** as variáveis forem **independentes**, a função densidade de probabilidade **condicional** é igual à **marginal** para todos os possíveis valores da variável a priori:

$$f(x|Y = y) = f(x), \qquad \forall y$$

$$f(y|X = x) = f(y), \qquad \forall x$$

Utilizando a definição da densidade condicional, podemos calcular a **densidade conjunta**, multiplicando a **densidade condicional** pela densidade da **variável a priori**:

$$f(x, y) = f(y|x). f(x)$$

E para calcular a **densidade marginal** de Y (a posteriori), integramos a densidade conjunta em relação a X, ou seja:

$$f(y) = \int_{-\infty}^{\infty} f(x, y). dx = \int_{-\infty}^{\infty} f(y|x). f(x). dx$$

A partir da densidade condicional, podemos calcular a **probabilidade condicional** associada a determinado intervalo, na referida condição. Para isso, integramos a densidade condicional nesse intervalo.

$$P(a < x < b|Y = y) = \int_a^b f(x|Y = y). dx$$

$$P(c < y < d|X = x) = \int_c^d f(y|X = x)\,.\,dy$$

Utilizando a densidade condicional $f(y|X = 0) = 2.y$, para $0 < y < 1$, podemos calcular a probabilidade de Y pertencer ao intervalo $y < \frac{1}{4}$, condicionada a $X = 0$:

$$P\left(y < \frac{1}{4}|X = 0\right) = \int_0^{\frac{1}{4}} 2y\,.\,dy = \left[2.\frac{y^2}{2}\right]_0^{\frac{1}{4}} = \left(\frac{1}{4}\right)^2 - (0)^2 = \frac{1}{16}$$

Além disso, podemos calcular as medidas de distribuição para a densidade condicional. A **esperança condicional** pode ser calculada como:

$$E(X|Y = y) = \int_{-\infty}^{\infty} x.f(x|Y = y)\,.\,dx$$

$$E(Y|X = x) = \int_{-\infty}^{\infty} y.f(y|X = x)\,.\,dy$$

Para o nosso exemplo, a esperança condicional $E(Y|X = 0)$ é dada por:

$$E(Y|X = 0) = \int_0^1 y.2y\,.\,dy = \int_0^1 2.y^2\,.\,dy = \left[2.\frac{y^3}{3}\right]_0^1 = 2.\frac{1^3 - 0^3}{3} = \frac{2}{3}$$

Também podemos calcular a **variância condicional** de X:

$$V(X|Y = y) = E(X^2|Y = y) - [E(X|Y = y)]^2$$

Em que o **segundo momento condicional** de X é:

$$E(X^2|Y = y) = \int_{-\infty}^{\infty} x^2.f(x|Y = y).dx$$

Similarmente, a **variância condicional** de **Y** é:

$$V(Y|X = x) = E(Y^2|X = x) - [E(Y|X = x)]^2$$

Em que o **segundo momento condicional** de **Y** é:

$$E(Y^2|X = x) = \int_{-\infty}^{\infty} y^2.f(y|X = x).dy$$

Para o nosso exemplo, o segundo momento condicional $E(Y^2|X = 0)$ é:

$$E(Y^2|X = 0) = \int_0^1 y^2.2y\,.\,dy = \int_0^1 2.y^3\,.\,dy = \left[2.\frac{y^4}{4}\right]_0^1 = \frac{1^4 - 0^4}{2} = \frac{1}{2}$$

Logo, a variância condicional de Y, dado X = 0, é:

$$V(Y|X=0)=\frac{1}{2}-\left[\frac{2}{3}\right]^2=\frac{1}{2}-\frac{4}{9}=\frac{9-8}{18}=\frac{1}{18}$$

Assim como vimos para o caso discreto, também podemos calcular as medidas da distribuição **marginal**, a partir das medidas da distribuição condicional.

A **esperança marginal** pode ser calculada a partir da esperança condicional:

$$E(X)=E_Y[E(X|Y)]$$

E a **variância marginal** pode ser calculada a partir da esperança e da variância condicionais:

$$Var(X)=E[Var(X|Y)]+Var[E(X|Y)]$$

Vamos supor que $X|Y$ seja uma distribuição normal com média $E(X|Y)=y$ e variância $Var(X|Y)=y^2$. Portanto, a **esperança marginal** de X é dada por:

$$E(X)=E_Y[E(X|Y)]=E(Y)$$

Supondo que Y siga uma distribuição contínua uniforme no intervalo (0, 1), a média é $E(Y)=\frac{0+1}{2}=\frac{1}{2}$, logo:

$$E(X)=E(Y)=\frac{1}{2}$$

E a **variância marginal** de X é dada por:

$$Var(X)=E[Var(X|Y)]+Var[E(X|Y)]=E(Y^2)+Var(Y)$$

A variância de uma distribuição contínua uniforme no intervalo (0, 1) é $Var(Y)=\frac{(1-0)^2}{12}=\frac{1}{12}$. Ademais, o segundo momento $E(Y^2)$ pode ser calculado a partir da fórmula da variância:

$$Var(Y)=E(Y^2)-[E(Y)]^2$$

$$E(Y^2)=Var(Y)+[E(Y)]^2=\frac{1}{12}+\left(\frac{1}{2}\right)^2=\frac{1}{12}+\frac{1}{4}=\frac{1+3}{12}=\frac{1}{3}$$

Portanto, a variância marginal de X é:

$$Var(X)=E(Y^2)+Var(Y)=\frac{1}{3}+\frac{1}{12}=\frac{4+1}{12}=\frac{5}{12}$$

(2019 – IBGE – *Adaptada*) Sendo X e Y duas variáveis aleatórias contínuas com função de densidade conjunta dada por:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{4}(2x+y), & 0 \le x \le 1, 0 \le y \le 2 \\ 0, & caso\ contrário \end{cases}$$

e considerando as variáveis aleatórias X e Y e a função de densidade conjunta, julgue:

$$E(X|Y=y) = \frac{4+3y}{6(1+y)}$$

Comentários:

A esperança condicionada é dada por:

$$E(X|Y=y) = \int_{-\infty}^{\infty} x.f(x|Y=y).dx$$

Ou seja, precisamos da densidade condicional, $f(x|Y=y)$, dada por:

$$f(x|Y=y) = \frac{f(x,y)}{f(Y=y)}$$

Sabendo que $f(y) = \frac{1}{4}(1+y)$, como já calculamos, temos:

$$f(x|Y=y) = \frac{f(x,y)}{f(Y=y)} = \frac{\frac{1}{4}(2x+y)}{\frac{1}{4}(1+y)} = \frac{(2x+y)}{(1+y)}$$

Agora, podemos calcular a esperança condicionada:

$$E(X|Y=y) = \int_0^1 x.\frac{(2x+y)}{(1+y)}.dx = \int_0^1 \left[\frac{2}{(1+y)}x^2 + \frac{y}{(1+y)}x\right].dx$$

$$E(X|Y=y) = \left[\frac{2}{(1+y)}.\frac{x^3}{3} + \frac{y}{(1+y)}.\frac{x^2}{2}\right]_0^1 = \frac{2}{3(1+y)}(1^3 - 0^3) + \frac{y}{2(1+y)}(1^2 - 0^2)$$

$$E(X|Y=y) = \frac{2}{3.(1+y)} + \frac{y}{2.(1+y)} = \frac{4+3y}{6.(1+y)}$$

Resposta: Certo.

Densidade Condicional de uma variável

Podemos calcular a **densidade condicional** de uma variável, dado que ela **própria** pertence a determinado intervalo de valores. Em outras palavras, agora vamos falar sobre uma **única** variável (e não 2).

Nessa situação, trabalhamos com a densidade **marginal** (não conjunta) da variável e **dividimos** pela probabilidade de a variável pertencer ao referido **intervalo** (evento a priori):

$$\boxed{f(x|l_{inf} < X < l_{sup}) = \frac{f(x)}{P(l_{inf} < X < l_{sup})}}$$

É importante se atentar para o intervalo em que a variável é definida.

Para ilustrar, vamos considerar o mesmo exemplo anterior, em que calculamos a densidade marginal da variável X: $f(x) = x + \frac{1}{2}$, para intervalo $0 < X < 1$.

Vamos calcular a **densidade condicional** dessa variável, dado que ela assume valores $X < 0{,}5$:

$$f(x|X < 0{,}5) = \frac{x + \frac{1}{2}}{P(X < 0{,}5)}$$

Para isso, precisamos calcular a **probabilidade do evento a priori** $P(X < 0{,}5)$, que corresponde à integral da função densidade marginal $f(x) = x + \frac{1}{2}$, no intervalo $X < 0{,}5$.

Sabendo que a variável assume valores no intervalo (0, 1), então o intervalo da integral será $0 < X < 0{,}5$:

$$P(X < 0{,}5) = \int_0^{0{,}5} \left(x + \frac{1}{2}\right) dx = \left[\frac{x^2}{2} + \frac{1}{2}x\right]_0^{0{,}5} = \frac{1}{2}[x^2 + x]_0^{0{,}5}$$

$$P(X < 0{,}5) = \frac{1}{2}(0{,}5^2 - 0^2 + 0{,}5 - 0) = \frac{1}{2}(0{,}25 + 0{,}5) = \frac{1}{2}.0{,}75 = \frac{1}{2}.\frac{3}{4} = \frac{3}{8}$$

Agora, podemos calcular a função densidade condicional:

$$f(x|X < 0{,}5) = \frac{x + \frac{1}{2}}{\frac{3}{8}} = \left(x + \frac{1}{2}\right).\frac{8}{3} = \frac{8}{3}x + \frac{4}{3}$$

Com base na densidade condicional, podemos calcular a **probabilidade** de a variável pertencer a determinado **intervalo**, nessa condição. Para isso, integramos a função densidade condicional.

$$P(a < x < b|l_{inf} < X < l_{sup}) = \int f(x|l_{inf} < X < l_{sup}).dx$$

Sabendo que a densidade condicional é a razão entre a **densidade** da variável e a probabilidade de ela pertencer a determinado **intervalo**, temos:

$$P(a < X < b | l_{inf} < X < l_{sup}) = \int \frac{f(x)}{P(l_{inf} < X < l_{sup})}.dx$$

Ademais, considerando que a probabilidade do evento a priori $P(l_{inf} < X < l_{sup})$ é uma constante, ou seja, não é uma função de x, podemos retirá-la da integral:

$$\boxed{P(a < X < b | l_{inf} < X < l_{sup}) = \frac{1}{P(l_{inf} < X < l_{sup})}.\int f(x).dx}$$

Ainda não falamos sobre o **intervalo da integração**: ele corresponde à **interseção** entre o intervalo desejado $a < X < b$ e o intervalo da condição $l_{inf} < X < l_{sup}$, atentando-se para o intervalo em que a variável é definida.

Para esclarecer isso, vamos esquecer, por ora, que estamos trabalhando com integrais e calcular a **probabilidade condicional** $P(X > 0{,}3 | X < 0{,}5)$.

Pela fórmula da probabilidade condicional, temos a razão entre a probabilidade da **interseção** dos eventos, dividida pela probabilidade do evento a **priori**:

$$P(X > 0{,}3 | X < 0{,}5) = \frac{P(X > 0{,}3 \text{ } E \text{ } X < 0{,}5)}{P(X < 0{,}5)} = \frac{P(0{,}3 < X < 0{,}5)}{P(X < 0{,}5)}$$

A integral da função densidade corresponde justamente ao **numerador**, no caso, $P(0{,}3 < X < 0{,}5)$ e, por isso, o seu intervalo deve ser a **interseção** entre o intervalo desejado e o intervalo do evento a priori.

Para ilustrar, vamos calcular a probabilidade condicional $P(X < 0{,}2 | X < 0{,}5)$, utilizando o mesmo exemplo, $f(x) = x + \frac{1}{2}$, para $0 < X < 1$.

Como já calculamos a densidade condicional $f(x | X < 0{,}5) = \frac{8}{3}x + \frac{4}{3}$, basta integrarmos essa função no intervalo correspondente à **interseção** entre o intervalo desejado $X < 0{,}2$ e o intervalo do evento a priori $X < 0{,}5$, atentando-se para o intervalo em que a variável é definida $0 < X < 1$.

Logo, devemos integrar a função no intervalo $0 < X < 0{,}2$:

$$P(X < 0{,}2 | X < 0{,}5) = \int_0^{0{,}2} \left(\frac{8}{3}x + \frac{4}{3}\right).dx = \left[\frac{8}{3}.\frac{x^2}{2} + \frac{4}{3}x\right]_0^{0{,}2} = \frac{4}{3}[x^2 + x]_0^{0{,}2}$$

$$P(X < 0{,}2 | X < 0{,}5) = \frac{4}{3}[0{,}2^2 - 0^2 + 0{,}2 - 0]_0^{0{,}2} = \frac{4}{3}(0{,}24) = 0{,}32$$

Que é o mesmo que calcular a integral da função densidade $f(x) = x + \frac{1}{2}$, no mesmo intervalo $0 < x < 0{,}2$, e **dividir** pela probabilidade do evento a priori $P(X < 0{,}5) = \frac{3}{8}$.

Com base na densidade condicional, também podemos calcular as medidas de distribuição. A **esperança condicional** é calculada pela integral da função densidade **condicional**, multiplicada pela variável, no **intervalo da condição**:

$$E(X|l_{inf} < X < l_{sup}) = \int_{l_{inf}}^{l_{sup}} x.f(x|l_{inf} < X < l_{sup}).dx = \int_{l_{inf}}^{l_{sup}} x.\frac{f(x)}{P(l_{inf} < X < l_{sup})}.dx$$

Sabendo que a probabilidade é uma constante (não é função da variável), podemos retirá-la da integral:

$$\boxed{E(X|l_{inf} < X < l_{sup}) = \frac{1}{P(l_{inf}<X<l_{sup})}\int_{l_{inf}}^{l_{sup}} x.f(x).dx}$$

Note que a integral $\int_{l_{inf}}^{l_{sup}} x.f(x).dx$ corresponde à **esperança** clássica da variável, calculada no **intervalo** da condição $l_{inf} < X < l_{sup}$.

Em outras palavras, a esperança condicionada a um intervalo corresponde à esperança clássica calculada nesse intervalo, dividida pela probabilidade associada ao intervalo.

Para o nosso exemplo, a esperança $E(X|X < 0{,}5)$ pode ser calculada pela integral da densidade condicional $f(x|X < 0{,}5) = \frac{8}{3}x + \frac{4}{3}$ que já calculamos, multiplicada por x, no intervalo $0 < x < 0{,}5$:

$$E(X|X < 0{,}5) = \int_0^{0{,}5} x.f(x|X < 0{,}5).dx = \int_0^{0{,}5} x.\left(\frac{8}{3}x + \frac{4}{3}\right).dx = \int_0^{0{,}5} \left(\frac{8}{3}x^2 + \frac{4}{3}x\right).dx$$

Agora, vamos integrar:

$$E(X|X < 0{,}5) = \left[\frac{8}{3}.\frac{x^3}{3} + \frac{4}{3}.\frac{x^2}{2}\right]_0^{0{,}5} = \left[\frac{8}{9}.x^3 + \frac{2}{3}.x^2\right]_0^{0{,}5}$$

E aplicar os limites:

$$E(X|X < 0{,}5) = \frac{8}{9}.[0{,}5^3 - 0^3] + \frac{2}{3}.[0{,}5^2 - 0^2] = \frac{8}{9}.0{,}125 + \frac{2}{3}.0{,}25 \cong 0{,}2778$$

Que é o mesmo que calcularmos a esperança no intervalo $0 < x < 0{,}5$ e **dividir** pela probabilidade do evento a priori $P(X < 0{,}5) = \frac{3}{8}$.

Também podemos calcular a **variância condicional**:

$$V(X|l_{inf} < X < l_{sup}) = E(X^2|l_{inf} < X < l_{sup}) - [E(X|l_{inf} < X < l_{sup})]^2$$

Em que o **segundo momento condicional** de X é a integral da densidade condicional, multiplicada por x^2, no intervalo da condição:

$$E\left(X^2|l_{inf} < X < l_{sup}\right) = \int_{l_{inf}}^{l_{sup}} x^2.f\left(x|l_{inf} < X < l_{sup}\right).dx = \int_{l_{inf}}^{l_{sup}} x^2.\frac{f(x)}{P(l_{inf} < X < l_{sup})}.dx$$

Aqui também podemos retirar a probabilidade do evento a priori $P(l_{inf} < X < l_{sup})$ da integral:

$$\boxed{E\left(X^2|l_{inf} < X < l_{sup}\right) = \frac{1}{P(l_{inf}<X<l_{sup})}\int_{l_{inf}}^{l_{sup}} x^2.f(x).dx}$$

Similarmente ao que vimos para a esperança condicional, o segundo momento condicionado a um intervalo corresponde ao **segundo momento** calculado no **intervalo**, **dividido** pela probabilidade do intervalo.

Para o nosso exemplo, já calculamos a função densidade condicional $f(x|X < 0{,}5) = \frac{8}{3}x + \frac{4}{3}$, então podemos multiplicá-la por x^2 para integrar no intervalo da condição $0 < x < 0{,}5$:

$$E(X^2|X < 0{,}5) = \int_0^{0{,}5} x^2.f(x|X < 0{,}5).dx = \int_0^{0{,}5} x^2.\left(\frac{8}{3}x + \frac{4}{3}\right).dx = \int_0^{0{,}5} \left(\frac{8}{3}x^3 + \frac{4}{3}x^2\right).dx$$

Agora, vamos integrar:

$$E(X^2|X < 0{,}5) = \left[\frac{8}{3}.\frac{x^4}{4} + \frac{4}{3}.\frac{x^3}{3}\right]_0^{0{,}5} = \left[\frac{2}{3}.x^4 + \frac{4}{9}.x^3\right]_0^{0{,}5}$$

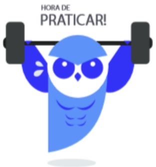
E aplicar os limites:

$$E(X^2|X < 0{,}5) = \frac{2}{3}.[0{,}5^4 - 0^4] + \frac{4}{9}.[0{,}5^3 - 0^3] = \frac{2}{3}.0{,}0625 + \frac{4}{9}.0{,}125 \cong 0{,}0972$$

Que é o mesmo que calcularmos o segundo momento no intervalo $0 < x < 0{,}5$ e **dividir** pela probabilidade do evento a priori $P(X < 0{,}5) = \frac{3}{8}$.

Com isso, podemos calcular a **variância condicional**:

$$V(X|X < 0{,}5) = E(X^2|X < 0{,}5) - [E(X|X < 0{,}5)]^2 \cong 0{,}0972 - [0{,}2778]^2 \cong 0{,}02$$



(FGV/2022 - TJDFT) Seja Y uma variável aleatória contínua com função de densidade de probabilidade $f_Y(y) = \begin{cases} \frac{5}{y^6}, & se\ y > 1 \\ 0; & caso\ contrário \end{cases}$. O valor de $E(Y|Y > 2)$ é

a) 2,1

b) 2,3

c) 2,5

d) 2,7

e) 2,9

Comentários:

Podemos calcular a esperança condicional de Y, dado que ela pertence ao intervalo Y > 2, dividindo a esperança da variável E(Y), pela probabilidade do evento a priori, no caso, P(Y > 2):

$$E(Y|Y > 2) = \int_2^{\infty} y.f(y|Y > 2).dy = \frac{1}{P(Y > 2)} \int_2^{\infty} y.f(y).dy$$

O intervalo da integral deve respeitar a condição e o intervalo da variável. Como a variável é definida no intervalo $y > 1$, ou seja, $(1, \infty)$, a condição $y > 2$ corresponde ao intervalo $(2, \infty)$.

Vamos começar calculando $P(Y > 2)$:

$$P(Y > 2) = \int_2^{\infty} f(y).dy = \int_2^{\infty} \frac{5}{y^6}.dy = \int_2^{\infty} 5.y^{-6}.dy = \left[5.\frac{y^{-5}}{-5}\right]_2^{\infty} = \left[-\frac{1}{y^5}\right]_2^{\infty}$$

Quando $y \to \infty$, a fração $\frac{1}{y^5}$ tende a zero:

$$P(Y > 2) = 0 - \left(-\frac{1}{2^5}\right) = \frac{1}{32}$$

E a esperança de Y, calculada no intervalo Y > 2, é:

$$E(Y > 2) = \int_2^{\infty} y.\frac{5}{y^6}.dy = \int_2^{\infty} 5.y^{-5}.dy = \left[5.\frac{y^{-4}}{-4}\right]_2^{\infty} = \left[-\frac{5}{4}.\frac{1}{y^4}\right]_2^{\infty}$$

$$E(Y > 2) = 0 - \left(-\frac{5}{4}.\frac{1}{2^4}\right) = \frac{5}{64}$$

Por fim, a esperança condicional é a razão entre a esperança e a probabilidade a priori:

$$E(Y|Y > 2) = \frac{\frac{5}{64}}{\frac{1}{32}} = \frac{5}{2} = 2{,}5$$

Gabarito: C

TRANSFORMAÇÃO DE VARIÁVEIS

Utilizamos a técnica da **transformação** de variáveis aleatórias (também chamada de **função** de uma variável aleatória), quando **conhecemos** a função densidade de probabilidade de uma variável X e queremos encontrar a função de **outra variável** Y, que possui uma **relação conhecida** com X, da forma $Y = h(X)$.

Essa transformação funciona tanto para variáveis discretas, quanto para variáveis contínuas.

Transformação de Variáveis Discretas

Para **variáveis discretas**, conhecemos as probabilidades de X, $P[X = x]$, e queremos calcular as probabilidades de Y, $P[Y = y]$.

Fazemos esse cálculo com base na relação entre Y e X que conhecemos, $Y = h(X)$:

$$P[Y = y] = P[h(X) = y] = P[X = h^{-1}(y)]$$

Para entender essa expressão, vamos supor a seguinte relação entre X e Y:

$$Y = 2X + 2$$

Então, a probabilidade $P(Y = y)$ corresponde à probabilidade:

$$P(Y = y) = P(2X + 2 = y)$$

Isolando X da expressão $2X + 2 = y$, temos:

$$X = \frac{y - 2}{2} = \frac{y}{2} - 1$$

Ou seja, a probabilidade $P(Y = y)$ corresponde à probabilidade:

$$P(Y = y) = P\left(X = \frac{y}{2} - 1\right)$$

Assim, para encontrar a função de probabilidade $P[Y = y]$ para $Y = h(X)$, percorremos os seguintes passos:

i) Primeiro, **invertemos a função** h, para encontrar a relação de **X em função de Y**: $X = h^{-1}(y)$.
ii) Em seguida, na expressão da probabilidade de **X**, **substituímos** o termo x pela expressão que calculamos no passo anterior, $\boldsymbol{h^{-1}(y)}$.

Por fim, os **possíveis valores** que Y pode assumir podem ser calculados a partir da **relação** de Y em função de X, $\boldsymbol{y = h(x)}$.

EXEMPLIFICANDO

Vamos considerar uma variável geométrica X, com probabilidade $p = 0{,}1$, ou seja:

$$P[X = x] = 0{,}1 \times 0{,}9^{x-1}, \text{ com } x = 1, 2, 3, \ldots$$

Vamos supor que a variável Y tenha a mesma relação com X: $\boldsymbol{Y = h(X) = 2X + 2}$.

Primeiro calculamos $X = h^{-1}(y)$, **isolando** a variável X:

$$X = \frac{Y}{2} + 1$$

Agora, na expressão de $P[X = x] = 0{,}1 \times 0{,}9^{x-1}$, substituímos x por $h^{-1}(y) = \boldsymbol{\frac{y}{2} + 1}$:

$$P[Y = y] = 0{,}1 \times 0{,}9^{\left(\frac{y}{2}+1\right)-1} = 0{,}1 \times 0{,}9^{\frac{y}{2}}$$

Essa é a função de **probabilidade de Y**.

Os possíveis valores de Y podem ser obtidos pela relação de Y em função de X:

$$\boldsymbol{y = 2x + 2}$$

Como os possíveis valores de $\boldsymbol{x = 1, 2, 3}$, ..., então os possíveis valores de Y são:

$$\boldsymbol{y = 2 \times 1 + 2 = 4}$$

$$\boldsymbol{y = 2 \times 2 + 2 = 6}$$

$$\boldsymbol{y = 2 \times 3 + 2 = 8}$$

...

Se as probabilidades de X, $P[X = x]$, **não** forem uma função de x, ou seja, se elas forem **fixas**, então não será necessário calcular a inversa $X = h^{-1}(y)$. Nesse caso, basta calcular os **possíveis valores de Y** pela relação $\boldsymbol{y_i = h(x_i)}$ e **substituir** x_i por y_i, **mantendo** as probabilidades correspondentes.

Por exemplo, considere os valores de X e as respectivas probabilidades indicadas na tabela a seguir:

x_i	$P(x_i)$
1	15%
3	25%
5	40%
7	20%

Considere também uma variável Y com a seguinte relação com a variável X: $Y = X^2 - 1$. Para encontrar a função de probabilidade de Y, basta calcular os valores de Y como $y_i = x_i^2 - 1$ e **manter** as probabilidades:

x_i	$y_i = x_i^2 - 1$	$P(x_i) = P(y_i)$
1	$1^2 - 1 = 0$	15%
3	$3^2 - 1 = 8$	25%
5	$5^2 - 1 = 24$	40%
7	$7^2 - 1 = 48$	20%

Transformação de Variáveis Contínuas

Para a **transformação (ou função) de variáveis contínuas**, temos uma situação parecida. Porém, como não existe função de probabilidade, $P[X = x]$, podemos fazer a transformação utilizando a **Função de Distribuição Acumulada** (chamado **Método da Função de Distribuição**) ou a **Função Densidade de Probabilidade** (chamado **Método Jacobiano**).

Método da Função de Distribuição

Pelo **Método da Função de Distribuição**, obtemos a f.d.a. de Y, $F_Y[y]$, da seguinte forma:

$$F_Y[y] = P[Y \leq y] = P[h(X) \leq y] = P[X \leq h^{-1}(y)] = F_X[h^{-1}(y)]$$

Vamos entender essa expressão. Por exemplo, sendo $Y = 2X$, a f.d.a. de Y corresponde à seguinte expressão:

$$F_Y[y] = P[Y \leq y] = P[2X \leq y]$$

Dividindo ambos os lados da inequação $2X \leq y$ por 2, isolamos X:

$$F_Y[y] = P\left(X \leq \frac{y}{2}\right)$$

Sabendo que a probabilidade $P\left(X \leq \frac{y}{2}\right)$ corresponde à f.d.a. de X, no ponto $\frac{y}{2}$, então:

$$F_Y[y] = F_X\left[\frac{y}{2}\right]$$

Essa expressão significa que, para encontrar a f.d.a. de Y, a partir da f.d.a. de X, sabendo que $Y = h(X)$, devemos percorrer os seguintes passos:

i) Primeiro, **invertemos a função** h, para encontrar a relação de **X em função de Y**: $X = h^{-1}(y)$.
ii) Em seguida, na função de distribuição **acumulada** de **X**, **substituímos** o termo x pela expressão que calculamos no passo anterior, $\boldsymbol{h^{-1}(y)}$.

