Aula 10

Banco do Brasil (Escriturário - Agente de Tecnologia) Probabilidade e Estatística - 2023 (Pós-Edital)

Autor:

Equipe Exatas Estratégia Concursos

03 de Janeiro de 2023

Índice

Olá, concurseiro(a)!

Agora, veremos **Distribuições Contínuas**, que são distribuições especiais, como a distribuição **normal**, que aparece muito, muito, muito nas provas. Embora esta aula esteja ligada a aulas anteriores, você consegue acompanhar esta aula, mesmo sem ter fixado muito bem as aulas anteriores.

Então respira, e vamos!

Luana Brandão

Posso te contar um pouquinho da minha trajetória? Sou Doutora em Engenharia de Produção, pela Universidade Federal Fluminense, e Auditora Fiscal da SEFAZ-RJ. Sou professora de Estatística do Estratégia, porque quero muito ajudá-lo(a) em sua trajetória rumo à aprovação!

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

professoraluanabrandao@gmail.com

@professoraluanabrandao

"Sorte é o que acontece quando a preparação encontra a oportunidade."

Sêneca

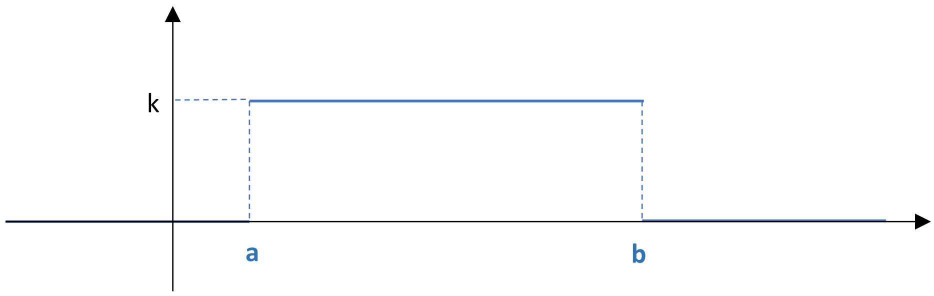
DISTRIBUIÇÕES CONTÍNUAS

Agora, veremos **distribuições teóricas** ou **especiais** de variáveis contínuas.

Distribuição Uniforme

Para variáveis **discretas**, as distribuições uniformes apresentam a **mesma probabilidade** para todos os resultados possíveis. Para variáveis uniformes **contínuas**, a situação é análoga: a função densidade de probabilidade (f.d.p.) apresenta um valor **constante** em **todo o intervalo** da variável.

A figura abaixo ilustra a f.d.p. de uma variável com distribuição uniforme, que assume um valor constante k, em um intervalo (a, b):



Ou seja, a função densidade de probabilidade para uma distribuição uniforme é da forma:

$$f(x) = \begin{Bmatrix} k, & se\ a < x < b \\ 0, & caso\ contrário \end{Bmatrix}$$

A probabilidade de todo o Espaço Amostral é de 100% = 1, ou seja, a **área** delimitada pela função é igual a 1:

$$Área = base \times altura = (\boldsymbol{b} - \boldsymbol{a}) \times k = 1$$

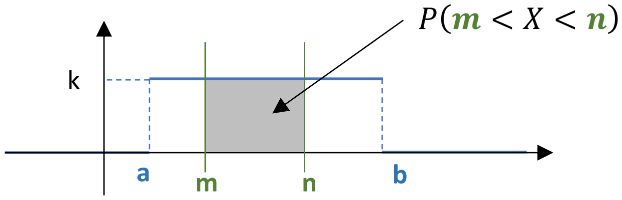
$$k = \frac{1}{(\boldsymbol{b} - \boldsymbol{a})}$$

Em outras palavras, conhecendo o intervalo $(\boldsymbol{a}, \boldsymbol{b})$, podemos calcular o valor da f.d.p. (k). Assim, os limites do intervalo $\boldsymbol{a}, \boldsymbol{b}$ são os únicos **parâmetros** da distribuição uniforme.