Por fim, o intervalo dos **possíveis valores** de Y pode ser calculado a partir da relação de Y em função de X, $\boldsymbol{y = h(x)}$.

EXEMPLIFICANDO

Vamos supor a seguinte f.d.a. de X, para $x \geq 1$:

$$F_X(x) = 1 - \frac{1}{x^2}$$

Considerando a relação $Y = h(X) = 2X$, primeiro, calculamos a função **inversa**, $h^{-1}(y)$, **isolando** X dessa expressão:

$$X = \frac{y}{2}$$

Assim, a f.d.a. de Y corresponde à f.d.a. de X no ponto $\frac{y}{2}$. Ou seja, devemos substituir o termo x por $\frac{y}{2}$, na função acumulada de X, $F_X(x) = 1 - \frac{1}{x^2}$,:

$$F_Y[y] = F_X\left[\frac{y}{2}\right] = 1 - \frac{1}{\left(\frac{y}{2}\right)^2} = 1 - \frac{1}{\frac{y^2}{4}} = 1 - \frac{4}{y^2}$$

Sabendo que $x \geq 1$ e que $Y = 2X$, então Y assume valores no intervalo $y \geq 2$.

Porém, a relação $F_Y[y] = F_X[h^{-1}(y)]$ está correta somente se $Y = h(X)$ for uma função **crescente**, ou seja, se X e Y crescerem no mesmo sentido. Caso contrário, a relação será a seguinte:

$$F_Y[y] = P[Y \leq y] = P[h(X) \leq y] = P[X \geq h^{-1}(y)] = 1 - F_X[h^{-1}(y)]$$

Vamos entender essa expressão. Por exemplo, sendo $Y = -2X$, temos:

$$F_Y[y] = P[Y \leq y] = P[-2X \leq y]$$

Dividindo ambos os lados da inequação por -2, precisamos inverter o sinal da inequação[1]:

$$F_Y[y] = P\left(X \geq -\frac{y}{2}\right)$$

[1] Se você não entende o porquê dessa inversão, suponha a seguinte inequação, por exemplo:

$$-X < -5$$

Se -X é um número menor que -5 (por exemplo, -6, -7,...), então podemos concluir que X é **maior** que 5:

$$X > 5$$

Ou seja, quando multiplicamos os lados da inequação por um número negativo, precisamos **inverter** o sinal da inequação.

A probabilidade $P\left(X \geq -\frac{y}{2}\right)$ é **complementar** à probabilidade $P\left(X \leq \frac{y}{2}\right)$:

$$F_Y[y] = 1 - P\left(X \leq \frac{y}{2}\right)$$

Sabendo que $P\left(X \leq \frac{y}{2}\right)$ corresponde à f.d.a. de X no ponto $\frac{y}{2}$, temos:

$$F_Y[y] = 1 - F_Y\left(\frac{y}{2}\right)$$

Se $Y = h(X)$ for uma **crescente**: $F_Y[y] = F_X[h^{-1}(y)]$
Se $Y = h(X)$ for uma **decrescente**: $F_Y[y] = 1 - F_X[h^{-1}(y)]$

Método Jacobiano

Se conhecermos a **função densidade de probabilidade** de X, f_X, e desejarmos calcular a função densidade de probabilidade de Y, f_Y, podemos fazer a transformação diretamente, sem precisar calcular as funções acumuladas, pelo **Método Jacobiano**.

$$\boxed{f_Y[y] = f_X[h^{-1}(y)].\left|\frac{dh^{-1}(y)}{dy}\right|}$$

Os passos para resolver a transformação por esse método são:

i) **Inverter a função** h, para encontrar a relação de **X em função de Y**, isto é, $X = h^{-1}(y)$;
ii) Na função densidade de probabilidade de **X**, **substituir** o termo x pela expressão calculada no passo anterior, $\boldsymbol{h^{-1}(y)}$;
iii) **Derivar** a função calculada no passo i, $\boldsymbol{h^{-1}(y)}$; e
iv) Multiplicar o **módulo** da derivada calculada no passo iii pela expressão do passo ii.

Por exemplo, considere a f.d.p. de X, $f_X(x) = 3x^2$, para $x \in [0,1]$ e $Y = h(X) = 2X$.

Primeiro, vamos calcular a função h^{-1}:

$$Y = 2X \rightarrow X = \frac{\boldsymbol{Y}}{\boldsymbol{2}}$$

Assim, $f_X[h^{-1}(y)]$ corresponde a:

$$f_X(x) = 3\boldsymbol{x}^2$$

$$f_X[h^{-1}(y)] = 3\left(\frac{\boldsymbol{y}}{\boldsymbol{2}}\right)^2 = \frac{3}{4}y^2$$

O valor da derivada de $h^{-1}(y)$ é:

$$\frac{dh^{-1}(y)}{dy} = \frac{d\left(\frac{\boldsymbol{y}}{\boldsymbol{2}}\right)}{dy} = \frac{1}{2}$$

Como $\frac{dh^{-1}(y)}{dy}$ é **positivo**, então ela equivale ao seu módulo $\left|\frac{dh^{-1}(y)}{dy}\right| = \frac{dh^{-1}(y)}{dy}$.

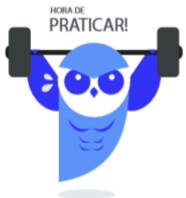
Assim, o produto $f_X[h^{-1}(y)] \times \left|\frac{dh^{-1}(y)}{dy}\right|$ é:

$$f_Y[y] = \frac{3}{4}y^2 \times \frac{1}{2} = \frac{3}{8}y^2$$

Como $x \in [0,1]$ e $y = 2x$, temos $y \in [0,2]$.

Para resolver questões dessa matéria, pode ser necessário conhecer o resultado da **derivada** do logaritmo na base *e*, que indicamos por **ln**:

$$\frac{d(\ln x)}{dx} = \frac{1}{x}$$



(CESPE/2010 – ABIN) A função de densidade da distribuição normal padrão Z é dada pela função $f(z) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)$, em que z é um número real.

Considerando a transformação Y = exp(Z), julgue o item a seguir.

A função de densidade da variável aleatória Y é $f(y) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{[\exp(y)]^2}{2}\right\}$, em que $y \geq 0$.

Comentários:

Para uma transformação da forma $Y = h(Z)$, vimos que a f.d.p. de Y é dada por:

$$f_Y[y] = f_Z[h^{-1}(y)].\left|\frac{dh^{-1}(y)}{dy}\right|$$

Sendo $Y = e^Z$, a inversa $h^{-1}(y)$ pode ser calculada aplicando a função logarítmica, na base e, em ambos os lados da equação:

$$\ln Y = \ln e^Z$$

Pela propriedade da função logarítmica, o logaritmo de e^Z na base e é justamente Z ($\ln e^Z = Z$):

$$\ln Y = Z$$

Logo, para obter $f_Z[h^{-1}(y)]$, substituímos **z** por $h^{-1}(y) = \mathbf{\ln y}$, na f.d.p. de Z:

$$f_Z[h^{-1}(y)] = f_Z[\ln y] = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(\mathbf{\ln y})^2}{2}\right)$$

Em seguida, multiplicamos essa expressão pelo módulo da derivada de $h^{-1}(y) = \ln y$:

$$\left|\frac{dh^{-1}(y)}{dy}\right| = \left|\frac{d(\ln y)}{dy}\right| = \left|\frac{1}{y}\right| = \frac{1}{|y|}$$

Ou seja, a f.d.p. de Y é dada por:

$$f_Y[y] = \frac{1}{\sqrt{2\pi}.|y|}\exp\left(-\frac{(\ln y)^2}{2}\right)$$

Que é **diferente** da expressão fornecida no item. Note que a f.d.p. de Y fornecida no item substituiu z por e^y, em vez de substituir a inversa $h^{-1}(y) = \ln y$, que é o correto.

Gabarito: Errado.

Uma transformação muito comum é a **transformação de lei quadrática**: $Y = X^2$.

Nesse caso, a função não é nem somente crescente e nem somente decrescente. Além disso, a inversa não possui valor único, pois $X = \pm\sqrt{Y}$. Nesse caso, temos:

$$\boxed{f_Y[y] = \frac{1}{2\sqrt{y}}\left[f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right)\right]}$$

Nessa expressão, $f_X(\sqrt{y})$ significa que, na f.d.p. de x, substituímos x por $\sqrt{y}$, e $f_X(-\sqrt{y})$ significa que substituímos x por $-\sqrt{y}$.

Se $\boldsymbol{X \geq 0}$, então $X = +\sqrt{Y}$. Logo, $f_X(-\sqrt{y}) = 0$. Nesse caso, temos:

$$f_Y[y] = \frac{1}{2\sqrt{y}}\left[f_X\left(\sqrt{y}\right)\right]$$

Analogamente, se $\boldsymbol{X \leq 0}$, então $X = -\sqrt{Y}$. Logo, $f_X(\sqrt{y}) = 0$. Nesse caso, temos:

$$f_Y[y] = \frac{1}{2\sqrt{y}}\left[f_X\left(-\sqrt{y}\right)\right]$$

Por exemplo, vamos considerar a mesma f.d.p. para X que vimos anteriormente:

$$f_X(x) = 3x^2, para\ x \in [0,1]$$

Nesse caso, temos:

$$f_X(\sqrt{y}) = 3(\sqrt{y})^2 = 3y$$

Como a função $f_X(x)$ está definida apenas para $x > 0$, então $f_X(-\sqrt{y}) = 0$. Logo:

$$f_Y[y] = \frac{1}{2\sqrt{y}}[3y]$$

INDO MAIS FUNDO!

De maneira geral, quando a função $Y = h(X)$ **não** for biunívoca e houver **dois valores de X** associados ao **mesmo** valor de Y, que é o caso da lei quadrática $Y = X^2$, é necessário **separar** a função inversa $h^{-1}(X)$ em **duas** funções $h_1^{-1}(X)$ e $h_2^{-1}(X)$ de modo que cada valor de Y esteja associado a um **único** valor de X em cada função.

Nessa situação, a função densidade de Y é definida como a **soma** da transformação de **ambas** as funções $h_1(X)$ e $h_2(X)$:

$$f_Y[y] = f_X[h_1^{-1}(y)] \times \left|\frac{dh_1^{-1}(y)}{dy}\right| + f_X[h_2^{-1}(y)] \times \left|\frac{dh_2^{-1}(y)}{dy}\right|$$

Vamos aplicar essa fórmula para a lei quadrática $Y = X^2$. As funções inversas são:

$$h_1^{-1}(y): x = +\sqrt{y}$$

$$h_2^{-1}(y): x = -\sqrt{y}$$

As derivadas dessas funções são:

$$\frac{dh_1^{-1}(y)}{dy} = \frac{d\left(y^{1/2}\right)}{dy} = \frac{1}{2}y^{-1/2} = \frac{1}{2\sqrt{y}}$$

$$\frac{dh_2^{-1}(y)}{dy} = \frac{d\left(-y^{1/2}\right)}{dy} = -\frac{1}{2}y^{-1/2} = -\frac{1}{2\sqrt{y}}$$

Como $\sqrt{y}$ é necessariamente positivo, então o **módulo** das derivadas de ambas as funções é igual a $\frac{1}{2\sqrt{y}}$. Assim, a função densidade de Y é aquela que havíamos visto anteriormente:

$$f_Y[y] = f_X[\sqrt{y}] \times \frac{1}{2\sqrt{y}} + f_X[-\sqrt{y}] \times \frac{1}{2\sqrt{y}} = \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})]$$

Transformação de Variáveis Multidimensionais

Também podemos utilizar a transformação para **variáveis multidimensionais** (ou para múltiplas variáveis).

Para **variáveis bidimensionais**, partimos da função densidade conjunta de duas variáveis X e Y e da relação entre essas variáveis e duas outras variáveis U e V, da forma $(U, V) = \boldsymbol{h}(X, Y)$, para obtermos a função densidade conjunta das variáveis U e V, dada por:

$$\boxed{f_{UV}(u, v) = f_{XY}[h^{-1}(u, v)].|J|}$$

Em que $|J|$, chamado **Jacobiano da transformação**, é o **módulo** do **determinante** da matriz formada pelas **derivadas parciais** de x e y em relação a u e v:

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix}$$

O determinante de uma matriz 2x2 é calculado multiplicando-se os elementos da diagonal principal e subtraindo-se o produto dos elementos da diagonal secundária:

$$J = \frac{\partial x}{\partial u}.\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}.\frac{\partial y}{\partial u}$$

ESCLARECENDO!

Utilizamos o **módulo** do determinante da matriz, isto é, o seu valor **absoluto**, desconsiderando-se eventual sinal negativo, pois a mera **inversão** de ordem das variáveis (o que **não** deve afetar o resultado da transformação) **inverte o sinal** do determinante:

$$J' = \begin{vmatrix} \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \\ \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \end{vmatrix} = \frac{\partial y}{\partial u}.\frac{\partial x}{\partial v} - \frac{\partial y}{\partial v}.\frac{\partial x}{\partial u}$$

Que corresponde ao mesmo resultado do determinante anterior J, porém com sinal contrário. Por isso, utilizamos o **módulo** do determinante.

Para obter $f_{XY}[h^{-1}(u, v)]$, substituímos x e y pelas respectivas expressões em função de $\boldsymbol{u}$ e $\boldsymbol{v}$, na função densidade conjunta f_{XY}.

Por exemplo, vamos supor que X e Y sejam variáveis aleatórias independentes com distribuição exponencial. Por serem independentes, a função densidade conjunta corresponde ao produto das funções densidade de cada variável:

$$f_{XY}(x,y) = f_X(x) \times f_Y(y) = \lambda e^{-\lambda x} \times \lambda e^{-\lambda y} = \lambda e^{-\lambda(x+y)}$$

Vamos considerar a seguinte relação entre (U,V) e (X,Y):

$$u = 3x + 5y \qquad\qquad v = x + 2y$$

O primeiro passo é calcular a **inversa** $(X,Y) = h^{-1}(U,V)$, encontrando as expressões de x e y em função de u e v. Para isso, começamos isolando x na primeira equação:

$$u = 3x + 5y$$

$$u - 5y = 3x$$

$$x = \frac{u - 5y}{3}$$

E fazemos o mesmo para a segunda equação:

$$v = x + 2y$$

$$x = v - 2y$$

▪

Em seguida, **igualamos** ambas as expressões de x; e **isolamos** y para encontrar a expressão de y em função de u e v:

$$\frac{u - 5y}{3} = v - 2y$$

$$\frac{u}{3} - \frac{5y}{3} = v - \frac{6y}{3}$$

$$\frac{y}{3} = v - \frac{u}{3}$$

$$y = 3v - u$$

Agora, **substituímos** essa expressão em uma das equações de x para encontrar a expressão de x em função de u e v:

$$x = v - 2y = v - 2(3v - u)$$

$$x = v - 6v + 2u$$

$$x = 2u - 5v$$

Assim, obtemos $f_{XY}[h^{-1}(u,v)]$, substituindo x por $2u - 5v$ e y por $3v - u$ na função conjunta f_{XY}:

$$f_{XY}(x,y) = \lambda e^{-\lambda(x+y)}$$

$$f_{XY}[h^{-1}(u,v)] = \lambda e^{-\lambda(2u-5v+3v-u)} = \lambda e^{-\lambda(u-2v)}$$

Agora, passemos ao **Jacobiano da transformação**.

Para isso, precisamos das **derivadas parciais** de x e y em função de $\boldsymbol{u}$ e $\boldsymbol{v}$. Para calcularmos a derivada parcial de uma variável, consideramos a outra variável como uma **constante**:

$$\frac{\partial x}{\partial u} = \frac{\partial(2u - 5v)}{\partial u} = 2$$

$$\frac{\partial x}{\partial v} = \frac{\partial(2u - 5v)}{\partial v} = -5$$

$$\frac{\partial y}{\partial u} = \frac{\partial(3v - u)}{\partial u} = -1$$

$$\frac{\partial y}{\partial v} = \frac{\partial(3v - u)}{\partial v} = 3$$

E o determinante da matriz das derivadas parciais é:

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 2 & -5 \\ -1 & 3 \end{vmatrix} = 2 \times 3 - (-5) \times (-1) = 6 - 5 = 1$$

As variáveis que constam nos **numeradores** das derivadas estão associadas às **linhas** da matriz Jacobiana, ou seja, em toda a primeira linha, temos ∂x como numerador e em toda a segunda linha, temos ∂y como numerador.

Já as variáveis que constam **denominadores** das derivadas estão associadas às **colunas** da matriz Jacobiana, ou seja, em toda a primeira coluna, temos ∂u como denominador e em toda a segunda coluna, temos ∂v como denominador.

Assim, a função densidade conjunta de U e V é dada por:

$$f_{UV}(u,v) = \lambda e^{-\lambda(u-2v)}.|1| = \lambda e^{-\lambda(u-2v)}$$

INDO MAIS FUNDO!

A fórmula da transformação pode ser estendida para **qualquer número de variáveis**. Para uma variável **tridimensional**, temos:

$$f_{UVW}(u,v,w) = f_{XYZ}[h^{-1}(u,v,w)].|J|$$

Em que $|J|$ é a transformação Jacobiana, que corresponde ao **módulo** do **determinante** da matriz formada pelas **derivadas parciais** de x, y e z em relação a u, v e w:

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}$$

Para calcular o determinante de uma matriz 3x3, replicamos as duas primeiras colunas, somamos os produtos das diagonais para a direita, indicadas em verde, e subtraímos os produtos das diagonais para a esquerda, indicadas em azul:

$$\begin{array}{ccccc} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} & \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} & \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} & \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{array} \qquad \begin{array}{ccccc} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} & \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} & \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} & \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{array}$$

$$J = +\frac{\partial x}{\partial u}.\frac{\partial y}{\partial v}.\frac{\partial z}{\partial w} + \frac{\partial x}{\partial v}.\frac{\partial y}{\partial w}.\frac{\partial z}{\partial u} + \frac{\partial x}{\partial w}.\frac{\partial y}{\partial u}.\frac{\partial z}{\partial v} - \frac{\partial x}{\partial w}.\frac{\partial y}{\partial v}.\frac{\partial z}{\partial u} - \frac{\partial x}{\partial u}.\frac{\partial y}{\partial w}.\frac{\partial z}{\partial v} - \frac{\partial x}{\partial v}.\frac{\partial y}{\partial u}.\frac{\partial z}{\partial w}$$

HORA DE PRATICAR!

(FGV/2019 – DPE-RJ - *Adaptada*) Seja a variável aleatória bidimensional (X,Y) que tem distribuição uniforme no quadrado 0 < x < 1 e 0 < y < 1 e Zero fora dele. Por uma transformação linear é definida a v.a. bidimensional (Z,W) da seguinte maneira:

Z = X + Y e W = X – Y

Então, sobre essa outra variável bidimensional, analise a afirmativa a seguir:

O Jacobiano da transformação de (X,Y) para (Z,W) vale 0,25.

Comentários:

O **Jacobiano da transformação** é o **módulo** do **determinante** da matriz formada pelas **derivadas parciais** de X e Y em função de Z e W:

$$J = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix}$$

Para obtê-la, precisamos representar X e Y em função de Z e W. Primeiro, **isolamos** x nas equações:

$$z = x + y \rightarrow x = z - y$$

$$w = x - y \rightarrow x = w + y$$

Agora, igualamos as equações para representarmos Y como uma função de Z e W:

$$z - y = w + y$$

$$z - w = 2y$$

$$y = \frac{z - w}{2}$$

Agora, **substituímos** essa expressão em uma das equações de X:

$$x = w + y = w + \frac{z - w}{2} = \frac{2w + z - w}{2} = \frac{z + w}{2}$$

Em seguida, calculamos as derivadas parciais:

$$\frac{\partial x}{\partial z} = \frac{\partial \left(\frac{z + w}{2}\right)}{\partial z} = \frac{1}{2}$$

$$\frac{\partial x}{\partial w} = \frac{\partial \left(\frac{z + w}{2}\right)}{\partial w} = \frac{1}{2}$$

$$\frac{\partial y}{\partial z} = \frac{\partial \left(\frac{z - w}{2}\right)}{\partial z} = \frac{1}{2}$$

$$\frac{\partial y}{\partial w} = \frac{\partial \left(\frac{z - w}{2}\right)}{\partial w} = -\frac{1}{2}$$

E o determinante da matriz é:

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = \frac{1}{2} \times \left(-\frac{1}{2}\right) - \frac{1}{2} \times \frac{1}{2} = -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2}$$

E o módulo é $|J| = \frac{1}{2} = 0{,}5$

Resposta: Errado

RESUMO DA AULA

Momentos:

- **k-ésimo momento teórico: $\mu_k = E(X^k)$**
  - Variáveis **Discretas**: $\mu_k = \sum_i (x_i)^k . P(X = x_i)$
  - Variáveis **Contínuas**: $\mu_k = \int_{x_I}^{x_S} (x)^k . f(x). dx$
  - **1º Momento = Esperança**; **Variância** = 2º momento – [1º momento]^2
- **k-ésimo momento central teórico: $\mu_{k_C} = E((X - \mu)^k)$**
  - 1º Momento Central = **0**; 2º Momento Central = **Variância**
- **k-ésimo momento amostral: $m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$**
  - **1º Momento Amostral = Média Amostral**

Função Geradora de Momentos (f.g.m.):

- Função a partir da qual **todos os momentos** podem ser calculados: $M_X(t) = E(e^{t.X})$
  - Variáveis **Discretas**: $M_X(t) = \sum_i e^{t.x_i} \times P(X = x_i)$
  - Variáveis **Contínuas**: $M_X(t) = \int_{x_I}^{x_S} e^{t.X} . f(x). dx$
- k-ésimo momento = k-ésima **derivada** no ponto **t = 0**: $M_X^{(k)}(0) = E(X^k)$
- **Propriedades da f.g.m.:**
  - $M_{a.X+b}(t) = e^{b.t} \times M_X(a.t)$
  - Para X e Y **independentes**: $M_{X+Y}(t) = M_X(t) \times M_Y(t)$
  - **Se $M_X(t) = M_Y(t)\ \forall t$ então X = Y**
- **F.g.m. de distribuições teóricas:**
  - **Distribuição Binomial: $M_{X_{Bi}}(t) = (p.e^t + q)^n$**
  - **Distribuição Geométrica: $M_{X_{Geo}}(t) = \frac{p.e^t}{1-q.e^t}$**
  - **Distribuição de Poisson: $M_{X_{Po}}(t) = e^{\lambda(e^t-1)}$**

- **Distribuição Exponencial:** $M_{X_{Exp}}(t) = \frac{\lambda}{\lambda - t}$
- **Distribuição Normal:** $M_{X_N}(t) = e^{\frac{\sigma^2 . t^2}{2} + \mu . t}$
- **Distribuição Qui-Quadrado:** $M_{X_{Qui}}(t) = \left(\frac{1}{1-2.t}\right)^{\frac{k}{2}}$

Distribuições Conjuntas:

- Variáveis **Discretas**: $P(X = x, Y = y)$;
  - Probabilidade **Marginal**: $P(X = x) = \sum_j p(x, y_j)$; $P(Y = y) = \sum_i p(x_i, y)$
  - Variáveis **independentes**: $P(X = x, Y = y) = P(X = x) \times P(Y = y),\ \forall (x, y)$
  - Probabilidade **condicional**: $P(Y = y|X = x) = \frac{P(X=x,Y=y)}{P(X=x)}$
  - **Esperança condicional:** $E(Y|X = x) = \sum_j y_j . P(Y = y_j|X = x)$
  - **2º momento condicional:** $E(Y^2|X = x) = \sum_j (y_j)^2 . P(Y = y_j|X = x)$
  - **Variância condicional**: $V(Y|X = x) = E(Y^2|X = x) - [E(Y|X = x)]^2$
  - **Esperança marginal**: $E(X) = E_Y[E(X|Y)] = \sum_j E(X|Y = y_j) \times P(Y = y_j)$
  - **Variância marginal:** $V(X) = E[V(X|Y)] + V[E(X|Y)]$
- **Distribuição multinomial:** generalização da distribuição binomial para k categorias
  - $p(n_1, n_2, \dots, n_k) = \frac{N!}{n_1! n_2! \dots n_k!} \times (p_1)^{n_1} \times (p_2)^{n_2} \times \dots \times (p_k)^{n_k}$
  - $\sum_{i=1}^{k} n_i = N, \qquad \sum_{i=1}^{k} p_i = 1$

Densidade Conjunta:

- Variáveis **Contínuas**: $f(x, y)$; $P(x_1 < X < x_2, y_1 < Y < y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} f(x, y) . dx . dy$
  - Densidade **Marginal**: $f(x) = \int_{-\infty}^{\infty} f(x, y) . dy$; $f(y) = \int_{-\infty}^{\infty} f(x, y) . dx$
  - Variáveis **independentes**: $\boldsymbol{f(x, y) = f(x) \times f(y)}$
  - **Densidade condicional:** $f(y|X = x) = \frac{f(x,y)}{f(X=x)}$

- **Esperança condicional:** $E(Y|X = x) = \int_{-\infty}^{\infty} y . f(y|X = x) . dy$
- **2º momento condicional:** $E(Y^2|X = x) = \int_{-\infty}^{\infty} y^2 . f(y|X = x) . dy$
- **Esperança marginal:** $E(X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x . f(x, y) . dy . dx$
- **2º momento marginal:** $E(X^2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 . f(x, y) . dy . dx$

Transformação:

- Cálculo da distribuição de Y a partir da distribuição de X e da relação $Y = h(X)$
- Variáveis **Discretas**: $P[Y = y] = P[X = h^{-1}(y)]$
- Variáveis **Contínuas**:
  - Transformação pela f.d.a., se $h(X)$ crescente: $F_Y[y] = F_X[h^{-1}(y)]$
  - Transformação pela f.d.a., se $h(X)$ decrescente: $F_Y[y] = 1 - F_X[h^{-1}(y)]$
  - Transformação pela f.d.p., se $f_Y[y] = f_X[h^{-1}(y)].\left|\frac{dh^{-1}(y)}{dy}\right|$
  - Transformação da **lei quadrática** ($Y = X^2$): $f_Y[y] = \frac{1}{2\sqrt{y}}\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right]$

DISTRIBUIÇÃO MULTINOMIAL

A distribuição **multinomial** é uma **generalização** da distribuição **binomial**, para quando houver **mais de dois** resultados possíveis em cada ensaio, ou seja, quando os ensaios **não** forem necessariamente de **Bernoulli**.

Por exemplo, se eu quisesse estudar a cor dos olhos de uma população, para utilizar a distribuição binomial, eu teria que dividir a população em apenas duas categorias, olhos claros e olhos escuros (isto é, olhos não claros). Já, utilizando a distribuição multinomial, posso dividir a população em quantas categorias interessar: olhos verdes, azuis, marrons, pretos, bicolores,...

Generalizando, na distribuição multinomial, a população é dividida em $\boldsymbol{k}$ categorias, com probabilidades $\boldsymbol{p_1, p_2, \dots, p_k}$, respectivamente, cuja soma é $\boldsymbol{\sum_{i=1}^{k} p_i = 1}$.

Assim como a distribuição binomial, serão realizados $\boldsymbol{N}$ **ensaios independentes**, mas os resultados desses ensaios poderão ser qualquer um dos $\boldsymbol{k}$ valores possíveis da população. O número de vezes em que obtemos, como **resultado**, cada uma das **categorias** $1, 2, \dots, k$ é representado pela respectiva variável aleatória $\boldsymbol{X_1, X_2, \dots, X_k}$, chamada **variável multinomial**.

Dessa forma, o conjunto das variáveis $X_1, X_2, \dots, X_k$ segue uma distribuição multinomial, em que a **probabilidade conjunta** de obtermos n_1 vezes a categoria 1 (isto é, $X_1 = n_1$), n_2 vezes a categoria 2 (isto é, $X_2 = n_2$), ..., e n_k vezes a categoria k (isto é, $X_k = n_k$), é dada por:

$$\boxed{\boldsymbol{p(n_1, n_2, \dots, n_k) = \frac{N!}{n_1! n_2! \dots n_k!} \times (p_1)^{n_1} \times (p_2)^{n_2} \times \dots \times (p_k)^{n_k}}}$$

Em que o número total de ensaios N equivale ao somatório dos números de ensaios de cada variável:

$$\sum_{i=1}^{k} n_i = N$$

Antes de fazermos um exemplo, vamos entender um pouquinho essa fórmula. Ela envolve a probabilidade de obter $X_1 = n_1, X_2 = n_2, \dots, X_k = n_k$, **nessa ordem**, dada por:

$$(p_1)^{n_1} \times (p_2)^{n_2} \times \dots \times (p_k)^{n_k}$$

Mas como a ordem não precisa ser essa, multiplicamos esse resultado pela **permutação** de N elementos, com **repetição** de $n_1, n_2, \dots, n_k$ elementos[1]:

$$P_N^{n_1,n_2,\dots,n_k} = \frac{N!}{n_1!\, n_2! \dots n_k!}$$

EXEMPLIFICANDO

> Suponha uma caixa com 10 bolas, das quais 4 são pretas, 3 são brancas, 2 são azuis e 1 é verde, e que vamos selecionar uma bola ao acaso, observar a sua cor e devolvê-la à caixa. Vamos repetir esse procedimento 5 vezes e calcular a probabilidade de obter 2 vezes uma bola preta e 1 vez as demais cores.
>
> Assim, X_1 representará o número de vezes que obtemos uma bola **preta**, X_2 representará o número de vezes que obtemos uma bola **branca**, X_3 representará o número de vezes que obtemos uma bola **azul** e X_4, o número de vezes que obtemos a bola **verde**.
>
> As respectivas probabilidades são:
>
> $$p_1 = \frac{4}{10} = \frac{2}{5}, \quad p_2 = \frac{3}{10}, \quad p_3 = \frac{2}{10} = \frac{1}{5}, \quad p_4 = \frac{1}{10}$$
>
> Logo, para $n_1 = 2$ e $n_2 = n_3 = n_4 = 1$, temos a seguinte probabilidade:
>
> $$p(n_1, n_2, n_3, n_4) = \frac{N!}{n_1!n_2!n_3!n_4!} \times (p_1)^{n_1} \times (p_2)^{n_2} \times (p_3)^{n_3} \times (p_4)^{n_4}$$
>
> $$p(2,1,1,1) = \frac{5!}{2!1!1!1!} \times \left(\frac{2}{5}\right)^2 \times \left(\frac{3}{10}\right)^1 \times \left(\frac{1}{5}\right)^1 \times \left(\frac{1}{10}\right)^1$$
>
> $$p(2,1,1,1) = 5 \times 4 \times 3 \times \frac{4}{25} \times \frac{3}{10} \times \frac{1}{5} \times \frac{1}{10} = \frac{144}{2500} = 5{,}76\%$$

Nesse exemplo, poderíamos utilizar a distribuição **binomial** caso quiséssemos calcular a probabilidade de obter **apenas um dos resultados**, por exemplo, **apenas 2** vezes a bola preta. Nesse caso, teríamos os parâmetros $n = 2$ e $p = {}^4/_{10}$.

Por outro lado, se **não houvesse reposição**, poderíamos calcular a probabilidade de obter 2 bolas pretas e 1 bola das demais cores, utilizando a distribuição **hipergeométrica**. Porém, como há reposição, precisamos efetuar os cálculos utilizando a multinomial.

---

[1] Se preferir, você pode pensar na partição de N elementos em grupos de $n_1, n_2, \dots, n_k$ elementos (a fórmula será a mesma).

Pontue-se que a distribuição multinomial é **reduzida** à binomial para $\boldsymbol{k = 2}$. Nesse caso, a fórmula da distribuição multinomial é dada por:

$$p(n_1, n_2) = \frac{N!}{n_1!\, n_2!} \times (p_1)^{n_1} \times (p_2)^{n_2}$$

Sabendo que $\sum_{i=1}^{k} p_i = 1$ e $\sum_{i=1}^{k} n_i = N$, então nesse caso, temos: $p_2 = 1 - p_1$ e $n_2 = N - n_1$:

$$p(n_1, N - n_1) = \frac{N!}{n_1!\,(N - n_1)!} \times (p_1)^{n_1} \times (1 - p_1)^{N - n_1}$$

Essa é a distribuição binomial, com os parâmetros N e $p = p_1$, em que normalmente denotamos n_1 (o número de resultados com o atributo sucesso) por k.

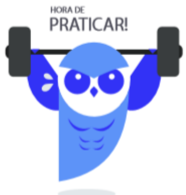
O **valor esperado** e a **variância** são calculados para **cada variável** $X_i, i = 1, 2, \dots, k$, como na distribuição **binomial**, considerando os parâmetros N e p_i:

$$E(X_i) = N.p_i$$

$$V(X_i) = N.p_i(1 - p_i)$$

A **covariância** entre duas variáveis X_u e X_v é dada por:

$$Cov(X_u, X_v) = -n.p_u.p_v$$



(FGV/2019 – DPE/RJ – *Adaptada*) O nível de escolaridade dos cidadãos que necessitam recorrer à Defensoria Pública do RJ segue uma distribuição multinomial com parâmetros p1 = 0,4, p2 = 0,3, p3 = 0,2 e p4 = 0,1, que são as probabilidades de que pertençam à classe menos instruída (Cp1) até a classe mais instruída (Cp4). Com base nessa distribuição, julgue o item a seguir.

Selecionando 10 pessoas ao acaso, com reposição, a probabilidade de obter as quantidades Cp1 = 4, Cp2 = 3, Cp3 = 2 e Cp4 = 1 é menor que 3%.

Comentários:

A probabilidade de $n_1 = 4, n_2 = 3, n_3 = 2, n_4 = 1$ é dada por

$$p(n_1, n_2, n_3, n_4) = \frac{N!}{n_1!\, n_2!\, n_3!\, n_4!} \times (p_1)^{n_1} \times (p_2)^{n_2} \times (p_3)^{n_3} \times (p_4)^{n_4}$$

$$p(4,3,2,1) = \frac{10!}{4!\,3!\,2!\,1!} \times (0{,}4)^4 \times (0{,}3)^3 \times (0{,}2)^2 \times (0{,}1)^1$$

$$p(4,3,2,1) = 10 \times 9 \times 4 \times 7 \times 5 \times 0{,}0256 \times 0{,}027 \times 0{,}04 \times 0{,}1 \cong 0{,}0348\ (ou\ 3{,}48\%)$$

Resposta: Errado.

(2014 – HC-UFPE – *Adaptada*) Um geneticista acredita que a cor dos olhos, classificada em três categorias (preto, marrom, verde/azul), ocorre, respectivamente, com as seguintes porcentagens: 30%, 50% e 20%. Com base nessa crença, julgue o item a seguir.

A probabilidade de selecionar ao acaso, com reposição, 3 pessoas com olhos pretos, 4 pessoas com olhos marrons e 1 pessoa com olhos verdes/azuis é inferior a 10%.

Comentários:

A probabilidade de $n_1 = 3, n_2 = 4, n_3 = 1$, é dada por

$$p(n_1, n_2, n_3) = \frac{N!}{n_1!\, n_2!\, n_3!} \times (p_1)^{n_1} \times (p_2)^{n_2} \times (p_3)^{n_3}$$

$$p(3{,}4{,}1) = \frac{8!}{4!\, 3!\, 1!} \times (0{,}3)^3 \times (0{,}5)^4 \times (0{,}2)^1$$

$$p(3{,}4{,}1) = 8 \times 7 \times 5 \times 0{,}027 \times 0{,}0625 \times 0{,}2 = 0{,}0945$$

Resposta: Certo.

QUESTÕES COMENTADAS – CEBRASPE

Momentos

1. (Cebraspe/2013 – Agente Nacional de Transportes Terrestres) Considere que a função geradora de momentos de uma variável aleatória discreta X seja dada pela relação $M_X(q) = (0,8 + 0,2e^q)^2$, em que q ∈ R. Com base nessas informações, julgue o item a seguir.

A probabilidade de ocorrência do evento [X = 0] é igual a 0,64.

Comentários:

Pela função geradora de momento descrita no enunciado, $M_X(t) = (0{,}8 + 0{,}2e^t)^2$, podemos concluir que a variável segue distribuição binomial, cuja f.g.m. é:

$$M_{X_{Bi}}(t) = (p.e^t + q)^n$$

Portanto, os parâmetros da distribuição são $p = 0{,}2$ ($q = 1 - 0{,}2 = 0{,}8$) e $n = 2$. A probabilidade $P[X = 0]$ é portanto:

$$P[X = k] = C_{n,k}.p^k.q^{n-k}$$

$$P[X = 0] = C_{2,0}.0{,}2^0.0{,}8^2 = 1 \times 1 \times 0{,}64 = 0{,}64$$

Gabarito: Certo.

2. (CESPE/2012 – TJ-RO) Na sequência $\{X_1, X_2, ..., X_n\}$, de variáveis aleatórias independentes e identicamente distribuídas, cada variável possui a função geradora de momentos na forma $\psi(t) = pe^t + 1 – p$, para $0 < p < 1$. Com base nessas informações, assinale a opção correta.

a) Independentemente do valor de p, a distribuição da soma de $\{X_1, X_2, ..., X_n\}$ será sempre simétrica.

b) A variância dessa variável aleatória é igual à sua média.

c) A variável aleatória é contínua.

d) A variância dessa variável aleatória é maior que a sua média.

e) A função geradora de momentos da soma dessas variáveis é dada por $\psi_S(t) = (pe^t + 1 – p)^n$, $0 < p < 1$.

Comentários:

Pela função geradora de momentos descrita no enunciado, $M_X(t) = pe^t + 1 - p$, podemos concluir que cada variável X_i segue distribuição de **Bernoulli**.

Em relação à alternativa A, a soma das n variáveis X_i segue **distribuição binomial**. Essa distribuição é simétrica para $p = 1 - p = 0{,}5$ e assimétrica para os demais valores de p. Por isso, a alternativa A está incorreta.

Em relação à alternativa E, que também trata da soma das variáveis, sabemos que a f.g.m. de uma variável com distribuição binomial é dada por:

$$M_{X_{Bi}}(t) = (p.e^t + q)^n$$

Em que $q = 1 - p$. Ademais p é a probabilidade de sucesso, ou seja, $0 < p < 1$. Portanto, a alternativa E está correta.

Em relação à alternativa B, a variância de uma variável com distribuição de Bernoulli é:

$$V(X) = p.q$$

E a média é:

$$E(X) = p$$

Logo, a variância não é igual à média e a alternativa B está incorreta.

Em relação à alternativa D, que também compara a variância e a média da distribuição de Bernoulli, sabemos a variância da distribuição é igual a média multiplicada por q (probabilidade de fracasso). Como $q < 1$ (por se tratar de uma probabilidade), a variância é, então, menor do que a média. Logo, a alternativa D está incorreta.

Em relação à alternativa C, sabemos que a variável com distribuição de Bernoulli assume os valores X = 0 ou X = 1, sendo, portanto, uma variável discreta. Por isso, a alternativa C está incorreta.

Gabarito: E

3. (CESPE/2009 – TRT-ES) Considere que Y seja uma variável aleatória de Bernoulli com parâmetro p, em que p é a probabilidade de uma ação judicial trabalhista ser julgada improcedente. De uma amostra aleatória simples de 1.600 ações judiciais trabalhistas, uma seguradora observou que, em média, 20% dessas ações foram julgadas improcedentes.

Com base nessa situação hipotética, julgue os próximos itens.

A estimativa de máxima verossimilhança para a função geratriz de momentos de Y é igual a 0,2 + 0,8exp(t), em que t é um número real e exp(.) denota a função exponencial.

Comentários:

O enunciado informa que a variável Y segue distribuição de Bernoulli com parâmetro p. Assim, a f.g.m. dessa variável é:

$$M_X(t) = pe^t + q$$

O enunciado informa que p é a probabilidade de uma ação judicial ser julgada improcedente. Considerando que 20% das ações foram julgada improcedentes, temos p = 0,2:

$$M_X(t) = 0{,}2.e^t + 0{,}8$$

Essa f.g.m. é diferente de $M_X(t) = 0{,}2 + 0{,}8.e^t$, indicada no item. Logo, o item está errado.

Gabarito: Errado.

4. (Cebraspe/2013 – Estatístico FUB) Considerando o coeficiente de assimetria que se define pelo terceiro momento central, é correto afirmar que a distribuição de Poisson exibe assimetria positiva.

Comentários:

Sabemos que o coeficiente de assimetria em distribuições de Poisson é dado por:

$$b_1 = \frac{1}{\sqrt{\lambda}}$$

Considerando que $\lambda > 0$, então o coeficiente é positivo.

Gabarito: Certo

5. (CESPE/2011 – Estatístico FUB) Considerando-se duas variáveis aleatórias contínuas X e Y, em que X tem função de densidade arbitrária f com função geradora de momentos M(t) e Y = exp(X), julgue o próximo item.

E(Y) = M(1)

Comentários:

A esperança de uma variável pode ser calculada a partir da função geradora de momentos, derivando-a e calculando-a para t = 0:

$$\frac{dM_Y(0)}{dy} = E(Y)$$

O enunciado informa que a f.g.m. dessa variável é a seguinte função da variável X:

$$M_Y(t) = e^X$$

A derivada dessa função em relação a Y é ela mesma:

$$\frac{dM_Y(t)}{dy} = e^X$$

Como essa função não depende de t, quando aplicamos o ponto $t = 0$, a função continua igual. Logo, a esperança de Y é:

$$E(Y) = \frac{dM_Y(0)}{dy} = e^X$$

Ademais, como a f.g.m. de Y, $M_Y(t) = e^X$, não depende de t, então essa f.g.m. no ponto $t = 1$ também é ela mesma:

$$M_Y(1) = e^X$$

Logo, temos $E(Y) = M_Y(1)$.

Gabarito: Certo

QUESTÕES COMENTADAS - CEBRASPE

Distribuições Conjuntas

1. (Cebraspe/2024 – ANTT)

	$Y=2$	$Y=3$	$Y=4$
$X=0$	$\frac{1}{16}$	$\frac{1}{6}$	$\frac{1}{8}$
$X=1$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$
$X=2$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{16}$

Considerando a tabela precedente, que apresenta a função massa de probabilidade conjunta de duas variáveis aleatórias discretas, X e Y, julgue o item que se segue.

As distribuições X e Y são independentes.

Comentários:

Para que duas variáveis sejam independentes, é necessário que a distribuição conjunta seja igual ao produto das distribuições marginais, para todos os possíveis valores das variáveis:

$$P(X=x,Y=y)=P(X=x)\times P(Y=y),\qquad \forall x,y$$

Para verificar, vamos calcular as probabilidades marginais de X = 0 (pela soma da 1ª linha) e de Y = 2 (pela soma da 1ª coluna):

$$P(X=0)=\frac{1}{16}+\frac{1}{6}+\frac{1}{8}=\frac{3+8+6}{48}=\frac{17}{48}$$

$$P(Y=2)=\frac{1}{16}+\frac{1}{16}+\frac{1}{6}=\frac{3+3+8}{48}=\frac{14}{48}=\frac{7}{24}$$

E o produto é:

$$P(X=0)\times P(Y=2)=\frac{17}{48}\times\frac{7}{24}$$

Que é diferente da probabilidade conjunta $P(X=0,Y=2)=\frac{1}{16}$. Assim, verificamos que as variáveis **não** são independentes.

Gabarito: Errado

2. (Cebraspe/2024 – ANTT)

	$Y=2$	$Y=3$	$Y=4$
$X=0$	$\frac{1}{16}$	$\frac{1}{6}$	$\frac{1}{8}$
$X=1$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$
$X=2$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{16}$

Considerando a tabela precedente, que apresenta a função massa de probabilidade conjunta de duas variáveis aleatórias discretas, X e Y, julgue o item que se segue.

$P(X=2)=\frac{19}{48}$.

Comentários:

Precisamos calcular a probabilidade marginal de $X=2$, pela soma da última linha da tabela:

$$P(X=2)=\frac{1}{6}+\frac{1}{6}+\frac{1}{16}=\frac{8+8+3}{48}=\frac{19}{48}$$

Gabarito: Certo

3. (CESPE/2022 – PC-RO) As pessoas que cumprem penas judiciais em determinado município são classificadas como reincidentes (R) ou não reincidentes (R^c). Além disso, essas mesmas pessoas são classificadas de acordo com o tipo do regime no cumprimento da pena: regime fechado (F) ou regime não fechado (F^c).

	R	R^c
F	0,5	0,1
F^c	0,2	0,2

Considerando-se que a tabela precedente mostra a distribuição de probabilidade conjunta dos eventos apresentados na situação hipotética anterior, é correto afirmar que o valor da probabilidade condicional $P(R|F)$ é igual a

a) 5/10.

b) 7/10.

c) 5/7.

d) 5/6.

e) 9/10.

Comentários:

A probabilidade condicional $P(R|F)$ é a razão entre a probabilidade da interseção dos eventos $P(R \cap F)$ e a probabilidade do evento a priori $P(F)$:

$$P(R|F) = \frac{P(R \cap F)}{P(F)}$$

Pela tabela, observamos que a probabilidade da interseção é $P(R \cap F) = 0{,}5$.

E a probabilidade marginal do evento F é a soma da linha correspondente $P(F) = 0{,}5 + 0{,}1 = 0{,}6$.

Logo, a razão é:

$$P(R|F) = \frac{0{,}5}{0{,}6} = \frac{5}{6}$$

Gabarito: D.

4. (CESPE/2022 – FUNPRESP-EXE)

Cliente	produto			Total
	X	Y	Z	
Tipo A	0,10	0,08	0,02	0,20
Tipo B	0,15	0,45	0,20	0,80
Total	0,25	0,53	0,22	1,00

Considerando a tabela precedente, que mostra as preferências de compras dos produtos X, Y e Z dos clientes de uma loja, classificados como tipos A e B, julgue o próximo item.

A probabilidade de um cliente do tipo A entrar ao acaso na loja é de 20%

Comentários:

O enunciado fornece a tabela de distribuição conjunta das variáveis, bem como as probabilidades marginais (totais), indicadas na última linha e na última coluna.

Pela tabela, observamos que a probabilidade marginal associada ao cliente do tipo A é de 20%. Isso significa, que ao deparar-se com um cliente ao acaso, a probabilidade de ele ser do tipo A é de 20%.

Gabarito: Certo.

5. (CESPE/2022 – FUNPRESP-EXE)

Cliente	produto			Total
	X	Y	Z	
Tipo A	0,10	0,08	0,02	0,20
Tipo B	0,15	0,45	0,20	0,80
Total	0,25	0,53	0,22	1,00

Considerando a tabela precedente, que mostra as preferências de compras dos produtos X, Y e Z dos clientes de uma loja, classificados como tipos A e B, julgue o próximo item.

A probabilidade de o produto X ser comprado por um cliente do tipo A é de 10%.

Comentários:

O enunciado pede a probabilidade de o produto X ser comprado por um cliente do tipo A. Em outras palavras, precisamos da probabilidade de compra do produto X, **sabendo**-se que o cliente é do tipo A, isto é, a probabilidade **condicional** de X dado A:

$$P(X|A) = \frac{P(X, A)}{P(A)}$$

Pela tabela de distribuição de probabilidade conjunta, observamos que a probabilidade conjunta de X e A é $P(X, A) = 0{,}10$ e sabemos que a probabilidade marginal de A é $P(A) = 0{,}20$, logo:

$$P(X|A) = \frac{0{,}10}{0{,}20} = 0{,}5$$

A probabilidade é de 50%, e não de 10%.

Gabarito: Errado.

6. (CESPE/2022 – TELEBRÁS)

Supondo que

$$\boldsymbol{P(Y = y|M = m) = \frac{e^{-m}m^{y}}{y!}}$$

para $y \in \{0, 1, 2, 3 \dots\}$, em que $m > 0$, e M é uma variável aleatória contínua cuja função densidade é dada por $f_m(m) = e^{-m}$, julgue o item a seguir.

$Var(Y = y|M = m) = m.$

Comentários:

O enunciado fornece a função de distribuição condicional de Y, dado M = m. Podemos observar que se trata de uma distribuição de Poisson $P(Y = k) = \frac{e^{-\lambda}\lambda^k}{k!}$, com $\lambda = m$ e $k = y$. Sabendo que a variância de Poisson é igual ao parâmetro λ e que, nessa distribuição, temos $\lambda = m$, então a variância dessa distribuição é:

$$V(Y|M = m) = \lambda = m$$

Gabarito: Certo

7. (CESPE/2022 – TELEBRÁS)

Supondo que

$$\boldsymbol{P(Y = y|M = m) = \frac{e^{-m}m^y}{y!}}$$

para $y \in \{0, 1, 2, 3 \dots\}$, em que $m > 0$, e M é uma variável aleatória contínua cuja função densidade é dada por $f_m(m) = e^{-m}$, julgue o item a seguir.

Y e M são variáveis aleatórias independentes.

Comentários:

O enunciado fornece a função de distribuição condicional de Y, dado $M = m$. Podemos observar que essa função **depende** de $M = m$. Por exemplo, se $M = m_1$, teremos uma função de probabilidade para Y e se $M = m_2$, teremos outra função de probabilidade para Y

$$P(Y = y|M = m_1) \neq P(Y = y|M = m_2)$$

Isso nos permite concluir que a distribuição de probabilidade de Y condicionada a $M = m$ é diferente da distribuição não condicionada:

$$P(Y = y|M = m_1) \neq P(Y = y)$$

Assim, concluímos que as variáveis são **dependentes**.

Gabarito: Errado

8. (CESPE/2022 – TELEBRÁS - *Adaptada*)

Supondo que

$$\boldsymbol{P(Y = y|M = m) = \frac{e^{-m}m^y}{y!}}$$

para $y \in \{0, 1, 2, 3 \dots\}$, em que $m > 0$, e M é uma variável aleatória contínua cuja função densidade é dada por $f_m(m) = m.e^{-m}$, julgue o item a seguir.

$P(Y > 0|M = m) = P(M \leq m)$.

Comentários:

O enunciado fornece a função de distribuição condicional de Y, dado M = m, que corresponde a uma distribuição de Poisson, com parâmetro m. A probabilidade de Y ser maior que 0, dado que M =m, pode ser calculada pela probabilidade complementar, sabendo que a variável de Poisson assume os valores {0, 1, 2, ...}:

$$P(Y > 0|M = m) = 1 - P(Y = 0|M = m)$$

A probabilidade de Y ser igual a 0, dado que o parâmetro é m, é:

$$P(Y = 0) = \frac{e^{-m}m^0}{0!} = \frac{e^{-m} \times 1}{1} = e^{-m}$$

Logo, a probabilidade de Y ser maior que 0, dado o parâmetro m é o complemento:

$$P(Y > 0|M = m) = 1 - e^{-m}$$

Para calcular a probabilidade $P(M \leq m)$, devemos observar que a função densidade da variável M corresponde a uma distribuição **exponencial** com parâmetro $m.$

A função **distribuição acumulada** dessa variável, que corresponde à probabilidade $P(M \leq m)$ é:

$$P(M \leq m) = F(m) - 1 - e^{-m}$$

Portanto, de fato, essas duas probabilidades são iguais.

Gabarito: Certo

9. (CESPE/2022 – ANP) Uma curva de regressão da variável aleatória Y sobre $X = x$ é dada por $E[Y|X = x] = 1 - x$, em que o par de variáveis aleatórias (X, Y) segue uma distribuição normal bivariada, a média de X é igual a zero, $Var[Y] = 4$ e $Var[X] = 1$.

Considerando essa situação hipotética, julgue o item a seguir.

A média de Y é igual a 1.

Comentários:

Para calcular a média de Y, vamos utilizar a Lei das Expectativas Iteradas, sabendo que $E[Y|X = x] = 1 - x$:

$$E(Y) = E_X[E(Y|X = x)] = E_X(1 - X) = 1 - E(X)$$

Sabendo que $E(X) = 0$, temos:

$$E(Y) = 1$$

Gabarito: Certo

10. (CESPE/2022 – ANP) Uma curva de regressão da variável aleatória Y sobre $X = x$ é dada por $E[Y|X = x] = 1 - x$, em que o par de variáveis aleatórias (X, Y) segue uma distribuição normal bivariada, a média de X é igual a zero, $Var[Y] = 4$ e $Var[X] = 1$. Considerando essa situação hipotética, julgue o item a seguir.

A correlação linear entre X e Y é igual a -1.