Por exemplo, para uma variável com distribuição contínua uniforme no intervalo $(\mathbf{1}, \mathbf{5})$, o valor da f.d.p. para $\mathbf{1} < \boldsymbol{x} < \mathbf{5}$ (ou seja, o valor de k) é:

$$k = \frac{1}{(\mathbf{5} - \mathbf{1})} = \frac{1}{4}$$

Já a probabilidade associada a um **intervalo** $(\boldsymbol{m}, \boldsymbol{n})$, com $a < m < n < b$, corresponde à **área** da região indicada abaixo:



$$Área = base \times altura = (\boldsymbol{n} - \boldsymbol{m}) \times k$$

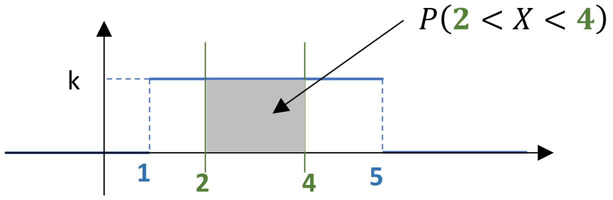
Sabendo que $k = \frac{1}{(b-a)}$, temos:

$$\boxed{P(\boldsymbol{m} < X < \boldsymbol{n}) = \frac{(\boldsymbol{n}-\boldsymbol{m})}{(\boldsymbol{b}-\boldsymbol{a})}}$$

Ou seja, a probabilidade de um intervalo em uma distribuição uniforme é a **razão** entre a amplitude desse **intervalo** e a amplitude **do intervalo total**.

Para o nosso exemplo, da variável com distribuição contínua uniforme no intervalo $(\mathbf{1}, \mathbf{5})$, a probabilidade $P(\mathbf{2} < X < \mathbf{4})$ é:

$$P(\mathbf{2} < X < \mathbf{4}) = \frac{(\mathbf{4} - \mathbf{2})}{(\mathbf{5} - \mathbf{1})} = \frac{2}{4} = 0{,}5$$



Ressalte-se que, por se tratar de uma variável contínua, temos:

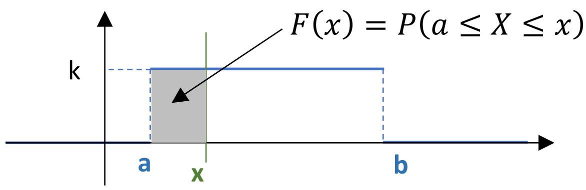
$$P(m < X < n) = P(m \leq X < n) = P(m < X \leq n) = P(m \leq X \leq n)$$

A **função de distribuição acumulada** no ponto x corresponde à probabilidade de a variável ser menor ou igual a x:

$$F(x) = P(X \leq x)$$

Para a distribuição uniforme, esse valor é igual à probabilidade de a variável estar entre $\boldsymbol{x}$ e o limite inferior do intervalo $\boldsymbol{a}$:

$$F(x) = P(X \leq x) = P(a \leq X \leq x)$$



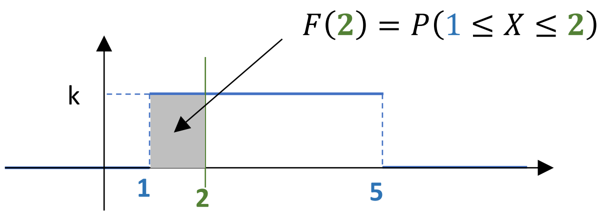
Logo, para $a < x < b$, a **função de distribuição acumulada** é dada por:

$$Área = base \times altura = (\boldsymbol{x} - \boldsymbol{a}) \times k$$

$$\boldsymbol{F}(\boldsymbol{x}) = \frac{(\boldsymbol{x}-\boldsymbol{a})}{(\boldsymbol{b}-\boldsymbol{a})}$$

Para o nosso exemplo, da variável contínua no intervalo $(\mathbf{1}, \mathbf{5})$, a f.d.a. para $\boldsymbol{x} = \mathbf{2}$, isto é, a probabilidade $P(X \leq \mathbf{2})$ é:

$$\boldsymbol{F}(\mathbf{2}) = \frac{(\mathbf{2} - \mathbf{1})}{(\mathbf{5} - \mathbf{1})} = \frac{1}{4} = 0{,}25$$



A maioria das distribuições teóricas conhecidas, tanto discretas quanto contínuas, fazem parte da chamada **família exponencial**, cuja **função densidade de probabilidade** pode ser descrita de forma similar, como as distribuições discretas: binomial, geométrica, hipergeométrica, de Poisson; e as contínuas: exponencial, Normal, Beta, Weibull.

A **distribuição uniforme contínua**, assim como a distribuição uniforme discreta, é uma distribuição importante que **não** pertence à família exponencial.

(FGV/2017 – IBGE) Uma variável aleatória contínua X é uniformemente distribuída no intervalo real [0 , 50]. A probabilidade de que X seja maior do que 20 é igual a:

a) 0,8

b) 0,6

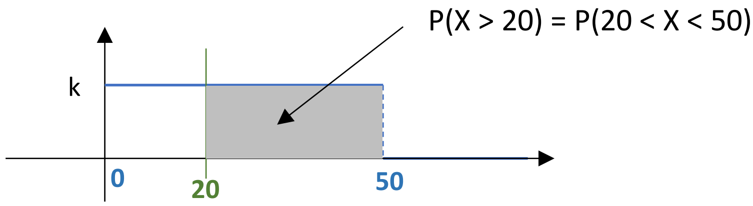
c) 0,4

d) 0,2

e) 0,1

Comentários:

O enunciado informa que X é uniformemente distribuída no intervalo de [0,50], logo a = 0 e b = 50. A probabilidade P(X > 20) é dada por:



$$P(X > 20) = base \times altura = (50 - 20) \times k = \frac{(50 - 20)}{(50 - 0)}$$

$$P(X > 20) = \frac{30}{50} = 0{,}6$$

Gabarito: B

(VUNESP/2015 – TJ-SP) Leia o texto para responder à questão.

A Cia. Alfa Auto-ônibus declara, em seus catálogos, que o tempo de viagem entre duas cidades é de 3 horas. No entanto o tempo real de viagem é uma variável aleatória x que se distribui uniformemente entre 175 e 190 minutos, ou seja,

$$f(x) = \begin{Bmatrix} \frac{1}{15} \quad para\ 175 \le x \le 190 \\ 0\ em\ qualquer\ outro\ lugar \end{Bmatrix}$$

Considere ainda que qualquer tempo x do intervalo tal que x > 180 é considerado como atraso.

A probabilidade de que a viagem não terá mais do que 5 minutos de atraso é

a) 1/2

b) 2/3

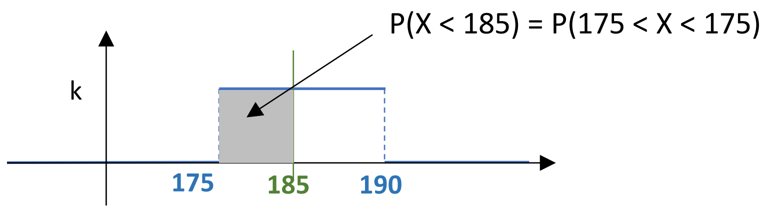
c) 1/8

d) 1/9

e) 1/15

Comentários:

O enunciado informa que o tempo de viagem é uma variável com distribuição uniforme entre 175 e 190 minutos, ou seja, a = 175 e b = 190. Considerando que a viagem deve ter 180 minutos, a probabilidade de ela não atrasar mais do que 5 minutos corresponde a P(X < 185), dada por:



$$P(X < 185) = base \times altura = (\mathbf{185} - \mathbf{175}) \times k = \frac{(\mathbf{185} - \mathbf{175})}{(\mathbf{190} - \mathbf{175})}$$

$$P(X < 185) = \frac{10}{15} = \frac{2}{3}$$

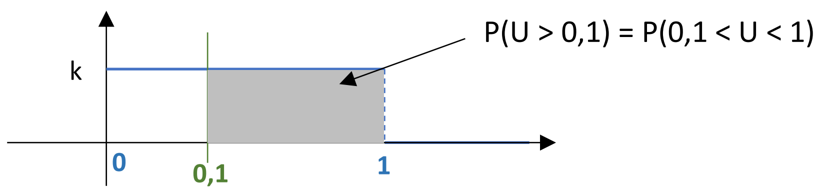
Gabarito: B

(CESPE/2016 – Auditor de Controle Externo do TCE/PA) A respeito de uma variável aleatória contínua *U*, uniformemente distribuída no intervalo [0, 1], julgue o seguinte item.

P(U > 1/10) = 0,9.

Comentários:

Para uma distribuição uniforme no intervalo a = 0 e b = 1, a probabilidade P(U > 1/10) = P(U > 0,1) é dada por:



$$P(U > 0{,}1) = base \times altura = (\mathbf{1} - \mathbf{0,1}) \times k = \frac{(\mathbf{1} - \mathbf{0,1})}{(\mathbf{1} - \mathbf{0})}$$

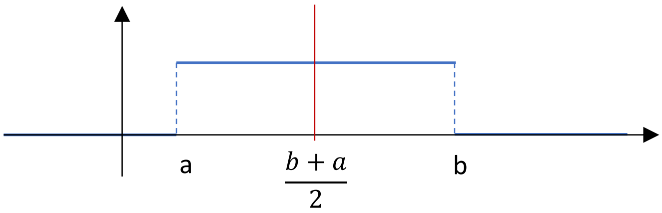
$$P(U > 0{,}1) = 0{,}9$$

Gabarito: Certo.

Esperança e Variância

A esperança de uma variável uniforme é a **média** aritmética dos extremos do seu intervalo $\boldsymbol{a}, \boldsymbol{b}$:

$$E(X) = \frac{\boldsymbol{b}+\boldsymbol{a}}{2}$$



Para o nosso exemplo da variável contínua no intervalo $(\mathbf{1}, \mathbf{5})$, a esperança é:

$$E(X) = \frac{\boldsymbol{b}+\boldsymbol{a}}{2} = \frac{\mathbf{1}+\mathbf{5}}{2} = 3$$

E a **variância** é dada por:

$$V(X) = \frac{(\boldsymbol{b}-\boldsymbol{a})^2}{12}$$

Para o nosso exemplo da variável contínua no intervalo $(\mathbf{1}, \mathbf{5})$, a variância é:

$$V(X) = \frac{(\boldsymbol{b}-\boldsymbol{a})^2}{12} = \frac{(\mathbf{5}-\mathbf{1})^2}{12} = \frac{(4)^2}{12} = \frac{16}{12} = \frac{4}{3}$$

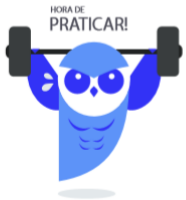
ESQUEMATIZANDO

Distribuição Uniforme

Função Densidade de Probabilidade: $f(x) = \frac{1}{(\boldsymbol{b}-\boldsymbol{a})}, \; se \; \boldsymbol{a} \le x \le \boldsymbol{b}$

Cálculo da Probabilidade: $P(\boldsymbol{m} < X < \boldsymbol{n}) = \frac{(\boldsymbol{n}-\boldsymbol{m})}{(\boldsymbol{b}-\boldsymbol{a})}$

Esperança: $E(X) = \frac{\boldsymbol{b}+\boldsymbol{a}}{2}$; **Variância:** $V(X) = \frac{(\boldsymbol{b}-\boldsymbol{a})^2}{12}$



(CESPE/2016 – Auditor de Controle Externo do TCE/PA) A respeito de uma variável aleatória contínua *U*, uniformemente distribuída no intervalo [0, 1], julgue o seguinte item.