Comentários:

O coeficiente de correlação linear é a razão entre a covariância e os desvios padrão das variáveis:

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)}.\sqrt{Var(Y)}}$$

Por sua vez, a covariância é dada por:

$$Cov(X, Y) = E(XY) - E(X).E(Y)$$

O enunciado informa que $E(X) = 0$, logo:

$$Cov(X, Y) = E(XY)$$

Para calcular $E(XY)$, vamos utilizar a Lei das Expectativas Iteradas, uma vez que o enunciado informa que $E[Y|X = x] = 1 - x$:

$$E(XY) = E_X[E(XY|X)] = E_X[X.E(Y|X)] = E_X[X.(1 - X)] = E_X[X - X^2] = E(X) - E(X^2)$$

Sabendo que $E(X) = 0$, temos:

$$E(XY) = -E(X^2)$$

E calculamos $E(X^2)$ a partir da variância, sabendo que $Var(X) = 1$:

$$Var(X) = E(X^2) - [E(X)]^2$$

$$1 = E(X^2) - 0^2$$

$$E(X^2) = 1$$

E a covariância é:

$$Cov(X,Y) = E(XY) = -E(X^2) = -1$$

Agora, podemos calcular o coeficiente de correlação:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}.\sqrt{Var(Y)}} = -\frac{1}{\sqrt{1}.\sqrt{4}} = -\frac{1}{2}$$

Que é diferente de -1.

Gabarito: Errado

11. (CESPE/2022 – ANP) Uma curva de regressão da variável aleatória Y sobre $X = x$ é dada por $E[Y|X = x] = 1 - x$, em que o par de variáveis aleatórias (X, Y) segue uma distribuição normal bivariada, a média de X é igual a zero, $Var[Y] = 4$ e $Var[X] = 1$.

Considerando essa situação hipotética, julgue o item a seguir.

$Var[X + Y] < 5$.

Comentários:

A variância da soma é dada por:

$$Var(X + Y) = Var(X) + Var(Y) + 2. Cov(X,Y)$$

Na questão anterior, calculamos $Cov(X,Y) = -1$. Sabendo que $Var(X) = 1$ e que $Var(Y) = 4$, temos:

$$Var(X + Y) = 1 + 4 + 2 \times (-1) = 5 - 2 = 3$$

Que, de fato, é menor que 5.

Gabarito: Certo

12. (CESPE/2022 – TRT/8ª Região) O número N de pessoas que chegam diariamente em uma fila de atendimento segue uma distribuição de Poisson com média igual a 500. Se Y representa o número diário dessas pessoas que recebem atendimento preferencial, e se

$$P(Y = y|N = n) = \binom{n}{y} 0,1^y 0,9^{n-y}$$

em que $0 \leq y \leq n$, então Var(Y) é igual a

a) 0,1

b) 9

c) 45

d) 50

e) 500

Comentários:

A variância marginal pode ser calculada a partir da esperança condicional e da variância condicional:

$$V(Y) = E[V(Y|N)] + V[E(Y|N)]$$

Pela função de probabilidade fornecida, observamos que $Y|N = n$ segue distribuição binomial com parâmetros n e probabilidade de sucesso $p = 0{,}1$.

A esperança e a variância condicionadas a n são, respectivamente:

$$E(Y|N) = n \times p = 0{,}1.n$$

$$V(Y|N) = n \times p \times (1 - p) = n \times 0{,}1 \times 0{,}9 = 0{,}09.n$$

Assim, a esperança da variância condicionada é dada por:

$$E[V(Y|N)] = E[0{,}09.n] = 0{,}09.E(N)$$

O enunciado informa que N é uma variável com distribuição de Poisson com média igual a 500, logo:

$$E[V(Y|N)] = 0{,}09 \times 500 = 45$$

E a variância da esperança condicionada é dada por:

$$V[E(Y|N)] = V[0{,}1.n] = 0{,}1^2.V(N)$$

Sabendo que, na distribuição de Poisson, a média e a variância são iguais, temos:

$$V[E(Y|N)] = 0{,}01 \times 500 = 5$$

E a soma corresponde à variância de Y:

$$V(Y) = 45 + 5 = 50$$

Gabarito: D

13. (CESPE/2022 – TRT/8ª Região) Considere duas variáveis X_1 e X_2 tais que:

$$\boldsymbol{E(X_1) = E(X_2) = 1 \quad e \quad Var(X_1) = Var(X_2) = 4}$$

Se X_1 e X_2 forem variáveis aleatórias independentes, então a variância do produto X_1e X_2 será igual a:

a) 4

b) 8

c) 16

d) 18

e) 24

Comentários:

A variância do produto de duas variáveis aleatórias é dada por:

$$Var(X_1.X_2) = E[(X_1X_2)^2] - [E(X_1X_2)]^2$$

Para variáveis independentes, temos $E[(X_1X_2)^2] = E(X_1^2).E(X_2^2)$ e $E(X_1X_2) = E(X_1).E(X_2)$, logo:

$$Var(X_1X_2) = E(X_1^2).E(X_2^2) - [E(X_1).E(X_2)]^2$$

Vamos calcular $E(X_1^2)$ e $E(X_2^2)$, a partir das variâncias:

$$Var(X_1) = E(X_1^2) - [E(X_1)]^2$$

Sabendo que $Var(X_1) = 4$ e $E(X_1) = 1$, temos:

$$E(X_1^2) = 4 + 1^2 = 5$$

Como a média e a variância de X_2 são iguais às de X_1, o resultado será igual: $E(X_2^2) = 5$. Logo:

$$Var(X_1X_2) = 5 \times 5 - (1 \times 1)^2 = 25 - 1 = 24$$

Gabarito: E

14. (CESPE/2021 – TCE-RJ)

		B	
		Sim	Não
A	Sim	22	8
	Não	8	12

Considerando que o cruzamento de duas variáveis categorizadas A e B cujos níveis de resposta são "Sim" e "Não" tenha produzido a tabela de contingência precedente, julgue o próximo item.

Se a tabela de contingência em questão for representada como

		Y	
		1	0
X	1	22	8
	0	8	12

na qual X e Y são variáveis quantitativas dicotômicas, o coeficiente de correlação linear de Pearson entre X e Y será igual a 1/3.

Comentários:

Para calcular o coeficiente de correlação linear, vamos começar calculando a tabela de **distribuição conjunta** das variáveis, dividindo os valores pelo total de 50 respostas:

		Y		Total
		1	0	
X	1	0,44	0,16	0,60
	0	0,16	0,24	0,40
Total		0,60	0,40	1,00

Vamos primeiro calcular a média de X, considerando as probabilidades marginais indicadas na última coluna:

$$E(X) = \sum x \times p(x) = 1 \times 0{,}60 + 0 \times 0{,}40 = 0{,}60$$

E a média de Y, considerando as probabilidades marginais da variável indicadas na última linha:

$$E(Y) = \sum y \times p(y) = 1 \times 0{,}60 + 0 \times 0{,}40 = 0{,}60$$

Para calcular as variâncias, vamos calcular o segundo momento central das variáveis, multiplicando o quadrado dos valores pelas probabilidades marginais correspondentes:

$$E(X^2) = \sum x^2 \times p(x) = 1^2 \times 0{,}60 + 0^2 \times 0{,}40 = 0{,}60$$

$$E(Y^2) = \sum y^2 \times p(y) = 1^2 \times 0{,}60 + 0^2 \times 0{,}40 = 0{,}60$$

Agora, podemos calcular as variâncias:

$$Var(X) = E(X^2) - [E(X)]^2 = 0{,}60 - (0{,}60)^2 = 0{,}60 - 0{,}36 = 0{,}24$$

$$Var(Y) = E(Y^2) - [E(Y)]^2 = 0{,}60 - (0{,}60)^2 = 0{,}60 - 0{,}36 = 0{,}24$$

Para calcular a covariância, precisamos de $E(X.Y)$. Para isso, multiplicamos os valores das variáveis pela **probabilidade conjunta** correspondente:

$$E(X.Y) = \sum x \times y \times p(x,y) = 1 \times 1 \times 0{,}44 + 1 \times 0 \times 0{,}16 + 0 \times 1 \times 0{,}16 + 0 \times 0 \times 0{,}24 = 0{,}44$$

Agora, podemos calcular a covariância:

$$Cov(X,Y) = E(X.Y) - E(X) \times E(Y) = 0{,}44 - 0{,}6 \times 0{,}6 = 0{,}44 - 0{,}36 = 0{,}08$$

Por fim, calculamos o coeficiente de correlação:

$$\rho = \frac{Cov(X,Y)}{\sqrt{V(X)}.\sqrt{V(Y)}} = \frac{0{,}08}{\sqrt{0{,}24}.\sqrt{0{,}24}} = \frac{0{,}08}{0{,}24} = \frac{1}{3}$$

Gabarito: Certo

15. (CESPE/2020 – ME)

		A			total (%)
		-1	0	1	
B	-1	10	5	5	20
	0	0	60	0	60
	1	10	5	5	20
total (%)		20	70	10	100

Considerando que a tabela precedente mostra o cruzamento de duas variáveis categorizadas A e B, que foram codificadas em três níveis numéricos de resposta: −1, 0 e 1, julgue o item que se segue.

A média da variável A é negativa.

Comentários:

Para calcular a média da variável A, a partir da distribuição conjunta apresentada na tabela, precisamos das probabilidades **marginais** da variável, indicadas na última linha (total). Com elas, podemos construir a seguinte tabela da distribuição de probabilidade de A:

A	-1	0	1
P(A)	20%	70%	10%

Sabendo que a média (ou esperança) marginal de A corresponde ao produto dos seus valores pelas respectivas probabilidades marginais, temos:

$$E(X) = \sum x \times p(x)$$

$$E(A) = -1 \times 0{,}2 + 0 \times 0{,}7 + 1 \times 0{,}1 = -0{,}2 + 0{,}1 = -0{,}1$$

Que, de fato, é negativo.

Gabarito: Certo

16. (CESPE/2020 – ME)

		A			total (%)
		-1	0	1	
B	-1	10	5	5	20
	0	0	60	0	60
	1	10	5	5	20
total (%)		20	70	10	100

Considerando que a tabela precedente mostra o cruzamento de duas variáveis categorizadas A e B, que foram codificadas em três níveis numéricos de resposta: −1, 0 e 1, julgue o item que se segue.

As modas e as medianas das variáveis A e B são iguais a zero.

Comentários:

Para calcular a moda e a mediana das variáveis, também precisamos das probabilidades marginais.

Na questão anterior, vimos que a tabela de distribuição de probabilidades marginais da variável A é:

A	-1	0	1
P(A)	20%	70%	10%

Com base nela, podemos observar que a moda de A é igual a 0, porque esse valor está associado à maior probabilidade (70%).

A mediana de A também é igual a 0, porque a probabilidade dos valores menores ou iguais a 0 (-1 e 0) é de pelo menos 50% (no caso, igual a 90%); e a probabilidade dos valores maiores ou iguais a 0 (0 e 1) também é de pelo menos 50% (no caso, igual a 80%).

Agora, vejamos a distribuição de probabilidades marginais da variável B, conforme os dados indicados na última coluna:

B	-1	0	1
P(B)	20%	60%	20%

Com base nessa tabela, podemos observar que a moda de B é igual a 0, porque esse valor está associado à maior probabilidade (60%).

A mediana de B também é igual a 0, porque a probabilidade dos valores menores ou iguais a 0 (-1 e 0) é de pelo menos 50% (no caso, igual a 80%); e a probabilidade dos valores maiores ou iguais a 0 (0 e 1) também é de pelo menos 50% (no caso, igual a 80%).

Assim, as modas e medianas de A e B, de fato, são iguais a 0.

Gabarito: Certo

17. (Cebraspe/2018 – EBSERH) Todo paciente que chega a determinado posto hospitalar é imediatamente avaliado no que se refere à prioridade de atendimento. Suponha que o paciente seja classificado como "emergente" (Y = 0) ou como "não emergente" (Y = 1), e que as quantidades X, diárias, de pacientes que chegam a esse posto sigam uma distribuição de Poisson com média igual a 20.

Considerando que W represente o total diário de pacientes emergentes, de tal sorte que $P(W = w|X = x) = \binom{x}{w} 0,1^w 0,9^{x-w}$, em que 0 ≤ *w* ≤ *x* e *x* ≥ 0, julgue o item subsequente.

Se, em determinado dia, 10 pacientes forem atendidos nesse posto hospitalar, então a probabilidade de se registrar, entre esses pacientes, exatamente um paciente emergente será igual a 0,1.

Comentários:

Sabendo que X representa as quantidades diárias de pacientes que chegam ao posto, então, pelos dados fornecidos no item, temos X = 10. Considerando que W representa o total de pacientes **emergentes**, então a probabilidade de termos W = 1 é, de acordo com a função de probabilidade fornecida:

$$P(W = w|X = x) = \binom{x}{w} 0,1^w 0,9^{x-w}$$

$$P(W = 1|X = 10) = \binom{10}{1} 0,1^{10} 0,9^{10-1}$$

$$P(W = 1|X = 10) = \frac{10!}{1!\,9!} 0,1^{10} 0,9^9 = 10 \times 0,1^{10} 0,9^9 \cong 3,8 \times 10^{-10}$$

Esse é um valor muito pequeno, muito menor que 0,1.

Gabarito: Errado.

18. (Cebraspe/2013 – TRT 17ª Região) No que se refere a distribuições discretas, julgue o seguinte item.

Para a distribuição conjunta $P(N = n, X = x) = \frac{1}{6} \times \binom{n}{x} p^x (1-p)^{n-x}$, em que n = 1, ..., 6 e x é uma contagem, as variáveis N e X são dependentes.

Comentários:

Podemos observar que a variável X segue uma distribuição binomial, com parâmetros p e N e sabemos que N pode assumir os valores n = 1, 2, ... 6. Logo, a probabilidade de X = x depende do valor de N.

Por exemplo, a probabilidade P(X = 0) para n = 1 terá probabilidade q = 1 – p e a probabilidade P(X = 0) para n = 2 terá probabilidade q^2.

Ademais, a probabilidade P(X = 6) para n = 6 será p^6, enquanto que, para os demais valores de n, a probabilidade P(X = 6) será nula. Portanto, temos:

$$P(X = x|N = n) \neq P(X = x)$$

Dessa forma, as variáveis **não** são **independentes**.

Gabarito: Certo

19. (Cebraspe/2013 – TRT 17ª Região) Considerando o conceito de distribuição de probabilidade, julgue o item a seguir.

Considere que, em um tribunal, os processos sejam classificados como urgentes (T1) e não urgentes (T2) e que os não urgentes sejam reclassificados como importantes (T2.1) ou não importantes (T2.2). Considere-se, ainda, que a proporção de processos do tipo T1 seja 0,5 e que, entre os processos do tipo T2, 0,2 sejam do tipo T2.1 e 0,8 do tipo T2.2. Se X, Y e Z forem, respectivamente, as contagens de processos de tipos T1, T2.1 e T2.2 em determinado momento, então a distribuição conjunta de (X, Y, Z) é uma multinomial com parâmetros 0,5, 0,1 e 0,4.

Comentários:

Considerando que há mais do que dois resultados possíveis, não temos uma distribuição binomial, mas sim uma distribuição multinomial. Sabendo que X está associado aos processos urgentes (T1), então $p_x = 0,5$ e a probabilidade dos processos não urgentes (T2) é igual a $1 – p_x = 0,5$. Considerando que Y está associada aos processos não urgentes e importantes e que corresponde 0,2 dos processos não urgentes, então $p_y = 0,2 \times 0,5 = 0,1$. Considerando que Z está associada aos processos não urgentes e não importantes e que corresponde a 0,8 dos processos não urgentes, então $p_z = 0,8 \times 0,5 = 0,4$.

Gabarito: Certo

20. (Cebraspe/2013 – CNJ) Uma máquina de café expresso precisa ser reiniciada algumas vezes durante o dia, devido ao uso excessivo. A tabela abaixo mostra a distribuição de probabilidade conjunta do número de vezes que ela é reiniciada na parte da manhã (M) e na parte da tarde (T).

M\T	0	1	2
0	0,10	0,10	0,30
1	0,04	0,06	0,12
2	0,06	0,10	x

Considerando essa tabela, julgue o próximo item.

O número médio de vezes que a máquina é reiniciada na parte da manhã é inferior a 1.

Comentários:

O número médio de vezes que a máquina é reiniciada na parte da manhã é a esperança da variável representada nas linhas. Para calculá-la, precisamos do valor de x primeiro. Esse valor pode ser obtido, considerando que o somatório dos campos da tabela de probabilidade conjunta é 1:

$$0{,}1 + 0{,}1 + 0{,}3 + 0{,}04 + 0{,}06 + 0{,}12 + 0{,}06 + 0{,}1 + x = 1$$

$$0{,}88 + x = 1$$

$$x = 0{,}12$$

A média pode ser calculado diretamente pela tabela da distribuição conjunta, multiplicando os valores de probabilidade da tabela pelo respectivo valor da variável (que podemos chamar de X), sem termos que calcular as probabilidades marginais, quais sejam, P(X = 0), P(X = 1) e P(X = 2):

$$E(X) = \sum_i \sum_j x_i . p(x_i, y_j)$$

$$E(X) = 0 \times 0{,}1 + 0 \times 0{,}1 + 0 \times 0{,}3 + 1 \times 0{,}04 + 1 \times 0{,}06 + 1 \times 0{,}12 + 2 \times 0{,}06 + 2 \times 0{,}1 + 2 \times 0{,}12$$

$$E(X) = 0{,}22 + 0{,}12 + 0{,}2 + 0{,}24 = 0{,}78$$

Logo, a média é, de fato, **inferior a 1**.

Gabarito: Certo.

21. (Cebraspe/2013 – CNJ) Uma máquina de café expresso precisa ser reiniciada algumas vezes durante o dia, devido ao uso excessivo. A tabela abaixo mostra a distribuição de probabilidade conjunta do número de vezes que ela é reiniciada na parte da manhã (M) e na parte da tarde (T).

M\T	0	1	2
0	0,10	0,10	0,30
1	0,04	0,06	0,12
2	0,06	0,10	x

Considerando essa tabela, julgue o próximo item.

O número de vezes que a máquina é reiniciada na parte da tarde depende do número de vezes que ela é reiniciada pela manhã.

Comentários:

Essa questão indaga quanto à **dependência** entre as variáveis aleatórias. Para verificar isso, é importante calcular as probabilidades marginais, em que X corresponde à manhã e Y, à tarde.

Sabendo que $x = 0{,}12$, como calculamos no item anterior, temos:

$$P(X = 0) = 0{,}1 + 0{,}1 + 0{,}3 = 0{,}50$$

$$P(X = 1) = 0{,}04 + 0{,}06 + 0{,}12 = 0{,}22$$

$$P(X = 2) = 0{,}06 + 0{,}1 + 0{,}12 = 0{,}28$$

$$P(Y = 0) = 0{,}1 + 0{,}04 + 0{,}06 = 0{,}20$$

$$P(Y = 1) = 0{,}1 + 0{,}06 + 0{,}1 = 0{,}26$$

$$P(Y = 2) = 0{,}3 + 0{,}12 + 0{,}12 = 0{,}54$$

Se X e Y são independentes, então o produto das probabilidades será igual à probabilidade conjunta:

$$P(X = x, Y = y) = P(X = x) \times P(Y = y)$$

Para X = 0 e Y = 0, temos $P(X = 0, Y = 0) = 0{,}10$ e o seguinte produto das probabilidades marginais:

$$P(X = 0) \times P(Y = 0) = 0{,}5 \times 0{,}2 = 0{,}10$$

Para X = 1 e Y = 0, temos $P(X = 1, Y = 0) = 0{,}04$ e:

$$P(X = 1) \times P(Y = 0) = 0{,}22 \times 0{,}20 = 0{,}044$$

Como os números da tabela apresentam 2 casas decimais, podemos aproximar $0{,}044 \cong 0{,}04$.

Como $P(X = 1, Y = 0)$ e $P(X = 2, Y = 0)$ seguem a proporcionalidade das probabilidades marginais, então $P(X = 3, Y = 0)$ também seguirá.

Para X = 0 e Y = 1, temos $P(X = 0, Y = 1) = 0{,}10$ e o seguinte produto das probabilidades marginais:

$$P(X = 0) \times P(Y = 1) = 0{,}5 \times 0{,}26 = 0{,}13$$

Como $0{,}13 \neq 0{,}10$, então as variáveis não são independentes, logo, o número de vezes em que a máquina é reiniciada de tarde depende no número de vezes em que ela é reiniciada de manhã (e vice versa).

Gabarito: Certo.

22. (CESPE/2011 – SEDUC/AM)

X	Y			
	-1	0	1	
-1	0	1/5	0	1/5
0	1/5	1/5	1/5	3/5
1	0	1/5	0	1/5
	1/5	3/5	1/5	1

Julgue o próximo item, considerando que o vetor aleatório (X, Y) possui distribuição conjunta de probabilidade conforme o quadro acima.

As variáveis aleatórias Z = X + Y e W = X – Y são dependentes.

Comentários:

As variáveis aleatórias Z e W são independentes se:

$$P(Z = z, W = w) = P(Z = z) \times P(W = w)$$

Sabemos que Z = X + Y e W = X – Y, logo a probabilidade conjunta $P(X + Y = z, X - Y = w)$ exige que ambas as equações sejam atendidas:

$$X + Y = z$$

$$X - Y = w$$

Somando ambas as equações, temos:

$$2X = z + w$$

$$X = \frac{z + w}{2}$$

Sabendo que $Y = X - w$, então:

$$Y = \frac{z + w}{2} - w = \frac{z - w}{2}$$

Logo:

$$P(Z = z, W = w) = P\left(X = \frac{z + w}{2}, Y = \frac{z - w}{2}\right)$$

Podemos, inclusive, dizer que $\frac{z+w}{2} = x$ e $\frac{z-w}{2} = y$:

$$P(Z = z, W = w) = P(X = x, Y = y)$$

Logo, se X e Y forem independentes, então Z e W também serão.

Pela tabela, observamos que as proporções das probabilidades marginais não se mantêm para as probabilidades conjuntas, ou seja:

$$P(X = x, Y = y) \neq P(X = x) \times P(Y = y)$$

Como X e Y são dependentes, Z e W também são.

Gabarito: Certo

23. (CESPE/2011 – SEDUC/AM)

X	Y			
	-1	0	1	
-1	0	1/5	0	1/5
0	1/5	1/5	1/5	3/5
1	0	1/5	0	1/5
	1/5	3/5	1/5	1

Julgue o próximo item, considerando que o vetor aleatório (X, Y) possui distribuição conjunta de probabilidade conforme o quadro acima.

A correlação linear entre X e Y é nula; disso se conclui que ambas são independentes.

Comentários:

Sabemos que o fato de a correlação linear ser nula não implica na independência de X e Y, por isso o item está errado.

Gabarito: Errado.

24. (CESPE/2011 – SEDUC/AM)

$$\boldsymbol{P(X = x, Y = y) = \frac{e^{-\lambda}\lambda^{y}p^{x}(1-p)^{y-x}}{(y-x)!\,x!}, y \geq 0, \quad 0 \leq x \leq y, \quad \lambda > 0, \quad 0 \leq p \leq 1}$$

A respeito da distribuição conjunta (XY), de variáveis aleatórias discretas, apresentada acima, julgue os itens a seguir.

A variância de X é menor que $\lambda . p^4$

Comentários:

Pela distribuição de probabilidade conjunta, podemos observar que Y segue uma distribuição de Poisson com parâmetro λ:

$$P(Y = y) = \frac{e^{-\lambda}\lambda^{y}}{y!}$$

E que X|Y segue uma distribuição Binomial, com parâmetros p e $Y = y$:

$$P(X = x|Y = y) = \binom{y}{x} p^x (1-p)^{y-x} = \frac{y!}{x!\,(y-x)!} p^x (1-p)^{y-x}$$

O produto dessas funções de probabilidade, que corresponde à probabilidade conjunta, resulta na fórmula fornecida no enunciado:

$$P(X = x, Y = y) = P(X = x|Y = y) \times P(Y = y)$$

$$P(X = x, Y = y) = \frac{y!}{x!\,(y-x)!} . p^x . (1-p)^{y-x} . \frac{e^{-\lambda}\lambda^y}{y!}$$

$$P(X = x, Y = y) = \frac{e^{-\lambda}\lambda^y . p^x . (1-p)^{y-x}}{x!\,(y-x)!}$$

A variância da distribuição binomial, V(X), pode ser calculada por a partir da esperança e variância condicionais:

$$V(X) = E[V(X|Y)] + V[E(X|Y)]$$

Como X|Y é uma distribuição binomial, com parâmetros p e $Y = y$, então a variância e a esperança condicionais são:

$$V(X|Y) = n.p.q = Y.p.q = Y.p.(1-p)$$

$$E(X|Y) = n.p = Y.p$$

Então, pelas propriedades da esperança e da variância, temos:

$$V(X) = E[Y.p.q] + V[Y.p] = p.(1-p).E(Y) + p^2 V(Y)$$

Como Y segue uma distribuição de Poisson, com parâmetro λ, então $E(Y) = V(Y) = \lambda$:

$$V(X) = p.(1-p).\lambda + p^2\lambda = p.\lambda - p^2.\lambda + p^2.\lambda = p.\lambda$$

Como $p \leq 1$, então $p \geq p^4$. Logo $V(X) \geq \lambda.p^4$.

Gabarito: Errado.

25. (CESPE/2011 – SEDUC/AM)

$$\boldsymbol{P(X = x, Y = y) = \frac{e^{-\lambda}\lambda^y p^x (1-p)^{y-x}}{(y-x)!\,x!}, y \geq 0, \quad 0 \leq x \leq y, \quad \lambda > 0, \quad 0 \leq p \leq 1}$$

A respeito da distribuição conjunta (XY), de variáveis aleatórias discretas, apresentada acima, julgue o item a seguir.

O valor esperado de X é negativo.

Comentários:

No item anterior, vimos que X segue distribuição binomial, cujo valor esperado é o produto de 2 parâmetros positivos E(X) = n.p, sendo, portanto, **positivo**. Mas vamos calculá-lo, a partir da esperança condicional:

$$E(X) = E[E(X|Y)]$$

Sabemos que X|Y é uma distribuição binomial, com parâmetros p e $Y = y$, logo:

$$E(X|Y) = n.p = Y.p$$

$$E(X) = E[Y.p] = p.E(Y)$$

Como Y segue uma distribuição de Poisson, com parâmetro λ, então $E(Y) = \lambda$:

$$E(X) = \lambda.p$$

Como $\lambda > 0, p \geq 0$, então $E(X) \geq 0$

Gabarito: Errado.