A variância de U é inferior a 1/10.

Comentários:

Para uma distribuição uniforme no intervalo a = 0 e b = 1, a variância é dada por:

$$V(X) = \frac{(b-a)^2}{12}$$

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

Como 1/12 é inferior a 1/10, o item está correto.

Gabarito: Certo.

(FCC/2015 – Analista Judiciário do TER/RR) Uma variável aleatória X tem distribuição uniforme contínua com média igual a 4 e variância igual a 12.

Nessas condições, P(X < 7) é igual a

a) 0,45.

b) 0,75.

c) 0,25.

d) 0,60.

e) 0,67.

Comentários:

Tratando-se de uma distribuição uniforme com média igual a 4, temos:

$$\mu = \frac{b+a}{2} = 4$$

$$b + a = 8$$

Sendo a variância igual a 12, temos[1]:

$$V(X) = \frac{(b-a)^2}{12} = 12$$

$$(b-a)^2 = 12 \times 12$$

$$b - a = 12$$

Somando as duas equações, $b + a = 8$ e $b - a = 12$, temos:

$$b + a + b - a = 8 + 12$$

$$2b = 20$$

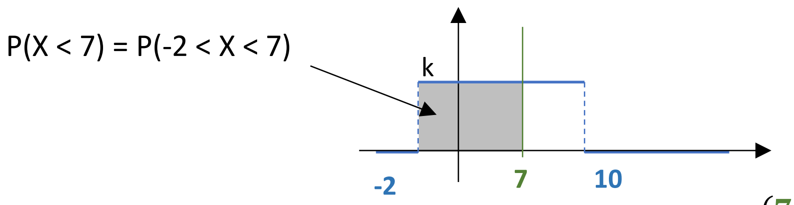
$$b = 10$$

Substituindo esse resultado em $b + a = 8$, temos:

$$10 + a = 8$$

$$a = -2$$

A probabilidade P(X < 7), sendo a = -2 e b = 10, é igual à probabilidade de X estar entre -2 e 7:



$$P(X < 7) = base \times altura = (\mathbf{7} - (\mathbf{-2})) \times k = \frac{(\mathbf{7} - (\mathbf{-2}))}{(\mathbf{10} - (\mathbf{-2}))}$$

$$P(X < 7) = \frac{9}{12} = 0{,}75$$

Gabarito: B

---

[1] A raiz de um número elevado ao quadrado é igual ao módulo do número:

$$\sqrt{x^2} = |x|$$

Isso porque x pode ser um número negativo, mas a raiz de um número é necessariamente um número positivo. Por exemplo, podemos ter $x = -3$, então:

$$\sqrt{(-3)^2} = \sqrt{9} = 3$$

Ou seja, para a nossa questão, temos:

$$\sqrt{(b-a)^2} = |b - a|$$

Como b > a, então b – a > 0 e, consequentemente:

$$|b - a| = b - a$$

(FCC/2011 – Analista de Controle Atuarial do TCE/PR) Sabe-se que a variável aleatória X tem distribuição uniforme contínua no intervalo [10, ß], ß > 10. Sabendo-se que a variância de X é igual a 3, o valor de K tal que P(X > K) = 0,3 é

a) 14,2

b) 13,8

c) 13,5

d) 13,1

e) 12,8

Comentários:

Sabendo que a variância de X, com distribuição discreta, no intervalo [10, ß], é 3, então[2]:

$$V(X) = \frac{(b-a)^2}{12} = 3$$

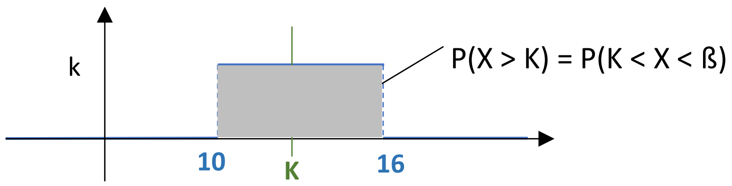
$$V(X) = \frac{(\beta - 10)^2}{12} = 3$$

$$(\beta - 10)^2 = 36$$

$$\beta - 10 = 6$$

$$\beta = 16$$

Assim, o valor de K para o qual P(X > K) = 0,3, considerando que essa probabilidade corresponde à probabilidade de X estar entre K e o limite superior do intervalo β = 16, é:



$$P(X > K) = base \times altura = (\mathbf{16} - \boldsymbol{K}) \times k = \frac{(\mathbf{16} - \boldsymbol{K})}{(\mathbf{16} - \mathbf{10})} = 0{,}3$$

$$16 - K = 6 \times 0{,}3 = 1{,}8$$

$$K = 14{,}2$$

Gabarito: A.

---

[2] Aqui temos a mesma situação, considerando que $\beta > 10$, então:

$$\sqrt{(\beta - 10)^2} = |\beta - 10| = \beta - 10$$

DISTRIBUIÇÃO EXPONENCIAL

A distribuição exponencial se caracteriza por ter uma **taxa de falha (ou de ocorrência) constante**, sendo normalmente associada a um **tempo**. Por exemplo, o **tempo de vida** de um micro-organismo ou a **vida útil** de uma lâmpada podem seguir uma distribuição exponencial.

A **função densidade de probabilidade** (f.d.p.) de uma variável X com distribuição exponencial é:

$$f(x) = \begin{cases} \boldsymbol{\lambda . e^{-\lambda x}}, & se\ \boldsymbol{x \geq 0} \\ 0, & se\ x < 0 \end{cases}$$

Em que e, número de Euler, é um número irracional, cujo valor aproximado é $e \cong 2{,}718$. É comum utilizar a seguinte notação para representar o expoente na base e:

$$e^{-\lambda x} = \exp\{-\lambda x\}$$

Observa-se que a f.d.p. apresenta algum valor $f(x) \neq 0$ apenas para valores **positivos** da variável $x \geq 0$.

Ressalte-se que $\boldsymbol{\lambda}$ é necessariamente **positivo** e representa a **taxa de falha por unidade de tempo**. Esse é o **único parâmetro** da distribuição (a f.d.p. só depende desse valor para ser caracterizada)

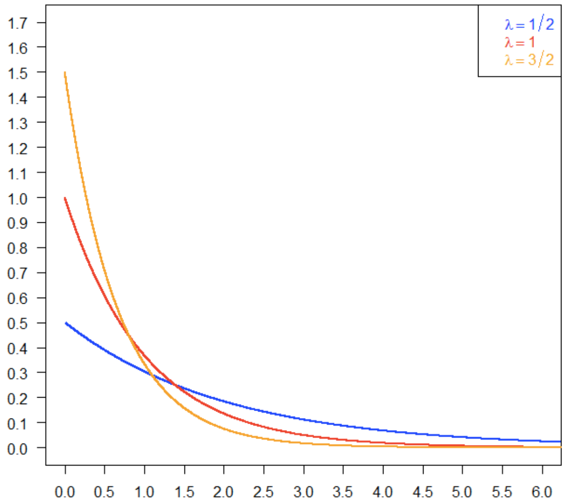
Por exemplo, o **tempo de vida** de um micro-organismo pode seguir uma distribuição exponencial com parâmetro $\boldsymbol{\lambda = \frac{1}{12}}$ **por dia**; a **vida útil** de uma lâmpada pode seguir uma distribuição exponencial com parâmetro $\boldsymbol{\lambda = \frac{1}{100}}$ **por hora**.