26. (CESPE/2010 – DETRAN/ES) Com relação a distribuições conjuntas, julgue o item que se segue.

Se X seguir uma distribuição de Bernoulli com parâmetro P e P seguir uma distribuição uniforme contínua entre 0 e 1, então a esperança de X será $\frac{3}{4}$.

Comentários:

O item informa que o parâmetro da distribuição de Bernoulli P segue uma distribuição uniforme contínua no intervalo (0,1).

Assim, a esperança marginal de X deve ser calculada a partir da esperança condicional:

$$E(X) = E[E(X|P)]$$

Sabendo que X segue distribuição de Bernoulli, a sua esperança é:

$$E(X|P) = P$$

Sabendo que P segue distribuição uniforme no interval (0,1), a sua esperança corresponde à média aritmética entre os extremos:

$$E(X) = E[P] = \frac{0+1}{2} = \frac{1}{2}$$

Logo, a esperança **não** é igual a ¾.

Gabarito: Errado.

27. (CESPE/2010 – INMETRO) As variáveis discretas X e Y possuem a distribuição conjunta na forma:

P(X = x, Y = y) = d(x + 2y), em que x ∈ {0, 1, 2} e y ∈ {0, 1, 2, 3, 4}.

Para os casos em que x ∉ {0, 1, 2} ou y ∉ {0, 1, 2, 3, 4}, P(X = x, Y = y) = 0.

Com base nessas informações, assinale a opção correta.

a) P(*X*=2,*Y*=3)=0,10.

b) *d* = 1/75.

c) As variáveis aleatórias *X* e *Y* são independentes.

d) P(*X*≥1,*Y*≥3)=7/15.

e) Com respeito à distribuição marginal, tem-se que P(*X* = 0) = 3/75.

Comentários:

Vamos representar a distribuição conjunta pela seguinte tabela:

X\Y	0	1	2	3	4
0	d.(0+2.0)=0	d.(0+2.1)=2d	d.(0+2.2)=4d	d.(0+2.3)=6d	d.(0+2.4)=8d
1	d.(1+2.0)=d	d.(1+2.1)=3d	d.(1+2.2)=5d	d.(1+2.3)=7d	d.(1+2.4)=9d
2	d.(2+2.0)=2d	d.(2+2.1)=4d	d.(2+2.2)=6d	d.(2+2.3)=8d	d.(2+2.4)=10d

Sabemos que a soma de todos os campos é igual a 1, logo:

0 + d + 2d + 2d + 3d + 4d + 4d + 5d + 6d + 6d + 7d + 8d + 8d + 9d + 10d = 75d = 1

$$d = \frac{1}{75}$$

Assim, a resposta é a alternativa B. Porém, vejamos as demais alternativas.

Em relação à alternativa A, temos:

$$P(X = 2, Y = 3) = 8d = \frac{8}{75}$$

Logo, a alternativa A está incorreta.

Em relação à alternativa C, para que as variáveis sejam independentes, devemos ter:

$$P(X = x, Y = y) = P(X = x) \times P(Y = y)$$

Para $X = 0$ e $Y = 0$, por exemplo, temos:

$$P(X=0)=0+2d+4d+6d+8d=20d=\frac{4}{15}$$

$$P(Y=0)=0+d+2d=3d=\frac{3}{75}=\frac{1}{25}$$

Logo:

$$P(X=0)\times P(Y=0)=\frac{4}{15}\times\frac{1}{25}=\frac{4}{375}$$

Como $P(X=0,Y=0)=0$, então temos $P(X=0,Y=0)\neq P(X=0)\times P(Y=0)$. Logo, as variáveis **não** são **independentes**. Assim, a alternativa C está errada.

Sabendo que $P(X=0)=\frac{4}{15}$, podemos concluir que a alternativa E está errada.

Em relação à alternativa D, temos:

$$P(X\geq 1,Y\geq 3)=P(X=1,Y=3)+P(X=1,Y=4)+P(X=2,Y=3)+P(X=2,Y=4)$$

$$P(X\geq 1,Y\geq 3)=7d+9d+8d+10d=34d=\frac{34}{75}$$

Assim, a alternativa D está errada.

Gabarito: B

QUESTÕES COMENTADAS – CEBRASPE

Densidade Conjunta

1. (CEBRASPE/2024 – CAPES) Julgue o próximo item, com base na distribuição de probabilidade condicional $P(X = x|W = w) = \frac{e^{-w}w^x}{x!}$, em que $x = 0, 1, 2, 3, \ldots, w > 0$ e W segue uma distribuição exponencial com média igual a 1.

$E(X|W = w) = w$.

Comentários:

Pela função de distribuição condicional fornecida no enunciado, podemos observar que a distribuição condicional da variável X dado um valor w para a variável W segue uma distribuição de Poisson com parâmetro w. A média da distribuição de Poisson corresponde ao próprio parâmetro, logo:

$$E(X|W = w) = w$$

Gabarito: Certo

2. (CEBRASPE/2024 – CAPES) Julgue o próximo item, com base na distribuição de probabilidade condicional $P(X = x|W = w) = \frac{e^{-w}w^x}{x!}$, em que $x = 0, 1, 2, 3, \ldots, w > 0$ e W segue uma distribuição exponencial com média igual a 1.

O valor esperado da variável aleatória X é igual a w.

Comentários:

A esperança marginal de X pode ser calculada a partir da esperança condicional de X dado W:

$$E(X) = E_W[E(X|W)]$$

No item anterior, vimos que a esperança condicional de X dado W é $E(X|W = w) = w$, logo:

$$E(X) = E_W[W]$$

O enunciado informa que W é uma variável exponencial com média (ou esperança) $E(W) = 1$, logo:

$$E(X) = E_W[W] = 1$$

Que é diferente do resultado indicado no item.

Gabarito: Errado

3. (CEBRASPE/2024 – CAPES) Julgue o próximo item, com base na distribuição de probabilidade condicional $P(X = x|W = w) = \frac{e^{-w}w^x}{x!}$, em que $x = 0, 1, 2, 3, \dots, w > 0$ e W segue uma distribuição exponencial com média igual a 1.

$E[Var(X|W)] = W$.

Comentários:

Nos itens anteriores, vimos que $X|W = w$ segue uma distribuição de Poisson com parâmetro w. Sabendo que a variância da distribuição de Poisson é igual ao próprio parâmetro, temos:

$$Var(X|W) = W$$

Portanto:

$$E[Var(X|W)] = E(W)$$

Sabendo que W segue distribuição exponencial com média $E(W) = 1$, temos:

$$E[Var(X|W)] = E(W) = 1$$

Que é diferente do resultado indicado no item.

Gabarito: Errado

4. (CEBRASPE/2024 – CAPES) Julgue o próximo item, com base na distribuição de probabilidade condicional $P(X = x|W = w) = \frac{e^{-w}w^x}{x!}$, em que $x = 0, 1, 2, 3, \dots, w > 0$ e W segue uma distribuição exponencial com média igual a 1.

$Var(X) = 1$.

Comentários:

A variância marginal pode ser calculada a partir da esperança e da variância condicionadas como:

$$Var(X) = E[Var(X|W)] + Var[E(X|W)]$$

No item anterior, calculamos a primeira expressão, sabendo que $X|W$ segue distribuição de Poisson com parâmetro w, que corresponde à própria variância da distribuição, e que W segue distribuição exponencial com média $E(W) = 1$:

$$E[Var(X|W)] = E(W) = 1$$

Também vimos que a esperança de $X|W$ é igual ao parâmetro w da distribuição de Poisson:

$$Var[E(X|W)] = Var(W)$$

Sabendo que W segue distribuição exponencial, com média $E(W) = 1$ e que a variância é igual ao quadrado da média (enquanto a média é o inverso do parâmetro, a variância é o inverso do quadrado do parâmetro), então $Var(W) = 1^2 = 1$:

$$Var[E(X|W)] = Var(W) = 1$$

Agora, podemos calcular a variância marginal de X:

$$Var(X) = E[Var(X|W)] + Var[E(X|W)] = 1 + 1 = 2$$

Que é diferente de 1.

Gabarito: Errado

5. (CEBRASPE/2024 – CAPES) Julgue o próximo item, com base na distribuição de probabilidade condicional $P(X = x|W = w) = \frac{e^{-w}w^x}{x!}$, em que $x = 0, 1, 2, 3, \dots, w > 0$ e W segue uma distribuição exponencial com média igual a 1.

$P(X = 3) = \frac{1}{16}$.

Comentários:

Para calcular a probabilidade marginal, a partir da probabilidade condicional, primeiro calculamos a função densidade conjunta, a partir da definição de probabilidade condicional:

$$P(X = 3|W = w) = \frac{f(X = 3, W = w)}{f(W = w)}$$

$$f(X = 3, W = w) = P(X = 3|W = w) \times f(w)$$

E a probabilidade marginal de X = 3 corresponde à integral da função densidade conjunta em relação a W, em todo o intervalo $w > 0$:

$$P(X = 3) = \int_0^\infty f(X = 3, W = w).dw = \int_0^\infty P(X = 3|W = w) \times f(w).dw$$

Sabendo que W segue distribuição exponencial com média igual a 1, ou seja, parâmetro $\lambda = \frac{1}{E(X)} = 1$, então a função densidade de probabilidade é:

$$f(w) = \lambda.e^{-\lambda.w} = e^{-w}$$

Considerando a função de probabilidade condicional fornecida, temos, para $X = 3$:

$$P(X = x|W = w) = \frac{e^{-w}.w^x}{x!}$$

$$P(X = 3|W = w) = \frac{e^{-w}.w^3}{3!}$$

Com isso, podemos calcular a probabilidade marginal desejada:

$$P(X = 3) = \int_0^{\infty} \frac{e^{-w}.w^3}{3!}.e^{-w}.dw = \frac{1}{6}\int_0^{\infty} e^{-2w}.w^3.dw$$

Para resolver essa integral, aplicamos a técnica da integração por partes:

$$\int u.dv = u.v - \int v.du$$

Em que $u = w^3$, logo $du = 3.w^2.dw$; e $dv = e^{-2w}$, logo[1] $v = \int e^{-2w}.dw = e^{-2w}.\left(-\frac{1}{2}\right)$:

$$\int e^{-2w}.w^3.dw = w^3.e^{-2w}.\left(-\frac{1}{2}\right) - \int e^{-2w}.\left(-\frac{1}{2}\right).3.w^2.dw = -\boldsymbol{\frac{1}{2}.w^3.e^{-2w}} + \boldsymbol{\frac{3}{2}\int e^{-2w}.w^2.dw}$$

Para resolver essa integral, aplicamos novamente a integração por partes, em que $u = w^2$, $du = 2.w.dw$; e $dv = e^{-2w}$, $v = e^{-2w}.\left(-\frac{1}{2}\right)$:

$$\int e^{-2w}.w^2.dw = w^2.e^{-2w}.\left(-\frac{1}{2}\right) - \int e^{-2w}.\left(-\frac{1}{2}\right).2.w.dw = -\boldsymbol{\frac{1}{2}.w^2.e^{-2w}} + \boldsymbol{\int e^{-2w}.w.dw}$$

Para resolver essa integral, aplicamos, pela última vez, a integração por partes, em $u = w$, logo $du = dw$; e $dv = e^{-2w}$, logo $v = e^{-2w}.\left(-\frac{1}{2}\right)$:

$$\int e^{-2w}.w.dw = w.e^{-2w}.\left(-\frac{1}{2}\right) - \int e^{-2w}.\left(-\frac{1}{2}\right).dw = -\boldsymbol{\frac{1}{2}.w.e^{-2w}} + \boldsymbol{\frac{1}{2}\int e^{-2w}.dw}$$

Sabemos que essa última integral é $\int e^{-2w}.dw = \boldsymbol{e^{-2w}}.\left(-\boldsymbol{\frac{1}{2}}\right)$.Por fim, consideramos todas essas expressões, ignorando os limites, nesse momento:

$$P(X = 3) = \frac{1}{6}\int e^{-2w}.w^3.dw = \frac{1}{6}\left(-\frac{1}{2}.w^3.e^{-2w} + \frac{3}{2}\int e^{-2w}.w^2.dw\right)$$

$$P(X = 3) = \frac{1}{6}\left(-\frac{1}{2}.w^3.e^{-2w} + \frac{3}{2}\left(-\frac{1}{2}.w^2.e^{-2w} + \int e^{-2w}.w.dw\right)\right)$$

$$P(X = 3) = \frac{1}{6}\left(-\frac{1}{2}.w^3.e^{-2w} + \frac{3}{2}\left(-\frac{1}{2}.w^2.e^{-2w} - \frac{1}{2}.w.e^{-2w} + \frac{1}{2}\int e^{-2w}.dw\right)\right)$$

$$P(X = 3) = \frac{1}{6}\left(-\frac{1}{2}.w^3.e^{-2w} + \frac{3}{2}\left(-\frac{1}{2}.w^2.e^{-2w} - \frac{1}{2}.w.e^{-2w} + \frac{1}{2}.e^{-2w}.\left(-\frac{1}{2}\right)\right)\right)$$

---

[1] Para confirmar esse resultado, vamos derivar e^{-2w}. Pela regra da cadeia, primeiro derivamos a função externa $e^{f(w)}$, que é igual a ela mesma, e multiplicamos pela derivada da função interna $f(w) = -2w$:

$$\frac{d(e^{-2w})}{dw} = e^{-2w} \times (-2) = -2.e^{-2w}$$

Se multiplicarmos por $-\frac{1}{2}$, temos:

$$\frac{d(-\frac{1}{2}.e^{-2w})}{dw} = \left(-\frac{1}{2}\right) \times (-2.e^{-2w}) = e^{-2w}$$

Por se tratar da operação inversa da derivada, verificamos que a integral de e^{-2w} é igual a $-\frac{1}{2}.e^{-2w}$.

E trabalhamos para simplificar esse resultado:

$$P(X=3)=\frac{1}{6}\left(-\frac{1}{2}.w^3.e^{-2w}+\frac{3}{2}\left[-\frac{1}{2}.w^2.e^{-2w}-\frac{1}{2}.w.e^{-2w}-\frac{1}{4}.e^{-2w}\right]\right)$$

$$P(X=3)=\frac{1}{6}\left(-\frac{1}{2}.w^3.e^{-2w}-\frac{3}{4}.w^2.e^{-2w}-\frac{3}{4}.w.e^{-2w}-\frac{3}{8}.e^{-2w}\right)$$

$$P(X=3)=-\frac{1}{12}.w^3.e^{-2w}-\frac{3}{24}.w^2.e^{-2w}-\frac{3}{24}.w.e^{-2w}-\frac{3}{48}.e^{-2w}$$

Agora, aplicamos os limites[2]:

$$\lim_{w\to\infty}P(X=3)=0$$

$$\lim_{w\to 0}P(X=3)=-\frac{3}{48}\times 1=-\frac{1}{16}$$

A probabilidade desejada corresponde à diferença:

$$P(X=3)=\lim_{w\to\infty}P(X=3)-\lim_{w\to 0}P(X=3)=0-\left(-\frac{1}{16}\right)=\frac{1}{16}$$

Gabarito: Certo

6. (Cebraspe/2022 – TELEBRÁS) Considerando que a função de densidade conjunta do par de variáveis aleatórias (X,Y), seja dada por

$$\boldsymbol{f(x,y)=\begin{Bmatrix}\dfrac{3(1-x^2)}{4}, & se\ |x|\le 1\ e\ 0\le y\le 1;\\ 0, & caso\ contrário\end{Bmatrix},}$$

julgue o próximo item.

$P(|X|\le y|Y=y)=\frac{y(3-y^2)}{2}$, em que $0\le y\le 1$.

Comentários:

[2] A expressão $\lim_{w\to\infty}w.e^{-2w}=\lim_{w\to\infty}\frac{w}{e^{2w}}$ é indeterminada, pois temos a razão ∞/∞. No entanto, e^{2w} tende a infinito mais rapidamente que w, o que torna a expressão igual a zero.

Isso pode ser verificado, aplicando-se o Teorema de L´Hôpital, segundo o qual o limite de uma razão $\frac{f(x)}{g(x)}$ é igual ao limite da razão entre as derivadas $\frac{f\prime(x)}{g\prime(x)}$:

$$\lim_{w\to\infty}\frac{w}{e^{2w}}=\lim_{w\to\infty}\frac{1}{e^{2w}.2}=0$$

Analogamente, temos $\lim_{w\to\infty}w^2.e^{-2w}=0$ e $\lim_{w\to\infty}w^3.e^{-2w}=0$.

O enunciado fornece a função densidade de probabilidade conjunta de X e Y, mas esta **não depende** de Y[3]:

$$f(x,y) = f(x)$$

Considerando que X não depende de Y, então a probabilidade condicionada é igual à **não** condicionada:

$$P(|X| \leq y | Y = y) = P(|X| \leq y) = P(-y \leq X \leq y)$$

Essa probabilidade é calculada integrando-se a função $f(x,y) = f(x)$, em relação a x, no intervalo $[-y, y]$:

$$P(-y \leq X \leq y) = \int_{-y}^{y} \frac{3(1-x^2)}{4} dx = \frac{3}{4} \int_{-y}^{y} (1-x^2) dx$$

$$P(-y \leq X \leq y) = \frac{3}{4}\left([x]_{-y}^{y} - \left[\frac{x^3}{3}\right]_{-y}^{y}\right) = \frac{3}{4}\left([y-(-y)] - \left[\frac{y^3}{3} - \frac{(-y)^3}{3}\right]\right)$$

$$P(-y \leq X \leq y) = \frac{3}{4}\left([y+y] - \left[\frac{y^3}{3} + \frac{y^3}{3}\right]\right) = \frac{3}{4}\left(2y - \frac{2y^3}{3}\right)$$

$$P(-y \leq X \leq y) = \frac{3}{4}\left(\frac{6y - 2y^3}{3}\right) = \frac{3y - y^3}{2} = \frac{y(3-y^2)}{2}$$

Gabarito: Certo

7. (Cebraspe/2022 – TELEBRÁS) Considerando que a função de densidade conjunta do par de variáveis aleatórias (X,Y), seja dada por

$$\boldsymbol{f(x,y) = \begin{cases} \dfrac{3(1-x^2)}{4}, & se\ |x| \leq 1\ e\ 0 \leq y \leq 1; \\ 0, & caso\ contrário \end{cases}},$$

[3] Se preferir utilizar a definição de probabilidade condicional, sem visualizar que a função densidade conjunta independe de y, o primeiro passo é calcular a função marginal de y, integrando a função conjunta, em relação a x, em todo o intervalo [-1, 1]:

$$f(y) = \int f(x,y).dx = \int_{-1}^{1} \frac{3(1-x^2)}{4} dx = \frac{3}{4}\left([x]_{-1}^{1} - \left[\frac{x^3}{3}\right]_{-1}^{1}\right) = \frac{3}{4}\left(2 - \frac{2}{3}\right) = \frac{3}{4}\left(\frac{6-2}{3}\right) = 1$$

A probabilidade condicional de X dado Y $f(x|y)$ é a razão entre a função densidade conjunta e a função marginal de Y:

$$f(x|y) = \frac{f(x,y)}{f(y)} = \frac{\frac{3(1-x^2)}{4}}{1} = \frac{3(1-x^2)}{4}$$

É essa a função (que corresponde à própria função conjunta) que precisa ser integrada, em relação a x, no intervalo [-y, y].

julgue o próximo item.

$E(X) > 0.$

Comentários:

Na questão anterior, vimos que a função densidade de probabilidade conjunta **não depende** de Y, ou seja[4]:

$$f(x,y) = f(x)$$

Assim, para calcularmos a esperança de X, basta multiplicarmos a função fornecida por x e integrar em relação a essa variável em todo o seu intervalo: $|x| \leq 1 \rightarrow -1 \leq x \leq 1$.

$$E(X) = \int x.f(x).dx = \int_{-1}^{1} x.\frac{3(1-x^2)}{4}dx = \frac{3}{4}\int_{-1}^{1}(x-x^3)dx$$

$$E(X) = \frac{3}{4}\left(\left[\frac{x^2}{2}\right]_{-1}^{1} - \left[\frac{x^4}{4}\right]_{-1}^{1}\right) = \frac{3}{4}\left(\left[\frac{1^2}{2} - \frac{(-1)^2}{2}\right] - \left[\frac{1^4}{4} - \frac{(-1)^4}{4}\right]\right)$$

$$E(X) = \frac{3}{4}\left(\left[\frac{1}{2} - \frac{1}{2}\right] - \left[\frac{1}{4} - \frac{1}{4}\right]\right) = \frac{3}{4}(0-0)$$

$$= 0$$

Ou seja, a esperança é **igual** a zero, e não maior.

Gabarito: Errado

8. (Cebraspe/2022 – TELEBRÁS) Considerando que a função de densidade conjunta do par de variáveis aleatórias (X,Y), seja dada por

$$\boldsymbol{f(x,y) = \left\{\begin{matrix} \frac{3(1-x^2)}{4}, & se\ |x| \leq 1\ e\ 0 \leq y \leq 1; \\ 0, & caso\ contrário \end{matrix}\right\},}$$

[4] Se não visualizar a independência da função densidade conjunta em relação a y, podemos calcular a função marginal de X, integrando a função conjunta, em relação a y (como se x fosse constante), em todo o intervalo [0, 1]:

$$f(x) = \int f(x,y).dy = \int_0^1 \frac{3(1-x^2)}{4}dy = \frac{3(1-x^2)}{4}\int_0^1 dy = \frac{3(1-x^2)}{4}(1-0) = \frac{3(1-x^2)}{4}$$

É essa a função (que corresponde à própria função conjunta) que precisa ser multiplicada por x e integrada em relação a essa variável, em seu intervalo [-1, 1], para calcular a sua esperança.

julgue o próximo item.

$Var(Y) = \frac{1}{12}$.

Comentários:

Esse item pede a variância de Y. Para isso, o primeiro passo é encontrar a função densidade de Y, integrando a função densidade conjunta dada no enunciado, em relação a x, em seu intervalo $[-1,1]$:

$$f(y) = \int f(x,y).dx = \int_{-1}^{1} \frac{3(1-x^2)}{4} dx = \frac{3}{4}\int_{-1}^{1}(1-x^2)dx$$

$$f(y) = \frac{3}{4}\left([x]_{-1}^{1} - \left[\frac{x^3}{3}\right]_{-1}^{1}\right) = \frac{3}{4}\left([1-(-1)] - \left[\frac{1^3}{3} - \frac{(-1)^3}{3}\right]\right)$$

$$f(y) = \frac{3}{4}\left([1+1] - \left[\frac{1}{3}+\frac{1}{3}\right]\right) = \frac{3}{4}\left(2-\frac{2}{3}\right) = \frac{3}{4}\left(\frac{6-2}{3}\right) = 1$$

Portanto, Y segue distribuição uniforme no intervalo [0, 1]. A variância dessa distribuição é dada por:

$$Var(Y) = \frac{(b-a)^2}{12} = \frac{(1-0)^2}{12} = \frac{1}{12}$$

Gabarito: Certo

9. (Cebraspe/2022 – TELEBRÁS) Considerando que a função de densidade conjunta do par de variáveis aleatórias (X,Y), seja dada por

$$\boldsymbol{f(x,y) = \begin{cases} \frac{3(1-x^2)}{4}, & se\ |x| \leq 1\ e\ 0 \leq y \leq 1; \\ 0, & caso\ contrário \end{cases}},$$

julgue o próximo item.

$P(Y = y \| X| \leq y) = y$, em que $0 \leq y \leq 1$.

Comentários:

Neste item, o enunciado pede a probabilidade de $Y = y$, dado que $|X| \leq y$. Por se tratar de uma variável contínua, a probabilidade de ela assumir um valor específico (no caso, y), é igual a zero[5].

Gabarito: Errado

10. (Cebraspe/2022 – TELEBRÁS) Considerando que a função de densidade conjunta do par de variáveis aleatórias (X,Y), seja dada por

$$\boldsymbol{f(x,y) = \begin{cases} \dfrac{3(1-x^2)}{4}, & se\ |x| \leq 1\ e\ 0 \leq y \leq 1; \\ 0, & caso\ contrário \end{cases}},$$

julgue o próximo item.

A correlação linear entre as variáveis é positiva.

Comentários:

Nas questões anteriores, vimos que $f(x,y) = f(x)$ e que $f(y) = 1$. Assim, o produto das funções densidade de probabilidade $f(x) \times f(y)$ corresponde à própria função densidade conjunta:

$$f(x) \times f(y) = f(x,y) \times 1 = f(x,y)$$

Assim, podemos concluir que as variáveis são **independentes**. Para variáveis independentes, a correlação linear é **nula**.

Gabarito: Errado

11. (CESPE/2022 – TRT/8ª Região) Considerando que a distribuição conjunta entre duas variáveis aleatórias contínuas X e Y é dada pela expressão

$$\boldsymbol{f(x,y) = \frac{8xy}{3} + x^2}$$

se $(x,y) \in [0,1] \times [0,1]$ e $f(x,y) = 0$, se $(x,y) \notin [0,1] \times [0,1]$, a probabilidade $P(X+Y>1)$ é igual a

a) 2/9

[5] Para calcular $P(Y = y||X| \leq y)$, sem considerar isso, integraríamos a função condicional $f(y|x)$ no intervalo [y, y]. Como o limite superior da integração é igual ao limite inferior, quando aplicássemos os limites, a diferença seria igual a zero.

b) 7/36

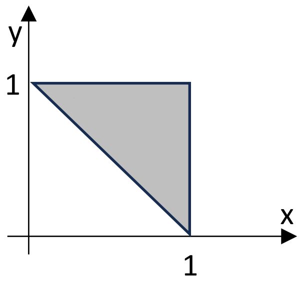
c) 29/36

d) 1/2

e) 7/9

Comentários:

Sabendo que X assume valores no intervalo [0,1] e que Y assume valores no intervalo [0,1], a região em que X + Y > 1 está indicada a seguir:



Assim, para calcular a probabilidade $P(X+Y>1)$, integramos a função no intervalo que corresponde a essa região, em que y varia de 0 a 1 e x varia de $1-y$ a 1:

- $$P(X+Y>1)=\int_0^1\int_{1-y}^1\left(\frac{8xy}{3}+x^2\right)dxdy$$

Calculando primeira a integral interna, em relação a x, temos:

$$\int_{1-y}^1\left(\frac{8y}{3}x+x^2\right)dx=\left[\frac{8y}{3}.\frac{x^2}{2}+\frac{x^3}{3}\right]_{1-y}^1=\frac{4y}{3}[1^2-(1-y)^2]+\frac{1}{3}[1^3-(1-y)^3]$$

Vamos calcular $(1-y)^3$, sabendo que $(1-y)^2=1-2y+y^2$:

$$(1-y)^3=(1-y).(1-2y+y^2)=1-2y+y^2-y+2y^2-y^3=1-3y+3y^2-y^3$$

Assim, o resultado da integral interna é:

$$\int_{1-y}^1\left(\frac{8y}{3}x+x^2\right)dx=\frac{4y}{3}[1-(1-2y+y^2)]+\frac{1}{3}[1-(1-3y+3y^2-y^3)]$$

$$\int_{1-y}^1\left(\frac{8y}{3}x+x^2\right)dx=\frac{4y}{3}[2y-y^2]+\frac{1}{3}[3y-3y^2+y^3)]$$

$$\int_{1-y}^{1} \left(\frac{8y}{3}x + x^2\right) dx = \frac{8}{3}y^2 - \frac{4}{3}y^3 + \frac{3}{3}y - \frac{3}{3}y^2 + \frac{1}{3}y^3 = y + \frac{5}{3}y^2 - y^3$$

Agora, calculamos a integral externa, em relação a y:

$$P(X+Y>1) = \int_0^1 \left(y + \frac{5}{3}y^2 - y^3\right) dy = \left[\frac{y^2}{2} + \frac{5}{3}.\frac{y^3}{3} - \frac{y^4}{4}\right]_0^1$$

$$P(X+Y>1) = \frac{1}{2}(1^2 - 0^2) + \frac{5}{9}(1^3 - 0^3) - \frac{1}{4}(1^4 - 0^4) = \frac{1}{2} + \frac{5}{9} - \frac{1}{4} = \frac{18+20-9}{36} = \frac{29}{36}$$

Gabarito: C

12. (Cebraspe/2021 – PF) Considere que X e Y sejam variáveis aleatórias contínuas que se distribuem conjuntamente conforme a função de densidade.

$$\boldsymbol{f(x,y) = x + y,}$$

Na qual $0 < x < 1$ e $0 < y < 1$.

Com base nessas informações, julgue o item seguinte.

X e Y são variáveis aleatórias independentes.

Comentários:

Duas variáveis contínuas são independentes quando a função densidade de probabilidade conjunta, $f(x,y)$, é o produto das funções densidade de probabilidade marginal das variáveis:

$$f(x,y) = f(x) \times f(y)$$

Ou seja, precisamos calcular as funções densidade de probabilidade marginal das variáveis. A função marginal de x é dada por:

$$f(x) = \int_{-\infty}^{\infty} f(x,y).dy$$

Aqui, estamos integrando em relação a y, ou seja, x é considerada constante. Calculando a integral, sem considerar os limites, temos:

$$\int (x+y).dy = \int x.dy + \int y.dy = x.y + \frac{y^{1+1}}{1+1} = x.y + \frac{y^2}{2}$$

Aplicando os limites, $0 < y < 1$, temos:

$$f(x) = \int_0^1 (x+y).dy = x \times 1 + \frac{1^2}{2} - \left(x \times 0 + \frac{0^2}{2}\right) = x + \frac{1}{2}$$

A função marginal de y é dada por:

$$f(y) = \int_{-\infty}^{\infty} f(x,y).dx$$

Ou seja, temos uma integral em relação a x, em que consideramos y constante. Calculando a integral, sem considerar os limites, temos:

$$\int (x+y).dx = \int x.dx + \int y.dx = \frac{x^{1+1}}{1+1} + y.x = \frac{x^2}{2} + y.x$$

Aplicando os limites, $0 < x < 1$, temos:

$$f(y) = \int_0^1 (x+y).dy = \frac{1^2}{2} + y.1 - \left(\frac{0^2}{2} + y.0\right) = \frac{1}{2} + y$$

O produto das funções marginais é:

$$f(x) \times f(y) = \left(x + \frac{1}{2}\right) \times \left(\frac{1}{2} + y\right) = \frac{1}{2}x + x.y + \frac{1}{4} + \frac{1}{2}y = x.y + \frac{1}{2}(x+y) + \frac{1}{4}$$

Podemos observar que essa função é **diferente** da função densidade conjunta fornecida. Por esse motivo, as variáveis **não** são independentes.

Gabarito: Errado.

13. (Cebraspe/2021 – PF) Considere que X e Y sejam variáveis aleatórias contínuas que se distribuem conjuntamente conforme a função de densidade.

$$\boldsymbol{f(x,y) = x + y,}$$

Na qual $0 < x < 1$ e $0 < y < 1$.

Com base nessas informações, julgue o item seguinte.

Y é uma variável uniforme no intervalo (0,1)

Comentários:

Na questão anterior, calculamos que a função densidade marginal de y é:

$$f(y) = \frac{1}{2} + y$$

Ou seja, quando $y = 0$ a função é $f(y) = \frac{1}{2}$ e quando $y = 1$ a função é $f(y) = \frac{1}{2} + 1$. Logo, o valor da função varia, não sendo, portanto, uniforme.

Gabarito: Errado.

14. (Cebraspe/2021 – PF) Considere que X e Y sejam variáveis aleatórias contínuas que se distribuem conjuntamente conforme a função de densidade.

$$\boldsymbol{f(x, y) = x + y,}$$

Na qual $0 < x < 1$ e $0 < y < 1$.

Com base nessas informações, julgue o item seguinte.

$E\left(X + Y|X = \frac{1}{2}\right) = \frac{14}{12}$.

Comentários:

A questão está pedindo pela esperança de Y condicionada a $X = \frac{1}{2}$. Antes de entrar nos cálculos, podemos aplicar a propriedade aditiva da integral para facilitar as nossas contas:

$$E\left(X + Y|X = \frac{1}{2}\right) = E\left(X|X = \frac{1}{2}\right) + E\left(Y|X = \frac{1}{2}\right)$$

Sabendo que $X = \frac{1}{2}$, a esperança de X é $E\left(X|X = \frac{1}{2}\right) = \frac{1}{2}$:

$$E\left(X + Y|X = \frac{1}{2}\right) = \frac{1}{2} + E\left(Y|X = \frac{1}{2}\right)$$

A esperança de Y, $E\left(Y|X = \frac{1}{2}\right)$, é dada por:

$$E(Y|X = x) = \int_{-\infty}^{\infty} y.f(y|X = x).dy$$

Em que $f(y|x)$ é a função densidade de Y condicionada a X, que é a razão entre a função densidade conjunta e a função marginal de x:

$$f(y|X = x) = \frac{f(x, y)}{f(X = x)}$$

Já calculamos $f(x) = x + \frac{1}{2}$, então, o valor de $f\left(X = \frac{1}{2}\right)$ é:

$$f\left(X = \frac{1}{2}\right) = \frac{1}{2} + \frac{1}{2} = 1$$

Portanto, a função de Y condicionada a $X = \frac{1}{2}$ é:

$$f\left(y|X = \frac{1}{2}\right) = \frac{x+y}{1} = \frac{1}{2} + y$$

Agora, podemos calcular a esperança de Y condicionada a $X = \frac{1}{2}$:

$$E\left(Y|X = \frac{1}{2}\right) = \int_0^1 y.(\frac{1}{2} + y).dy = \int_0^1 \left(\frac{1}{2}.y + y^2\right).dy$$

Calculando as integrais em separado, sem considerar os limites, temos:

$$\int \frac{1}{2}.y.dy = \frac{1}{2} \times \frac{y^{1+1}}{1+1} = \frac{1}{2} \times \frac{y^2}{2} = \frac{y^2}{4}$$

$$\int y^2.dy = \frac{y^{2+1}}{2+1} = \frac{y^3}{3}$$

Juntando as expressões, temos:

$$\int \left(\frac{1}{2}.y + y^2\right).dy = \frac{y^2}{4} + \frac{y^3}{3}$$

Aplicando os limites, temos:

$$E\left(Y|X = \frac{1}{2}\right) = \int_0^1 \left(\frac{1}{2}.y + y^2\right).dy = \frac{1^2}{4} + \frac{1^3}{3} - \left(\frac{0^2}{4} + \frac{0^3}{3}\right) = \frac{3+4}{12} = \frac{7}{12}$$

Por fim, somamos $E\left(X|X = \frac{1}{2}\right) = \frac{1}{2}$ para obter $E\left(X + Y|X = \frac{1}{2}\right)$:

$$E\left(X + Y|X = \frac{1}{2}\right) = \frac{1}{2} + E\left(Y|X = \frac{1}{2}\right) = \frac{6+7}{12} = \frac{13}{12}$$

O resultado é **diferente** de $\frac{14}{12}$.

Gabarito: Errado.

15. (CESPE/2020 – TJ/PA) Uma distribuição condicional é dada por

$P(X = x|Y = y) = y^x(1-y)^{1-x}$, em que $x = 0$ ou $x = 1$ e $0 \leq y \leq 1$.

Considerando-se que Y segue uma distribuição contínua no intervalo [0, 1], é correto afirmar, a respeito da distribuição condicional $Y|X = x$, que $E(Y|X = x)$ é igual a:

a) $\frac{1}{3} + \frac{x}{3}$

b) $\frac{1}{2} + \frac{x}{2}$

c) $\frac{1}{4} + \frac{x}{4}$.

d) $\frac{x}{4}$.

e) $\frac{1}{2}$

Comentários:

O enunciado fornece a função de distribuição de probabilidade condicional de X dado Y = y; e pede a esperança da distribuição condicional de Y dado X = x, invertendo assim os eventos a priori e a posteriori.

A partir da função de distribuição fornecida, podemos calcular a função densidade conjunta das variáveis:

$$f(x, y) = P(X = x|Y = y) \times f(y)$$

Sabendo que $P(X = x|Y = y) = y^x(1-y)^{1-x}$ e que Y segue distribuição uniforme no intervalo [0, 1], ou seja, $f(y) = 1$, então a função densidade conjunta é:

$$f(x, y) = P(X = x|Y = y) \times f(y) = y^x(1-y)^{1-x} \times 1 = y^x(1-y)^{1-x}$$

Para calcular a função densidade condicional de Y dado X = x, dividimos a função conjunta pela função marginal de X = x:

$$f(y|x) = \frac{f(x, y)}{P(X = x)}$$

Por sua vez, a função marginal de X = x é calculada integrando-se a função conjunta, em relação a y, em todo o seu intervalo [0, 1]:

$$P(X = x) = \int_0^1 f(x, y)dy = \int_0^1 y^x(1-y)^{1-x}.dy$$

Assim, a função densidade condicional de Y dado X = x é dada por:

$$f(y|x) = \frac{y^x(1-y)^{1-x}}{\int_0^1 y^x(1-y)^{1-x}.dy}$$

Aqui, temos uma distribuição beta $f(y) = \frac{1}{\mathrm{B}(\alpha,\beta)} y^{\alpha-1}(1-y)^{\beta-1}$, em que:

- $x = \alpha - 1$, logo $\alpha = x + 1$;
- $1 - x = \beta - 1$, logo $\beta = 2 - x$; e
- $\int_0^1 y^x (1-y)^{1-x}.dy$ corresponde à constante de normalização $\mathrm{B}(\alpha,\beta)$.

A média dessa distribuição é:

$$E(Y) = \frac{\alpha}{\alpha+\beta} = \frac{x+1}{x+1+2-x} = \frac{x+1}{3} = \frac{1}{3} + \frac{x}{3}$$

Gabarito: A

16. (CESPE/2010 – MPU) Uma corporação é formada por uma empresa de mineração e uma empresa petrolífera. Considere que X represente as despesas mensais, em milhões de reais, com a recuperação de danos ao meio ambiente causados pela empresa de mineração e que Y represente as despesas mensais, também em milhões de reais, com a recuperação de danos ambientais causados pela empresa petrolífera.

Considere, ainda, que X e Y sejam variáveis aleatórias cuja distribuição conjunta é dada por

$f(x,y) = \frac{xy}{96}$, em que 0 < x < 4 e 0 < y < 5.

Com base nessas informações, julgue o próximo item.

A equação matemática que descreve o valor esperado de X em função de Y depende unicamente de y.

Comentários:

A equação que descreve o valor esperados de X em função de Y corresponde à esperança condicional de X dado Y, denotada por E(X|Y = y). Ela pode ser calculada por:

$$E(X|Y=y) = \int_{-\infty}^{\infty} x.f(x|y).dx$$

Em que $f(x|y)$ é a função densidade condicional de X dado Y = y. Essa função, por sua vez, pode ser calculada a partir da função densidade conjunta $f(x,y)$ fornecida no enunciado:

$$f(x|y) = \frac{f(x,y)}{f(y)}$$

Em que o denominador é a função densidade marginal de Y, a qual também pode ser calculada a partir da função densidade conjunta:

$$f(y) = \int_{-\infty}^{\infty} f(x,y).dx$$

Portanto, a esperança condicional pode ser calculada a partir das informações do enunciado. Por ser uma integral em relação a x, essa variável será substituída ao aplicarmos os limites da integral (no caso, 0 < x < 4). Logo, a esperança condicional de X dado Y = y, de fato, é uma função que depende unicamente de y.

Gabarito: Certo.

17. (CESPE/2010 – MPU) Uma corporação é formada por uma empresa de mineração e uma empresa petrolífera. Considere que X represente as despesas mensais, em milhões de reais, com a recuperação de danos ao meio ambiente causados pela empresa de mineração e que Y represente as despesas mensais, também em milhões de reais, com a recuperação de danos ambientais causados pela empresa petrolífera.

Considere, ainda, que X e Y sejam variáveis aleatórias cuja distribuição conjunta é dada por

$f(x,y) = \frac{xy}{96}$**, em que 0 < x < 4 e 0 < y < 5.**

Com base nessas informações, julgue o próximo item.

O valor esperado de Y é superior a 2,5 milhões de reais.

Comentários:

Podemos calcular o valor esperado de Y a partir da sua função densidade, dada por:

$$f(y) = \int_{-\infty}^{\infty} f(x,y).dx$$

O enunciado informa que $f(x,y) = \frac{xy}{96}$ e que 0 < x < 4, logo:

$$f(y) = \int_0^4 \frac{xy}{96}.dx = \frac{y}{96}\left[\frac{x^2}{2}\right]_0^4 = \frac{y}{192}.(4^2 - 0^2) = \frac{16}{192}y = \frac{1}{12}y$$

E a esperança marginal de Y é:

$$E(y) = \int_{-\infty}^{\infty} y.f(y).dy$$

Sabendo que $f(y) = \frac{1}{12}y$ e que 0 < y < 5, temos:

$$E(Y) = \int_0^5 \frac{1}{12}y^2.dy = \frac{1}{12}\left[\frac{y^3}{3}\right]_0^5 = \frac{1}{36}.(5^3 - 0^3) = \frac{125}{36} \cong 3{,}5$$

Que é superior a 2,5.

Gabarito: Certo.

18. (CESPE/2010 – DETRAN/ES) Com relação a distribuições conjuntas, julgue o item que se segue.

Caso a função de densidade conjunta de X e Y seja dada por F(x,y) = λy.exp(-y(λ + x)), y ≥ 0, então X e Y serão independentes.

Comentários:

O item informa que a função densidade conjunta de X e Y é:

$$f(x,y) = \lambda.y.e^{-y(\lambda+x)}, \qquad com\ y \geq 0$$

Podemos observar que temos duas distribuições exponenciais:

$$f(y) = \lambda.e^{-\lambda y}$$

$$f(x|y) = y.e^{-y.x}$$

O produto dessas funções densidade, que corresponde à densidade conjunta, resulta na fórmula descrita no item:

$$f(x,y) = f(x|y) \times f(y) = y.e^{-y.x} \times \lambda.e^{-\lambda y} = \lambda.y.e^{-y(\lambda+x)}$$

Pela função densidade $f(x|y) = y.e^{-y.x}$, observamos que X é uma variável exponencial com parâmetro y (por isso, o item mencionou $y \geq 0$). Logo, a variável X depende de Y, isto é, X e Y não são independentes.

Gabarito: Errado.

QUESTÕES COMENTADAS – CEBRASPE

Transformação de Variáveis

1. (Cebraspe/2024 – ANTT) Com relação a probabilidade e variáveis aleatórias, julgue o item a seguir.

Se X é uma variável aleatória exponencial de parâmetro unitário e $U = 1 - e^{-X}$, então U é uma variável aleatória uniforme no intervalo [0, 1].

Comentários:

Para resolver essa questão, vamos utilizar o método da Função de Distribuição. Primeiro, vamos verificar se a função $U = h(X) = 1 - e^{-X}$ é crescente ou decrescente. Quanto maior o valor de X, menor o valor de e^{-X} e **maior** o valor de $1 - e^{-X}$. Sendo a função $h(X)$ crescente, calculamos a função da distribuição acumulada de U como:

$$F_U(u) = F_X[h^{-1}(x)]$$

Para obter $h^{-1}(x)$, vamos inverter a função $U = h(X) = 1 - e^{-X}$, isolando X:

$$U = 1 - e^{-X}$$

$$e^{-X} = 1 - U$$

$$-X = \ln(1 - U)$$

$$X = -\ln(1 - U)$$

Essa função deve ser aplicada na função de distribuição de X. Sabendo que essa variável segue distribuição exponencial com parâmetro $\lambda = 1$, a sua função de distribuição acumulada é:

$$F_X(x) = 1 - e^{-\lambda x} = 1 - e^{-x}$$

Aplicando $x = -\ln(1 - u)$ nessa função, temos:

$$F_U(u) = F_X[h^{-1}(x)] = 1 - e^{-(-\ln(1-u))} = 1 - (1 - u) = u$$

Essa função, de fato, corresponde à função de distribuição acumulada de uma variável uniforme no intervalo (0,1), uma vez que a sua derivada $f_U(u) = 1$ corresponde à função densidade de probabilidade dessa variável.

De todo modo, podemos confirmar que U assume valores entre 0 e 1, sabendo que X assume valores entre 0 e ∞: Por um lado, quando $x = 0$, $e^{-x} = 1$ e $u = 1 - e^{-x} = 1$; e, por outro, quando $x \to \infty$, $e^{-x} \to 0$ e $u = 1 - e^{-x} \to 1$. Alternativamente, podemos notar que $U = 1 - e^{-X}$ corresponde à própria função de distribuição acumulada da distribuição exponencial com parâmetro unitário. Sabemos que essa função assume valores no intervalo de 0 a 1 para $0 \leq x < \infty$, como é o caso.

Gabarito: Certo

2. (Cebraspe/2024 – ANTT - *Adaptada*) Com relação a probabilidade e variáveis aleatórias, julgue o item a seguir.

Suponha que X seja uma variável aleatória contínua com função densidade de probabilidade $f(x) = \frac{3}{2}\exp(-3|x|)$, para todo $x \in \mathbb{R}$. Nesse caso, se $Y = X^4$, então a função densidade de probabilidade para Y é $g(y) = \frac{3}{4}.\frac{\exp(-3\sqrt[4]{y})}{\sqrt[4]{y^3}}$ para todo $y > 0$.

Comentários:

Para encontrar a função densidade de probabilidade de $Y = X^4$, vamos utilizar o método Jacobiano:

$$g(y) = f[h^{-1}(y)]\left|\frac{dh^{-1}(y)}{dy}\right|$$

O primeiro passo é inverter a função $y = h(x) = x^4$, isolando x

$$y = x^4$$

$$x = \pm\left(y^{\frac{1}{4}}\right) = \pm\sqrt[4]{y}$$

Utilizamos o sinal $\pm$ porque x pode assumir tanto valores positivos quanto valores negativos. Considerando que $\pm(y^{1/4})$ não é uma função, pois um único valor do domínio (no caso, y) está associado a mais de um valor do contradomínio, precisamos utilizar 2 funções:

$$h_1^{-1}(y) = y^{1/4} = \sqrt[4]{y}, \qquad h_2^{-1}(y) = -y^{\frac{1}{4}} = -\sqrt[4]{y}$$

Dessa forma, a função densidade de Y corresponde à soma:

$$g(y) = f[h_1^{-1}(y)]\left|\frac{dh_1^{-1}(y)}{dy}\right| + f[h_2^{-1}(y)]\left|\frac{dh_2^{-1}(y)}{dy}\right|$$

Vamos agora calcular a derivada das funções inversas:

$$\frac{dh_1^{-1}(y)}{dy} = \frac{dy^{1/4}}{dy} = \frac{1}{4}.y^{\frac{1}{4}-1} = \frac{1}{4}.y^{-\frac{3}{4}} = \frac{1}{4.y^{3/4}} = \frac{1}{4.\sqrt[4]{y^3}}$$

Em relação à função $h_2^{-1}(y) = -y^{1/4}$, obtemos o mesmo resultado, mas com sinal contrário:

$$\frac{dh_2^{-1}(y)}{dy} = -\frac{1}{4.\sqrt[4]{y^3}}$$

Como precisamos do módulo da derivada, o resultado é o mesmo:

$$\left|\frac{dh_1^{-1}(y)}{dy}\right| = \left|\frac{dh_2^{-1}(y)}{dy}\right| = \frac{1}{4.\sqrt[4]{y^3}}$$

Agora, aplicamos as funções inversas na função densidade de $f(x) = \frac{3}{2}\exp(-3|x|)$, no lugar de x:

$$f[h_1^{-1}(y)] = \frac{3}{2}\exp(-3|\sqrt[4]{y}|) = \frac{3}{2}\exp(-3\sqrt[4]{y})$$

$$f[h_2^{-1}(y)] = \frac{3}{2}\exp(-3|-\sqrt[4]{y}|) = \frac{3}{2}\exp(-3\sqrt[4]{y})$$

Como a função densidade $f(x)$ utiliza o módulo de x, o resultado é o mesmo para ambas as inversas.

Agora, podemos encontrar a expressão de $g(y)$:

$$g(y) = \frac{3}{2}\exp\left(-3\sqrt[4]{y}\right)\frac{1}{4.\sqrt[4]{y^3}} + \frac{3}{2}\exp\left(-3\sqrt[4]{y}\right)\frac{1}{4.\sqrt[4]{y^3}} = \frac{3}{4}.\frac{\exp\left(-3\sqrt[4]{y}\right)}{\sqrt[4]{y^3}}$$

Resposta: Certo

3. (CEBRASPE/2024 – CAPES) Considerando que as variáveis aleatórias X e Y sejam normais, mutuamente independentes e identicamente distribuídas, e supondo que μ e σ representem, respectivamente, a média e o desvio padrão dessas distribuições, julgue o item subsequente.

Se $S = X + Y$ e $D = X - Y$, então S e D são variáveis aleatórias independentes.

Comentários:

O enunciado informa que X e Y são independentes e constrói duas novas variáveis a partir delas $S = X + Y$ e $D = X - Y$. A soma em nada interfere na diferença entre variáveis e vice-versa. Em outras palavras, conhecendo a soma, nada podemos dizer a respeito da diferença, ou o contrário. Como as operações são independentes, assim como as variáveis originais X e Y, então as novas variáveis S e D também são independentes.

Gabarito: Certo

4. (CEBRASPE/2024 – CAPES) Uma variável aleatória contínua possui função de distribuição acumulada dada pela expressão a seguir, na qual a é um parâmetro tal que $a \in (0,1)$.

$$F(x) = \begin{Bmatrix} 1 - a^x, & se\ x \geq 0 \\ 0, & se\ x < 0 \end{Bmatrix}$$

A partir dessas informações, julgue o item a seguir.

Na transformação $Y = 1 - a^x$, a variável aleatória Y segue uma distribuição contínua com média 1/2 e variância 1/12.

Comentários:

Vamos resolver essa questão pelo método da Função de Distribuição. Como $h(y) = 1 - a^x$ é crescente, porque corresponde à função de distribuição acumulada de x, que é necessariamente crescente, então a função de distribuição acumulada de Y pode ser calculada como:

$$F_Y(y) = F_X[h^{-1}(y)]$$

Para aplicar essa fórmula, primeiro invertemos a função $h(y)$:

$$y = 1 - a^x$$

$$a^x = 1 - y$$

$$x = \log_a 1 - y$$

Portanto, $h^{-1}(y) = \log_a 1 - y$.

Agora, aplicamos essa expressão, no lugar de x, na função de distribuição acumulada de X:

$$F_Y(y) = F_X[h^{-1}(y)] = 1 - a^{\log_a 1-y} = 1 - (1 - y) = y$$

Essa é a função de distribuição acumulada de Y. A sua função densidade é a derivada:

$$f(y) = \frac{dy}{dy} = 1$$

Que corresponde a uma variável **uniforme** contínua. Como uma função de distribuição acumulada qualquer assume valores entre 0 e 1, então Y, que corresponde à função de distribuição acumulada de X, assume valores entre 0 e 1.

A média dessa distribuição corresponde à média aritmética dos extremos:

$$E(Y) = \frac{0+1}{2} = \frac{1}{2}$$

E a variância é dada por:

$$Var(Y) = \frac{(b-a)^2}{12} = \frac{(1-0)^2}{12} = \frac{1}{12}$$

Gabarito: Certo

5. (Cebraspe/2023 – TJ/ES) Uma amostra aleatória simples de tamanho n é retirada de uma distribuição exponencial X com a função de densidade de probabilidade representada a seguir.

$$\boldsymbol{f(x) = \begin{Bmatrix} 5e^{-5x}, & se\ x \geq 0 \\ 0, & se\ x < 0 \end{Bmatrix}}$$

Representando essa amostra aleatória simples com $X_1, \dots, X_n$, julgue o item subsequente.

A variável $D = e^{-5X}$ segue distribuição uniforme no intervalo [0, 1].

Comentários:

Vamos resolver essa questão pelo método da Função de Distribuição.

O primeiro passo é inverter a função $D = e^{-5X}$. Para isso, vamos aplicar o logaritmo natural:

$$\ln D = \ln e^{-5X}$$

$$\ln D = -5X$$

$$X = -\frac{\ln D}{5}$$

Como a função $D = e^{-5X}$ é decrescente (quanto maior o valor de X, menor o valor de D), a função de distribuição de D é dada por:

$$F_D(d) = 1 - F_X[h^{-1}(x)]$$

Sabendo que x segue distribuição exponencial com parâmetro 5, a sua função de distribuição acumulada é:

$$F_X(x) = 1 - e^{-5x}$$

Aplicando $x = -\frac{\ln d}{5}$ nessa função, temos:

$$F_X[h^{-1}(x)] = 1 - e^{-5\left(-\frac{\ln d}{5}\right)} = 1 - e^{\ln d} = 1 - d$$

E a função de distribuição de D é:

$$F_D(d) = 1 - F_X[h^{-1}(x)] = 1 - (1 - d) = d$$

A derivada de $F(d) = d$ é $f(d) = 1$, que corresponde à função densidade de probabilidade de uma variável uniforme no intervalo (0,1).

Gabarito: Certo

6. (Cebraspe/2023 – TJ/ES) Uma amostra aleatória simples de tamanho n é retirada de uma distribuição exponencial X com a função de densidade de probabilidade representada a seguir.

$$\boldsymbol{f(x) = \begin{Bmatrix} 5e^{-5x}, & se\ x \geq 0 \\ 0, & se\ x < 0 \end{Bmatrix}}$$

Representando essa amostra aleatória simples com $X_1, \dots, X_n$, julgue o item subsequente.

A seguir, é apresentada a função densidade da variável $W = 5X$.

$$f(w) = \begin{Bmatrix} 5e^{-w}, & se\ w \geq 0 \\ 0, & se\ w < 0 \end{Bmatrix}$$

Comentários:

Nessa questão, vamos utilizar o método Jacobiano para calcular a função densidade de probabilidade de w:

$$f_W(w) = f_X[h^{-1}(w)].\left|\frac{dh^{-1}(w)}{dw}\right|$$

O primeiro passo é inverter $W = 5X$:

$$X = \frac{1}{5}W$$

Que corresponde à função $h^{-1}(w)$. A sua derivada é:

$$\frac{dh^{-1}(w)}{dw} = \frac{1}{5}$$

Em seguida, aplicamos $x = \frac{1}{5}w$ na função densidade de x:

$$f[h^{-1}(w)] = 5e^{-5.\left(\frac{1}{5}w\right)} = 5e^{-w}$$

Por fim, a função densidade de w corresponde ao produto:

$$f_W(w) = 5e^{-w}.\frac{1}{5} = e^{-w}$$

Que é diferente do que consta no item.

Gabarito: Errado.

7. (Cebraspe/2022 – TRT/8ª Região) Se X e Y forem duas variáveis aleatórias independentes tais que $X \sim Bernoulli(p = 0,8)$ e $Y \sim Binomial(n = 3, p = 0,8)$, então $P(X^4 + Y = 1)$ é igual a

a) 0,0064

b) 0,0256

c) 0,0512

d) 0,1024

e) 0,4096

Comentários:

Sabemos que X segue distribuição de Bernoulli, com probabilidade de sucesso p = 0,8, assim temos X = 0, com probabilidade q = 1 – p = 0,2 e X = 1 com probabilidade p = 0,8. Assim, a distribuição de X^4 é:

X	X^4	p
0	$0^4 = 0$	0,2
1	$1^4 = 1$	0,8

E sabemos que Y segue distribuição binomial, com n = 3 e p = 0,8, podendo assumir os valores Y = 0, Y = 1, Y= 2 e Y = 3.

Desse modo, para obtermos X^4 + Y = 1, é necessário termos X^4 = 1 e Y = 0 ou X^4 = 1 e Y = 0.

No primeiro caso, sabemos que P(X^4 = 1) = 0,8 e que a P(Y = 0) é dada por:

$$P(Y = 0) = C_{3,0} \times 0{,}8^0 \times 0{,}2^3 = 1 \times 1 \times 0{,}2^3 = 0{,}2^3$$

E a probabilidade da interseção é:

$$P(X^4 = 1, Y = 0) = 0{,}8 \times 0{,}2^3$$

Em relação ao segundo caso, sabemos que P(X^4 = 0) = 0,2 e que a P(Y = 1) é dada por:

$$P(Y = 1) = C_{3,1} \times 0{,}8^1 \times 0{,}2^2 = 3 \times 0{,}8 \times 0{,}2^2$$

E a probabilidade da interseção é:

$$P(X^4 = 0, Y = 1) = 0{,}2 \times 3 \times 0{,}8 \times 0{,}2^2 = 3 \times 0{,}8 \times 0{,}2^3$$

Por serem eventos mutuamente excludentes, a probabilidade da união é a soma:

$$P(X^4 + Y = 1) = P(X^4 = 1, Y = 0) + P(X^4 = 0, Y = 1)$$

$$P(X^4 + Y = 1) = 0{,}8 \times 0{,}2^3 + 3 \times 0{,}8 \times 0{,}2^3 = 4 \times 0{,}8 \times 0{,}2^3 = 0{,}0256$$

Gabarito: B

8. (CEBRASPE/2021 – MJSP) Considerando que X representa uma variável aleatória contínua cuja função de densidade de probabilidade é $f(x) = exp(-\pi x^2)$, na qual $x \in \mathbf{R}$ e π é constante matemática, julgue o seguinte item.

Se $Y = \pi X^2$, então Y segue distribuição exponencial.

Comentários:

Essa questão trabalha com a transformação de variáveis, porque apresenta a função densidade de probabilidade da variável X, bem como a relação entre Y e X, e pede a função densidade da variável Y.

Para transformar uma função densidade em outra, vamos utilizar o método Jacobiano. O primeiro passo é encontrar a **relação inversa** $h^{-1}(y)$, isolando X na relação $Y = \pi X^2$:

$$Y = \pi X^2 \to X = \pm\sqrt{\frac{Y}{\pi}}$$

Observe que a relação inversa $X = \pm\sqrt{\frac{Y}{\pi}}$ **não** atende à definição de função, porque para um mesmo valor de X existe mais de um valor de Y. Isso ocorre porque a relação original $Y = \pi X^2$ **não** é **biunívoca** para $x \in \mathrm{R}$.

Por isso, vamos precisar separar essa relação inversa em duas funções:

$$h_1^{-1}(y) = \sqrt{\frac{y}{\pi}}$$

$$h_2^{-1}(y) = -\sqrt{\frac{y}{\pi}}$$

Agora, vamos derivar a primeira função:

$$\frac{dh_1^{-1}(y)}{dy} = \frac{d\left[\left(\frac{y}{\pi}\right)^{0,5}\right]}{dy} = 0{,}5.\left(\frac{y}{\pi}\right)^{-0,5} = 0{,}5.\left(\frac{\pi}{y}\right)^{0,5} = 0{,}5.\sqrt{\frac{\pi}{y}}$$

Como essa função é necessariamente positiva, o seu módulo é ela mesma.

Em relação à segunda função, a sua derivada é $-0{,}5.\sqrt{\frac{\pi}{y}}$ e o seu módulo é $0{,}5.\sqrt{\frac{\pi}{y}}$, igual ao da derivada da primeira função. Agora, utilizamos a função densidade de x fornecida $f_X(x) = e^{-\pi x^2}$ e substituímos x por cada uma das funções. Em relação à primeira função $h_1^{-1}(y) = \sqrt{\frac{y}{\pi}}$, temos:

$$f_X[h_1^{-1}(y)] = e^{-\pi\left(\sqrt{\frac{y}{\pi}}\right)^2} = e^{-\pi.\frac{y}{\pi}} = e^{-y}$$

Em relação à segunda função $h_2^{-1}(y) = -\sqrt{\frac{y}{\pi}}$, também temos:

$$f_X[h_2^{-1}(y)] = e^{-\pi\left(-\sqrt{\frac{y}{\pi}}\right)^2} = e^{-\pi.\frac{y}{\pi}} = e^{-y}$$

E a função densidade de Y é dada por:

$$f_Y[y] = f_X[h_1^{-1}(y)] \times \left|\frac{dh_1^{-1}(y)}{dy}\right| + f_X[h_2^{-1}(y)] \times \left|\frac{dh_2^{-1}(y)}{dy}\right|$$

$$f_Y[y] = e^{-y} \times 0{,}5.\sqrt{\frac{\pi}{y}} + e^{-y} \times 0{,}5.\sqrt{\frac{\pi}{y}} = \sqrt{\frac{\pi}{y}}e^{-y}$$

Que **não** corresponde à distribuição exponencial.

Gabarito: Errado

9. (CEBRASPE/2020 – TJ/PA) A função de densidade de probabilidade de uma variável aleatória contínua X é expressa por:

$$f_X(x) = \begin{Bmatrix} \frac{e^{-2\sqrt{x}}}{\sqrt{x}}, & se\ x \geq 0 \\ 0, & se\ x < 0 \end{Bmatrix}$$

Se $Y = \sqrt{X}$, então a função de densidade da variável Y para $y \geq 0$ é expressa por:

a) $f_Y(y) = y^{-1}e^{-2y}$.

b) $f_Y(y) = 2e^{-2y}$.

c) $f_Y(y) = y^{-2}e^{-2y^2}$.

d) $f_Y(y) = 2ye^{-2y}$.

e) $f_Y(y) = e^{-2y}$.

Comentários:

A questão informa a função **densidade** da variável X, bem como a **relação** entre as variáveis Y e X, e pede a função **densidade** da variável Y.

Para responder essa questão, vamos utilizar o **método Jacobiano** de transformação de variáveis, para transformar a função densidade de X na função densidade de Y:

$$f_Y[y] = f_X[h^{-1}(y)].\left|\frac{dh^{-1}(y)}{dy}\right|$$

O primeiro passo é encontrar a **relação inversa** $h^{-1}(y)$, isolando X na relação $Y = \sqrt{X}$:

$$Y = \sqrt{X} \rightarrow X = Y^2$$

Agora, derivamos $h^{-1}(y) = y^2$:

$$\frac{dh^{-1}(y)}{dy} = \frac{dy^2}{dy} = 2y$$

Como $y \geq 0$, então $2y \geq 0$, ou seja, o módulo dessa função é ela própria.

Agora, utilizamos a função densidade de x fornecida e substituímos x por $h^{-1}(y) = y^2$:

$$f_X[h^{-1}(y)] = \frac{e^{-2\sqrt{y^2}}}{\sqrt{y^2}} = \frac{e^{-2y}}{y}$$

Por fim, multiplicamos ambos os resultados:

$$f_Y[y] = \frac{e^{-2y}}{y} \times 2y = 2e^{-2y}$$

Gabarito: B

10. (CEBRASPE/2016 – TCE/PA) Considerando que, em um circuito elétrico, a corrente I siga uma distribuição uniforme no intervalo (0,1) e que a potência W desse circuito seja expressa por W = I^2, julgue o item a seguir relativo às transformações de variáveis.

A função de distribuição acumulada de W é F(w) = P(W ≤ w) = w^2 , em que 0 ≤ w ≤ 1.

Comentários:

Essa questão trabalha com a transformação de variáveis aleatórias, pois define a variável W em relação à variável I:

$$W = I^2$$

Como o item pede a função de distribuição acumulada, vamos utilizar o método da função de distribuição. Para isso, devemos primeiro inverter a função $W = I^2$, para encontrar a relação da variável I em função da variável W:

$$W = I^2 \rightarrow I = \sqrt{W}$$

Em seguida, precisamos da função de distribuição acumulada da variável I. Por se tratar de uma função uniforme no intervalo (0, 1), a sua função **densidade** é:

$$f(i) = \frac{1}{b-a} = \frac{1}{1-0} = 1$$

A sua função de distribuição acumulada é a **integral** desse resultado:

$$F(i) = \int_0^i 1\, di = i$$

Por fim, substituímos a relação $i = \sqrt{w}$ nessa função de distribuição acumulada:

$$F(w) = \sqrt{w}$$

Essa é a função de distribuição acumulada de w (e não $F(W) = w^2$).

Gabarito: Errado

11. (CEBRASPE/2016 – TCE/PA) Considerando que, em um circuito elétrico, a corrente I siga uma distribuição uniforme no intervalo (0,1) e que a potência W desse circuito seja expressa por W = I^2, julgue o item a seguir relativo às transformações de variáveis.

Nesse circuito, a potência esperada é igual a 1/3.

Comentários:

Para calcular a potência esperada, isto é, o valor esperado da variável W, precisamos da sua função **densidade** de probabilidade. Para isso, vamos derivar a função de distribuição acumulada $F(w) = \sqrt{w}$, que calculamos na questão anterior:

$$f(w) = \frac{dF(w)}{dw} = \frac{d(\sqrt{w})}{dw} = \frac{d(w^{0,5})}{dw} = 0{,}5.w^{-0,5}$$

Agora, multiplicamos essa função por w e integramos no intervalo da variável w.

$$E(W) = \int_{w_{min}}^{w_{max}} w.0{,}5.w^{-0{,}5}.dw = 0{,}5\int_{w_{min}}^{w_{max}} w^{0{,}5}.dw = 0{,}5.\left[\frac{w^{1{,}5}}{1{,}5}\right]_{w_{min}}^{w_{max}}$$

Agora, precisamos do intervalo de w. Sabendo que $w = i^2$ e que o intervalo de i é (0, 1), então:

$$w_{min} = 0^2 = 0$$

$$w_{max} = 1^2 = 1$$

Agora, podemos calcular a esperança de W:

$$E(W) = \frac{0{,}5}{1{,}5}(1^{1{,}5} - 0^{1{,}5}) = \frac{1}{3} \times 1 = \frac{1}{3}$$

Gabarito: Certo.

12. (CEBRASPE/2016 – TCE/PA) Considerando que, em um circuito elétrico, a corrente I siga uma distribuição uniforme no intervalo (0,1) e que a potência W desse circuito seja expressa por W = I², julgue o item a seguir relativo às transformações de variáveis.

A variância de W é maior que 0,10.

Comentários:

Para calcular a variância de W, precisamos de $E(W^2)$. Para isso, multiplicamos a função densidade de W, que calculamos na questão anterior $f(w) = 0{,}5.w^{-0{,}5}$, por w^2, no intervalo (0, 1):

$$E(W^2) = \int_0^1 w^2.0{,}5.w^{-0{,}5}.dw = 0{,}5.\int_0^1 w^{1{,}5}.dw$$

$$E(W^2) = 0{,}5.\left[\frac{w^{2{,}5}}{2{,}5}\right]_0^1 = \frac{0{,}5}{2{,}5}(1^{2{,}5} - 0^{2{,}5}) = \frac{1}{5} \times 1 = \frac{1}{5}$$

Para calcular a variância, subtraímos desse resultado o quadrado da média, que calculamos na questão anterior $E(W) = \frac{1}{3}$:

$$V(W) = E(W^2) - [E(W)]^2$$

$$V(W) = \frac{1}{5} - \left(\frac{1}{3}\right)^2 = \frac{1}{5} - \frac{1}{9} = \frac{9-5}{45} = \frac{4}{45} \cong 0{,}089$$

Logo, a variância é **menor** que 0,1.

Gabarito: Errado.

QUESTÕES COMENTADAS – CEBRASPE

Multinomial

1. (Cebraspe/2013 – TRT 17ª Região) Considerando o conceito de distribuição de probabilidade, julgue o item a seguir.

Considere que, em um tribunal, os processos sejam classificados como urgentes (T1) e não urgentes (T2) e que os não urgentes sejam reclassificados como importantes (T2.1) ou não importantes (T2.2). Considere-se, ainda, que a proporção de processos do tipo T1 seja 0,5 e que, entre os processos do tipo T2, 0,2 sejam do tipo T2.1 e 0,8 do tipo T2.2. Se X, Y e Z forem, respectivamente, as contagens de processos de tipos T1, T2.1 e T2.2 em determinado momento, então a distribuição conjunta de (X, Y, Z) é uma multinomial com parâmetros 0,5, 0,1 e 0,4.

Comentários:

Considerando que há mais do que dois resultados possíveis, não temos uma distribuição binomial, mas sim uma distribuição multinomial. Sabendo que X está associado aos processos urgentes (T1), então $p_x = 0,5$ e a probabilidade dos processos não urgentes (T2) é igual a $1 - p_x = 0,5$.

Considerando que Y está associada aos processos não urgentes e importantes e que corresponde 0,2 dos processos não urgentes, então $p_y = 0,2 \times 0,5 = 0,1$. Considerando que Z está associada aos processos não urgentes e não importantes e que corresponde a 0,8 dos processos não urgentes, então $p_z = 0,8 \times 0,5 = 0,4$.

Gabarito: Certo

2. (Cebraspe/2021 – ALE/CE) Um semáforo é composto por três círculos de luzes coloridas: vermelho, amarelo e verde. Do funcionamento desse semáforo, o sinal vermelho representa 40% do tempo, o sinal verde representa 50% do tempo e o sinal amarelo corresponde a 10% do tempo. O semáforo funciona de modo ininterrupto e não é possível que ele indique mais de um sinal ao mesmo tempo.

Caso um veículo passe 10 vezes por esse semáforo, de forma aleatória, a probabilidade de o veículo se deparar, exatamente, com 4 sinais vermelhos, 5 sinais verdes e 1 sinal amarelo é igual a

a) 0,02

b) 0,42

c) 0,1008

d) 0,00008

e) 0,15685

Comentários:

Considerando que existem **3 possibilidades** de cores de luz no semáforo e que o veículo vai passar **10 vezes** pelo semáforo, de modo que os resultados (cor da luz observada) são **independentes** entre si, então temos uma distribuição multinomial com 3 categorias. Nessa situação, a probabilidade é dada por:

$$p(n_{Vm}, n_{Vd}, n_{Am}) = \frac{N!}{n_{Vm}!\, n_{Vd}!\, n_{Am}!} \times (p_{Vm})^{n_{Vm}} \times (p_{Vd})^{n_{Vd}} \times (p_{Am})^{n_{Am}}$$

O enunciado informa que:

- O sinal vermelho representa 40% do tempo: $p_{Vm} = 0{,}4$;
- O sinal verde representa 50% do tempo: $p_{Vd} = 0{,}5$;
- O sinal amarelo representa 10% do tempo: $p_{Am} = 0{,}1$;
- O veículo vai passar 10 vezes pelo semáforo: $N = 10$

E pede a probabilidade de obter:

- 4 sinais vermelhos: $n_{Vm} = 4$
- 5 sinais verdes: $n_{Vd} = 5$
- 1 sinal amarelo: $n_{Am} = 1$

Substituindo esses dados na fórmula, temos:

$$p(4{,}5{,}1) = \frac{10!}{4!\,5!\,1!} \times (0{,}4)^4 \times (0{,}5)^5 \times (0{,}1)^1$$

Calculando essa fórmula em separado, temos:

$$\frac{10!}{4!\,5!\,1!} = \frac{10 \times 9 \times 8 \times 7 \times 6 \times 5!}{4!\,5!\,1!} = \frac{10 \times 9 \times 8 \times 7 \times 6}{4 \times 3 \times 2 \times 1} = 10 \times 9 \times 2 \times 7 = 1260$$

$$(0{,}4)^4 \times (0{,}5)^5 \times (0{,}1)^1 = 0{,}0256 \times 0{,}03125 \times 0{,}1 = 0{,}00008$$

E a probabilidade é o produto:

$$p(4{,}5{,}1) = 1260 \times 0{,}00008 = 0{,}1008$$

Observação: O produto $(0{,}4)^4 \times (0{,}5)^5 \times (0{,}1)^1 = 0{,}00008$ corresponde à probabilidade de obter os 4 sinais vermelhos, 5 verdes e 1 amarelo em **determinada ordem**. Para calcular a probabilidade de obter esse resultado em **qualquer ordem**, precisamos multiplicar pelo número de maneiras de **reordenar** esses resultados. Por isso, multiplicamos esse produto por $\frac{10!}{4!5!1!} = \frac{10!}{4!5!}$, que pode ser compreendido como a permutação de 10 sinais com repetição de 4 verdes e de 5 vermelhos.

Gabarito: C

Lista de Questões – Cebraspe

Momentos

1. (Cebraspe/2013 – Agente Nacional de Transportes Terrestres) Considere que a função geradora de momentos de uma variável aleatória discreta X seja dada pela relação $M_X(q) = (0,8 + 0,2e^q)^2$, em que $q \in R$. Com base nessas informações, julgue o item a seguir.

A probabilidade de ocorrência do evento [X = 0] é igual a 0,64.

2. (CESPE/2012 – TJ-RO) Na sequência $\{X_1, X_2, ..., X_n\}$, de variáveis aleatórias independentes e identicamente distribuídas, cada variável possui a função geradora de momentos na forma $\psi(t) = pe^t + 1 – p$, para $0 < p <1$. Com base nessas informações, assinale a opção correta.

a) Independentemente do valor de p, a distribuição da soma de $\{X_1, X_2, ..., X_n\}$ será sempre simétrica.

b) A variância dessa variável aleatória é igual à sua média.

c) A variável aleatória é contínua.

d) A variância dessa variável aleatória é maior que a sua média.

e) A função geradora de momentos da soma dessas variáveis é dada por $\psi_S(t) = (pe^t + 1 – p)^n$, $0 < p < 1$.

3. (CESPE/2009 – TRT-ES) Considere que Y seja uma variável aleatória de Bernoulli com parâmetro p, em que p é a probabilidade de uma ação judicial trabalhista ser julgada improcedente. De uma amostra aleatória simples de 1.600 ações judiciais trabalhistas, uma seguradora observou que, em média, 20% dessas ações foram julgadas improcedentes.

Com base nessa situação hipotética, julgue os próximos itens.

A estimativa de máxima verossimilhança para a função geratriz de momentos de Y é igual a 0,2 + 0,8exp(t), em que t é um número real e exp(.) denota a função exponencial.

4. (Cebraspe/2013 – Estatístico FUB) Considerando o coeficiente de assimetria que se define pelo terceiro momento central, é correto afirmar que a distribuição de Poisson exibe assimetria positiva.

5. (CESPE/2011 – Estatístico FUB) Considerando-se duas variáveis aleatórias contínuas X e Y, em que X tem função de densidade arbitrária f com função geradora de momentos M(t) e Y = exp(X), julgue o próximo item.

E(Y) = M(1)

GABARITO

1. CERTO
2. LETRA E
3. ERRADO
4. CERTO
5. CERTO

LISTA DE QUESTÕES - CEBRASPE

Distribuições Conjuntas

1. (Cebraspe/2024 – ANTT)

	$Y = 2$	$Y = 3$	$Y = 4$
$X = 0$	$\frac{1}{16}$	$\frac{1}{6}$	$\frac{1}{8}$
$X = 1$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$
$X = 2$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{16}$

Considerando a tabela precedente, que apresenta a função massa de probabilidade conjunta de duas variáveis aleatórias discretas, X e Y, julgue o item que se segue.

As distribuições X e Y são independentes.

2. (Cebraspe/2024 – ANTT)

	$Y = 2$	$Y = 3$	$Y = 4$
$X = 0$	$\frac{1}{16}$	$\frac{1}{6}$	$\frac{1}{8}$
$X = 1$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$
$X = 2$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{16}$

Considerando a tabela precedente, que apresenta a função massa de probabilidade conjunta de duas variáveis aleatórias discretas, X e Y, julgue o item que se segue.

$P(X = 2) = \frac{19}{48}$.

3. (CESPE/2022 – PC-RO) As pessoas que cumprem penas judiciais em determinado município são classificadas como reincidentes (R) ou não reincidentes (R^c).

Além disso, essas mesmas pessoas são classificadas de acordo com o tipo do regime no cumprimento da pena: regime fechado (F) ou regime não fechado (F^c).

	R	R^c
F	0,5	0,1
F^c	0,2	0,2

Considerando-se que a tabela precedente mostra a distribuição de probabilidade conjunta dos eventos apresentados na situação hipotética anterior, é correto afirmar que o valor da probabilidade condicional $P(R|F)$ é igual a

a) 5/10.

b) 7/10.

c) 5/7.

d) 5/6.

e) 9/10.

4. (CESPE/2022 – FUNPRESP-EXE)

Cliente	produto			Total
	X	Y	Z	
Tipo A	0,10	0,08	0,02	0,20
Tipo B	0,15	0,45	0,20	0,80
Total	0,25	0,53	0,22	1,00

Considerando a tabela precedente, que mostra as preferências de compras dos produtos X, Y e Z dos clientes de uma loja, classificados como tipos A e B, julgue o próximo item.

A probabilidade de um cliente do tipo A entrar ao acaso na loja é de 20%

5. (CESPE/2022 – FUNPRESP-EXE)

Cliente	produto			Total
	X	Y	Z	
Tipo A	0,10	0,08	0,02	0,20
Tipo B	0,15	0,45	0,20	0,80
Total	0,25	0,53	0,22	1,00

Considerando a tabela precedente, que mostra as preferências de compras dos produtos X, Y e Z dos clientes de uma loja, classificados como tipos A e B, julgue o próximo item.

A probabilidade de o produto X ser comprado por um cliente do tipo A é de 10%.

6. (CESPE/2022 – TELEBRÁS)

Supondo que $P(Y = y|M = m) = \frac{e^{-m}m^y}{y!}$

para $y \in \{0, 1, 2, 3 \dots\}$, em que $m > 0$, e M é uma variável aleatória contínua cuja função densidade é dada por $f_m(m) = e^{-m}$, julgue o item a seguir.

$Var(Y = y|M = m) = m.$

7. (CESPE/2022 – TELEBRÁS)

Supondo que $P(Y = y|M = m) = \frac{e^{-m}m^y}{y!}$

para $y \in \{0, 1, 2, 3 \dots\}$, em que $m > 0$, e M é uma variável aleatória contínua cuja função densidade é dada por $f_m(m) = e^{-m}$, julgue o item a seguir.

Y e M são variáveis aleatórias independentes.

8. (CESPE/2022 – TELEBRÁS - *Adaptada*) Supondo que $P(Y = y|M = m) = \frac{e^{-m}m^y}{y!}$ para $y \in \{0, 1, 2, 3 \dots\}$, em que $m > 0$, e M é uma variável aleatória contínua cuja função densidade é dada por $f_m(m) = m.e^{-m}$, julgue o item a seguir.

$P(Y > 0|M = m) = P(M \leq m).$

9. (CESPE/2022 – ANP) Uma curva de regressão da variável aleatória Y sobre $X = x$ é dada por $E[Y|X = x] = 1 - x$, em que o par de variáveis aleatórias (X, Y) segue uma distribuição normal bivariada, a média de X é igual a zero, $Var[Y] = 4$ e $Var[X] = 1$. Considerando essa situação hipotética, julgue o item a seguir.

A média de Y é igual a 1.

10. (CESPE/2022 – ANP) Uma curva de regressão da variável aleatória Y sobre $X = x$ é dada por $E[Y|X = x] = 1 - x$, em que o par de variáveis aleatórias (X, Y) segue uma distribuição normal bivariada, a média de X é igual a zero, $Var[Y] = 4$ e $Var[X] = 1$. Considerando essa situação hipotética, julgue o item a seguir.

A correlação linear entre X e Y é igual a -1.

11. (CESPE/2022 – ANP) Uma curva de regressão da variável aleatória Y sobre $X = x$ é dada por $E[Y|X = x] = 1 - x$, em que o par de variáveis aleatórias (X, Y) segue uma distribuição normal bivariada, a média de X é igual a zero, $Var[Y] = 4$ e $Var[X] = 1$.

Considerando essa situação hipotética, julgue o item a seguir.

$Var[X + Y] < 5.$

12. (CESPE/2022 – TRT/8ª Região) O número N de pessoas que chegam diariamente em uma fila de atendimento segue uma distribuição de Poisson com média igual a 500. Se Y representa o número diário dessas pessoas que recebem atendimento preferencial, e se

$$P(Y = y|N = n) = \binom{n}{y} 0,1^y 0,9^{n-y}$$

em que $0 \leq y \leq n$, então Var(Y) é igual a

a) 0,1

b) 9

c) 45

d) 50

e) 500

13. (CESPE/2022 – TRT/8ª Região) Considere duas variáveis X_1 e X_2 tais que:

$$E(X_1) = E(X_2) = 1 \quad e \quad Var(X_1) = Var(X_2) = 4$$

Se X_1 e X_2 forem variáveis aleatórias independentes, então a variância do produto X_1 e X_2 será igual a:

a) 4

b) 8

c) 16

d) 18

e) 24

14. (CESPE/2021 – TCE-RJ)

		B	
		Sim	Não
A	Sim	22	8
	Não	8	12

Considerando que o cruzamento de duas variáveis categorizadas A e B cujos níveis de resposta são "Sim" e "Não" tenha produzido a tabela de contingência precedente, julgue o próximo item.

Se a tabela de contingência em questão for representada como

		Y	
		1	0
X	1	22	8
	0	8	12

na qual X e Y são variáveis quantitativas dicotômicas, o coeficiente de correlação linear de Pearson entre X e Y será igual a 1/3.

15. (CESPE/2020 – ME)

		A			total (%)
		-1	0	1	
B	-1	10	5	5	20
	0	0	60	0	60
	1	10	5	5	20
total (%)		20	70	10	100

Considerando que a tabela precedente mostra o cruzamento de duas variáveis categorizadas A e B, que foram codificadas em três níveis numéricos de resposta: −1, 0 e 1, julgue o item que se segue.

A média da variável A é negativa.

16. (CESPE/2020 – ME)

		A			total (%)
		-1	0	1	
B	-1	10	5	5	20
	0	0	60	0	60
	1	10	5	5	20
total (%)		20	70	10	100

Considerando que a tabela precedente mostra o cruzamento de duas variáveis categorizadas A e B, que foram codificadas em três níveis numéricos de resposta: −1, 0 e 1, julgue o item que se segue.

As modas e as medianas das variáveis A e B são iguais a zero.

17. (Cebraspe/2018 – EBSERH) Todo paciente que chega a determinado posto hospitalar é imediatamente avaliado no que se refere à prioridade de atendimento. Suponha que o paciente seja classificado como “emergente” (Y = 0) ou como “não emergente” (Y = 1), e que as quantidades X, diárias, de pacientes que chegam a esse posto sigam uma distribuição de Poisson com média igual a 20. Considerando que W represente o total diário de pacientes emergentes, de tal sorte que $P(W = w|X = x) = \binom{x}{w} 0,1^w 0,9^{x-w}$, em que 0 ≤ *w* ≤ *x* e *x* ≥ 0, julgue o item subsequente.

Se, em determinado dia, 10 pacientes forem atendidos nesse posto hospitalar, então a probabilidade de se registrar, entre esses pacientes, exatamente um paciente emergente será igual a 0,1.

18. (Cebraspe/2013 – TRT 17ª Região) No que se refere a distribuições discretas, julgue o seguinte item.

Para a distribuição conjunta $P(N = n, X = x) = \frac{1}{6} \times \binom{n}{x} p^x (1-p)^{n-x}$, em que n = 1, ..., 6 e x é uma contagem, as variáveis N e X são dependentes.

19. (Cebraspe/2013 – TRT 17ª Região) Considerando o conceito de distribuição de probabilidade, julgue o item a seguir.

Considere que, em um tribunal, os processos sejam classificados como urgentes (T1) e não urgentes (T2) e que os não urgentes sejam reclassificados como importantes (T2.1) ou não importantes (T2.2). Considere-se, ainda, que a proporção de processos do tipo T1 seja 0,5 e que, entre os processos do tipo T2, 0,2 sejam do tipo T2.1 e 0,8 do tipo T2.2. Se X, Y e Z forem, respectivamente, as contagens de processos de tipos T1, T2.1 e T2.2 em determinado momento, então a distribuição conjunta de (X, Y, Z) é uma multinomial com parâmetros 0,5, 0,1 e 0,4.

20. (Cebraspe/2013 – CNJ) Uma máquina de café expresso precisa ser reiniciada algumas vezes durante o dia, devido ao uso excessivo. A tabela abaixo mostra a distribuição de probabilidade conjunta do número de vezes que ela é reiniciada na parte da manhã (M) e na parte da tarde (T).

M\T	0	1	2
0	0,10	0,10	0,30
1	0,04	0,06	0,12
2	0,06	0,10	x

Considerando essa tabela, julgue o próximo item.

O número médio de vezes que a máquina é reiniciada na parte da manhã é inferior a 1.

21. (Cebraspe/2013 – CNJ) Uma máquina de café expresso precisa ser reiniciada algumas vezes durante o dia, devido ao uso excessivo. A tabela abaixo mostra a distribuição de probabilidade conjunta do número de vezes que ela é reiniciada na parte da manhã (M) e na parte da tarde (T).

M\T	0	1	2
0	0,10	0,10	0,30
1	0,04	0,06	0,12
2	0,06	0,10	x

Considerando essa tabela, julgue o próximo item.

O número de vezes que a máquina é reiniciada na parte da tarde depende do número de vezes que ela é reiniciada pela manhã.

22. (CESPE/2011 – SEDUC/AM)

X	Y			
	-1	0	1	
-1	0	1/5	0	1/5
0	1/5	1/5	1/5	3/5
1	0	1/5	0	1/5
	1/5	3/5	1/5	1

Julgue o próximo item, considerando que o vetor aleatório (X, Y) possui distribuição conjunta de probabilidade conforme o quadro acima.

As variáveis aleatórias Z = X + Y e W = X – Y são dependentes.

23. (CESPE/2011 – SEDUC/AM)

X	Y			
	-1	0	1	
-1	0	1/5	0	1/5
0	1/5	1/5	1/5	3/5
1	0	1/5	0	1/5
	1/5	3/5	1/5	1

Julgue o próximo item, considerando que o vetor aleatório (X, Y) possui distribuição conjunta de probabilidade conforme o quadro acima.

A correlação linear entre X e Y é nula; disso se conclui que ambas são independentes.

24. (CESPE/2011 – SEDUC/AM)

$$P(X = x, Y = y) = \frac{e^{-\lambda}\lambda^{y}p^{x}(1-p)^{y-x}}{(y-x)!\,x!}, y \geq 0, \quad 0 \leq x \leq y, \quad \lambda > 0, \quad 0 \leq p \leq 1$$

A respeito da distribuição conjunta (XY), de variáveis aleatórias discretas, apresentada acima, julgue os itens a seguir.

A variância de X é menor que $\lambda.p^4$

25. (CESPE/2011 – SEDUC/AM)

$$P(X = x, Y = y) = \frac{e^{-\lambda}\lambda^{y}p^{x}(1-p)^{y-x}}{(y-x)!\,x!}, y \geq 0, \quad 0 \leq x \leq y, \quad \lambda > 0, \quad 0 \leq p \leq 1$$

A respeito da distribuição conjunta (XY), de variáveis aleatórias discretas, apresentada acima, julgue o item a seguir.

O valor esperado de X é negativo.

26. (CESPE/2010 – DETRAN/ES) Com relação a distribuições conjuntas, julgue o item que se segue.

Se X seguir uma distribuição de Bernoulli com parâmetro P e P seguir uma distribuição uniforme contínua entre 0 e 1, então a esperança de X será $\frac{3}{4}$.

27. (CESPE/2010 – INMETRO) As variáveis discretas X e Y possuem a distribuição conjunta na forma:

P(X = x, Y = y) = d(x + 2y), em que x ∈ {0, 1, 2} e y ∈ {0, 1, 2, 3, 4}.

Para os casos em que x ∉ {0, 1, 2} ou y ∉ {0, 1, 2, 3, 4}, P(X = x, Y = y) = 0.

Com base nessas informações, assinale a opção correta.

a) P(*X*=2,*Y*=3)=0,10.

b) $d = 1/75$.

c) As variáveis aleatórias X e Y são independentes.

d) $P(X \geq 1, Y \geq 3) = 7/15$.

e) Com respeito à distribuição marginal, tem-se que $P(X = 0) = 3/75$.

GABARITO

1. ERRADO
2. CERTO
3. LETRA D
4. CERTO
5. ERRADO
6. CERTO
7. ERRADO
8. CERTO
9. CERTO
10. ERRADO
11. CERTO
12. LETRA D
13. LETRA E
14. CERTO
15. CERTO
16. CERTO
17. ERRADO
18. CERTO
19. CERTO
20. CERTO
21. CERTO
22. CERTO
23. ERRADO
24. ERRADO
25. ERRADO
26. ERRADO
27. LETRA B

▪

LISTA DE QUESTÕES – CEBRASPE

Densidade Conjunta

1. (CEBRASPE/2024 – CAPES) Julgue o próximo item, com base na distribuição de probabilidade condicional $P(X = x|W = w) = \frac{e^{-w}w^x}{x!}$, em que $x = 0, 1, 2, 3, \dots, w > 0$ e W segue uma distribuição exponencial com média igual a 1.

$E(X|W = w) = w$.

2. (CEBRASPE/2024 – CAPES) Julgue o próximo item, com base na distribuição de probabilidade condicional $P(X = x|W = w) = \frac{e^{-w}w^x}{x!}$, em que $x = 0, 1, 2, 3, \dots, w > 0$ e W segue uma distribuição exponencial com média igual a 1.

O valor esperado da variável aleatória X é igual a w.

3. (CEBRASPE/2024 – CAPES) Julgue o próximo item, com base na distribuição de probabilidade condicional $P(X = x|W = w) = \frac{e^{-w}w^x}{x!}$, em que $x = 0, 1, 2, 3, \dots, w > 0$ e W segue uma distribuição exponencial com média igual a 1.

$E[Var(X|W)] = W$.

4. (CEBRASPE/2024 – CAPES) Julgue o próximo item, com base na distribuição de probabilidade condicional $P(X = x|W = w) = \frac{e^{-w}w^x}{x!}$, em que $x = 0, 1, 2, 3, \dots, w > 0$ e W segue uma distribuição exponencial com média igual a 1.

$Var(X) = 1$.

5. (CEBRASPE/2024 – CAPES) Julgue o próximo item, com base na distribuição de probabilidade condicional $P(X = x|W = w) = \frac{e^{-w}w^x}{x!}$, em que $x = 0, 1, 2, 3, \dots, w > 0$ e W segue uma distribuição exponencial com média igual a 1.

$P(X = 3) = \frac{1}{16}$.

6. (Cebraspe/2022 – TELEBRÁS) Considerando que a função de densidade conjunta do par de variáveis aleatórias (X,Y), seja dada por

$$\boldsymbol{f(x,y)=\left\{\begin{array}{lc}\frac{3(1-x^2)}{4}, & se\ |x|\leq 1\ e\ 0\leq y\leq 1;\\ 0, & caso\ contrário\end{array}\right\},}$$

julgue o próximo item.

$P(|X|\leq y|Y=y)=\frac{y(3-y^2)}{2}$, em que $0\leq y\leq 1$.

7. (Cebraspe/2022 – TELEBRÁS) Considerando que a função de densidade conjunta do par de variáveis aleatórias (X,Y), seja dada por

$$\boldsymbol{f(x,y)=\left\{\begin{array}{lc}\frac{3(1-x^2)}{4}, & se\ |x|\leq 1\ e\ 0\leq y\leq 1;\\ 0, & caso\ contrário\end{array}\right\},}$$

julgue o próximo item.

$E(X)>0$.

8. (Cebraspe/2022 – TELEBRÁS) Considerando que a função de densidade conjunta do par de variáveis aleatórias (X,Y), seja dada por

$$\boldsymbol{f(x,y)=\left\{\begin{array}{lc}\frac{3(1-x^2)}{4}, & se\ |x|\leq 1\ e\ 0\leq y\leq 1;\\ 0, & caso\ contrário\end{array}\right\},}$$

julgue o próximo item.

$Var(Y)=\frac{1}{12}$.

9. (Cebraspe/2022 – TELEBRÁS) Considerando que a função de densidade conjunta do par de variáveis aleatórias (X,Y), seja dada por

$$\boldsymbol{f(x,y)=\left\{\begin{array}{lc}\frac{3(1-x^2)}{4}, & se\ |x|\leq 1\ e\ 0\leq y\leq 1;\\ 0, & caso\ contrário\end{array}\right\},}$$

julgue o próximo item.

$P(Y=y||X|\leq y)=y$, em que $0\leq y\leq 1$.

10. (Cebraspe/2022 – TELEBRÁS) Considerando que a função de densidade conjunta do par de variáveis aleatórias (X,Y), seja dada por

$$f(x,y) = \begin{Bmatrix} \frac{3(1-x^2)}{4}, & se\ |x| \leq 1\ e\ 0 \leq y \leq 1; \\ 0, & caso\ contrário \end{Bmatrix},$$

julgue o próximo item.

A correlação linear entre as variáveis é positiva.

11. (CESPE/2022 – TRT/8ª Região) Considerando que a distribuição conjunta entre duas variáveis aleatórias contínuas X e Y é dada pela expressão

$$f(x,y) = \frac{8xy}{3} + x^2$$

se $(x,y) \in [0,1] \times [0,1]$ e $f(x,y) = 0$, se $(x,y) \notin [0,1] \times [0,1]$, a probabilidade $P(X+Y>1)$ é igual a

a) 2/9

b) 7/36

c) 29/36

d) 1/2

e) 7/9

12. (Cebraspe/2021 – PF) Considere que X e Y sejam variáveis aleatórias contínuas que se distribuem conjuntamente conforme a função de densidade.

$f(x,y) = x + y$, na qual $0 < x < 1$ e $0 < y < 1$.

Com base nessas informações, julgue o item seguinte.

X e Y são variáveis aleatórias independentes.

13. (Cebraspe/2021 – PF) Considere que X e Y sejam variáveis aleatórias contínuas que se distribuem conjuntamente conforme a função de densidade.

$f(x,y) = x + y$, na qual $0 < x < 1$ e $0 < y < 1$. Com base nessas informações, julgue o item seguinte.

Y é uma variável uniforme no intervalo (0,1)

14. (Cebraspe/2021 – PF) Considere que X e Y sejam variáveis aleatórias contínuas que se distribuem conjuntamente conforme a função de densidade.

$f(x,y) = x + y$, na qual $0 < x < 1$ e $0 < y < 1$. Com base nessas informações, julgue o item seguinte.

$E\left(X + Y|X = \frac{1}{2}\right) = \frac{14}{12}$.

15. (CESPE/2020 – TJ/PA) Uma distribuição condicional é dada por

$P(X = x|Y = y) = y^x(1-y)^{1-x}$, em que $x = 0$ ou $x = 1$ e $0 \leq y \leq 1$.

Considerando-se que Y segue uma distribuição contínua no intervalo [0, 1], é correto afirmar, a respeito da distribuição condicional $Y|X = x$, que $E(Y|X = x)$ é igual a:

a) $\frac{1}{3} + \frac{x}{3}$

b) $\frac{1}{2} + \frac{x}{2}$

c) $\frac{1}{4} + \frac{x}{4}$.

d) $\frac{x}{4}$.

e) $\frac{1}{2}$

16. (CESPE/2010 – MPU) Uma corporação é formada por uma empresa de mineração e uma empresa petrolífera. Considere que X represente as despesas mensais, em milhões de reais, com a recuperação de danos ao meio ambiente causados pela empresa de mineração e que Y represente as despesas mensais, também em milhões de reais, com a recuperação de danos ambientais causados pela empresa petrolífera. Considere, ainda, que X e Y sejam variáveis aleatórias cuja distribuição conjunta é dada por

$f(x,y) = \frac{xy}{96}$, em que 0 < x < 4 e 0 < y < 5.

Com base nessas informações, julgue o próximo item.

A equação matemática que descreve o valor esperado de X em função de Y depende unicamente de y.

17. (CESPE/2010 – MPU) Uma corporação é formada por uma empresa de mineração e uma empresa petrolífera. Considere que X represente as despesas mensais, em milhões de reais, com a recuperação de danos ao meio ambiente causados pela empresa de mineração e que Y represente as despesas mensais, também em milhões de reais, com a recuperação de danos ambientais causados pela empresa petrolífera. Considere, ainda, que X e Y sejam variáveis aleatórias cuja distribuição conjunta é dada por

$f(x,y) = \frac{xy}{96}$, **em que 0 < x < 4 e 0 < y < 5.**

Com base nessas informações, julgue o próximo item.

O valor esperado de Y é superior a 2,5 milhões de reais.

18. (CESPE/2010 – DETRAN/ES) Com relação a distribuições conjuntas, julgue o item que se segue.

Caso a função de densidade conjunta de X e Y seja dada por F(x,y) = λy.exp(-y(λ + x)), y ≥ 0, então X e Y serão independentes.

GABARITO

1. CERTO
2. ERRADO
3. ERRADO
4. ERRADO
5. CERTO
6. CERTO
7. ERRADO
8. CERTO
9. ERRADO
10. ERRADO
11. LETRA C
12. ERRADO
13. ERRADO
14. ERRADO
15. LETRA A
16. CERTO
17. CERTO
18. ERRADO

LISTA DE QUESTÕES – CEBRASPE

Transformação de Variáveis

1. (Cebraspe/2024 – ANTT) Com relação a probabilidade e variáveis aleatórias, julgue o item a seguir.

Se X é uma variável aleatória exponencial de parâmetro unitário e $U = 1 - e^{-X}$, então U é uma variável aleatória uniforme no intervalo [0, 1].

2. (Cebraspe/2024 – ANTT - *Adaptada*) Com relação a probabilidade e variáveis aleatórias, julgue o item a seguir.

Suponha que X seja uma variável aleatória contínua com função densidade de probabilidade $f(x) = \frac{3}{2}\exp(-3|x|)$, para todo $x \in \mathbb{R}$. Nesse caso, se $Y = X^4$, então a função densidade de probabilidade para Y é $g(y) = \frac{3}{4}.\frac{\exp(-3\sqrt[4]{y})}{\sqrt[4]{y^3}}$ para todo $y > 0$.

3. (CEBRASPE/2024 – CAPES) Considerando que as variáveis aleatórias X e Y sejam normais, mutuamente independentes e identicamente distribuídas, e supondo que μ e σ representem, respectivamente, a média e o desvio padrão dessas distribuições, julgue o item subsequente.

Se $S = X + Y$ e $D = X - Y$, então S e D são variáveis aleatórias independentes.

4. (CEBRASPE/2024 – CAPES) Uma variável aleatória contínua possui função de distribuição acumulada dada pela expressão a seguir, na qual a é um parâmetro tal que $a \in (0, 1)$.

$$F(x) = \begin{Bmatrix} 1 - a^x, & se\ x \geq 0 \\ 0, & se\ x < 0 \end{Bmatrix}$$

A partir dessas informações, julgue o item a seguir.

Na transformação $Y = 1 - a^x$, a variável aleatória Y segue uma distribuição contínua com média 1/2 e variância 1/12.

5. (Cebraspe/2023 – TJ/ES) Uma amostra aleatória simples de tamanho n é retirada de uma distribuição exponencial X com a função de densidade de probabilidade representada a seguir.

$$f(x) = \begin{Bmatrix} 5e^{-5x}, & se\ x \geq 0 \\ 0, & se\ x < 0 \end{Bmatrix}$$

Representando essa amostra aleatória simples com $X_1, \ldots, X_n$, julgue o item subsequente.

A variável $D = e^{-5X}$ segue distribuição uniforme no intervalo [0, 1].

6. (Cebraspe/2023 – TJ/ES) Uma amostra aleatória simples de tamanho n é retirada de uma distribuição exponencial X com a função de densidade de probabilidade representada a seguir.

$$\boldsymbol{f(x) = \begin{Bmatrix} 5e^{-5x}, & se\ x \geq 0 \\ 0, & se\ x < 0 \end{Bmatrix}}$$

Representando essa amostra aleatória simples com $X_1, \ldots, X_n$, julgue o item subsequente.

A seguir, é apresentada a função densidade da variável $W = 5X$.

$$f(w) = \begin{Bmatrix} 5e^{-w}, & se\ w \geq 0 \\ 0, & se\ w < 0 \end{Bmatrix}$$

7. (Cebraspe/2022 – TRT/8ª Região) Se X e Y forem duas variáveis aleatórias independentes tais que $X \sim Bernoulli(p = 0,8)$ e $Y \sim Binomial(n = 3, p = 0,8)$, então $P(X^4 + Y = 1)$ é igual a

a) 0,0064

b) 0,0256

c) 0,0512

d) 0,1024

e) 0,4096

8. (CEBRASPE/2021 – MJSP) Considerando que X representa uma variável aleatória contínua cuja função de densidade de probabilidade é $f(x) = exp(-\pi x^2)$, na qual $x \in \mathbb{R}$ e π é constante matemática, julgue o seguinte item.

Se $Y = \pi X^2$, então Y segue distribuição exponencial.

9. (CEBRASPE/2020 – TJ/PA) A função de densidade de probabilidade de uma variável aleatória contínua X é expressa por:

$$\boldsymbol{f_X(x) = \begin{Bmatrix} \frac{e^{-2\sqrt{x}}}{\sqrt{x}}, & se\ x \geq 0 \\ 0, & se\ x < 0 \end{Bmatrix}}$$

Se $Y = \sqrt{X}$, então a função de densidade da variável Y para $y \geq 0$ é expressa por:

a) $f_Y(y) = y^{-1}e^{-2y}$.

b) $f_Y(y) = 2e^{-2y}$.

c) $f_Y(y) = y^{-2}e^{-2y^2}$.

d) $f_Y(y) = 2ye^{-2y}$.

e) $f_Y(y) = e^{-2y}$.

10. (CEBRASPE/2016 – TCE/PA) Considerando que, em um circuito elétrico, a corrente I siga uma distribuição uniforme no intervalo (0,1) e que a potência W desse circuito seja expressa por $W = I^2$, julgue o item a seguir relativo às transformações de variáveis.

A função de distribuição acumulada de W é $F(w) = P(W \leq w) = w^2$, em que $0 \leq w \leq 1$.

11. (CEBRASPE/2016 – TCE/PA) Considerando que, em um circuito elétrico, a corrente I siga uma distribuição uniforme no intervalo (0,1) e que a potência W desse circuito seja expressa por $W = I^2$, julgue o item a seguir relativo às transformações de variáveis.

Nesse circuito, a potência esperada é igual a 1/3.

12. (CEBRASPE/2016 – TCE/PA) Considerando que, em um circuito elétrico, a corrente I siga uma distribuição uniforme no intervalo (0,1) e que a potência W desse circuito seja expressa por $W = I^2$, julgue o item a seguir relativo às transformações de variáveis.

A variância de W é maior que 0,10.

GABARITO

1. CERTO
2. CERTO
3. CERTO
4. CERTO
5. CERTO
6. ERRADO
7. LETRA B
8. ERRADO
9. LETRA B
10. ERRADO
11. CERTO
12. ERRADO

LISTA DE QUESTÕES – CEBRASPE

Multinomial

1. (Cebraspe/2013 – TRT 17ª Região) Considerando o conceito de distribuição de probabilidade, julgue o item a seguir.

Considere que, em um tribunal, os processos sejam classificados como urgentes (T1) e não urgentes (T2) e que os não urgentes sejam reclassificados como importantes (T2.1) ou não importantes (T2.2). Considere-se, ainda, que a proporção de processos do tipo T1 seja 0,5 e que, entre os processos do tipo T2, 0,2 sejam do tipo T2.1 e 0,8 do tipo T2.2. Se X, Y e Z forem, respectivamente, as contagens de processos de tipos T1, T2.1 e T2.2 em determinado momento, então a distribuição conjunta de (X, Y, Z) é uma multinomial com parâmetros 0,5, 0,1 e 0,4.

2. (Cebraspe/2021 – ALE/CE) Um semáforo é composto por três círculos de luzes coloridas: vermelho, amarelo e verde. Do funcionamento desse semáforo, o sinal vermelho representa 40% do tempo, o sinal verde representa 50% do tempo e o sinal amarelo corresponde a 10% do tempo. O semáforo funciona de modo ininterrupto e não é possível que ele indique mais de um sinal ao mesmo tempo.

Caso um veículo passe 10 vezes por esse semáforo, de forma aleatória, a probabilidade de o veículo se deparar, exatamente, com 4 sinais vermelhos, 5 sinais verdes e 1 sinal amarelo é igual a

a) 0,02

b) 0,42

c) 0,1008

d) 0,00008

e) 0,15685

GABARITO

1. CERTO

2. LETRA C

ESSA LEI TODO MUNDO CONHECE:
PIRATARIA É CRIME.

Mas é sempre bom revisar o porquê e como você pode ser prejudicado com essa prática.

1 Professor investe seu tempo para elaborar os cursos e o site os coloca à venda.

2 Pirata divulga ilicitamente (grupos de rateio), utilizando-se do anonimato, nomes falsos ou laranjas (geralmente o pirata se anuncia como formador de "grupos solidários" de rateio que não visam lucro).

3 Pirata cria alunos fake praticando falsidade ideológica, comprando cursos do site em nome de pessoas aleatórias (usando nome, CPF, endereço e telefone de terceiros sem autorização).

4 Pirata compra, muitas vezes, clonando cartões de crédito (por vezes o sistema anti-fraude não consegue identificar o golpe a tempo).

5 Pirata fere os Termos de Uso, adultera as aulas e retira a identificação dos arquivos PDF (justamente porque a atividade é ilegal e ele não quer que seus fakes sejam identificados).

6 Pirata revende as aulas protegidas por direitos autorais, praticando concorrência desleal e em flagrante desrespeito à Lei de Direitos Autorais (Lei 9.610/98).

7 Concurseiro(a) desinformado participa de rateio, achando que nada disso está acontecendo e esperando se tornar servidor público para exigir o cumprimento das leis.

8 O professor que elaborou o curso não ganha nada, o site não recebe nada, e a pessoa que praticou todos os ilícitos anteriores (pirata) fica com o lucro.

Deixando de lado esse mar de sujeira, aproveitamos para agradecer a todos que adquirem os cursos honestamente e permitem que o site continue existindo.