O valor de x corresponde ao **tempo** até a falha (ou ocorrência). Por exemplo, podemos calcular a probabilidade de o micro-organismo **viver menos de $\boldsymbol{x = 3}$ dias**; ou a probabilidade de a lâmpada **durar mais que $\boldsymbol{x = 100}$ horas e menos que $\boldsymbol{x = 200}$ horas**.

O gráfico a seguir[1] apresenta as funções densidade de probabilidade para variáveis com distribuição exponencial, para os parâmetros $\lambda = 0{,}5$, $\lambda = 1$ e $\lambda = 1{,}5$. Podemos observar que a f.d.p. é **decrescente** e que assume valores no intervalo $x \in (0, \infty)$.

---

[1] Gráfico obtido na seção de Distribuição Exponencial do Portal Action, disponível em http://www.portalaction.com.br/probabilidades/612-distribuicao-exponencial



A probabilidade X estar em um **intervalo** $P(a < X < b)$ é calculada como[2]:

$$\boxed{P(\boldsymbol{a} < X < \boldsymbol{b}) = e^{-\lambda.\boldsymbol{a}} - e^{-\lambda.\boldsymbol{b}}}$$

Ou seja, a probabilidade de a lâmpada durar entre 100 e 200 horas é dada por:

$$P(\mathbf{100} < X < \mathbf{200}) = e^{-\frac{\mathbf{1}}{\mathbf{100}}\times\mathbf{100}} - e^{-\frac{\mathbf{1}}{\mathbf{100}}\times\mathbf{200}} = e^{-\mathbf{1}} - e^{-\mathbf{2}} \cong 0{,}5$$

E a probabilidade de o micro-organismo viver **menos que $\boldsymbol{x} = \mathbf{3}$ dias**?

A f.d.p. exponencial assume algum valor somente para $\boldsymbol{x} \geq \mathbf{0}$ (realmente, não tem como o tempo ser negativo). Então, a probabilidade de ele durar menos de $x = 3$ dias equivale à probabilidade de durar mais que $x = 0$ e menos que $x = 3$ dias:

$$P(X < 3) = P(0 < X < 3)$$

Agora, aplicamos a fórmula que vimos antes. Considerando $\lambda = \frac{1}{12}$ por dia, temos:

$$P(\mathbf{0} < X < \mathbf{3}) = e^{-\frac{\mathbf{1}}{\mathbf{12}}\times\mathbf{0}} - e^{-\frac{\mathbf{1}}{\mathbf{12}}\times\mathbf{3}} = \boldsymbol{e}^{\mathbf{0}} - e^{-\frac{1}{4}} = \mathbf{1} - e^{-\frac{1}{4}}$$

---

[2] Pontue-se que o extremo inferior do intervalo, $\boldsymbol{a}$, antecede o termo superior, $\boldsymbol{b}$, no cálculo da probabilidade da função exponencial porque a função é **decrescente**, como se pode observar no gráfico acima.

Logo, o valor de $e^{-\lambda.\boldsymbol{a}}$ é maior do que o valor de $e^{-\lambda.\boldsymbol{b}}$, para $\boldsymbol{a} < \boldsymbol{b}$.

Em geral, para calcular a probabilidade $P(X < x)$ para algum valor de x em uma distribuição exponencial, fazemos:

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

A probabilidade $\boldsymbol{P(X < x)}$ é igual à **função de distribuição acumulada** no ponto $\boldsymbol{x}$. Então, podemos dizer que a f.d.a. da variável exponencial é:

$$F(x) = 1 - e^{-\lambda x}$$

Quando uma variável exponencial representa a **taxa de falha** de um sistema, a **função de confiabilidade** do sistema é **complementar** à **função de distribuição acumulada** e representa a probabilidade de o sistema **não falhar** no período x.

$$R(x) = 1 - F(x)$$

Essa expressão é bem utilizada quando um sistema é composto por mais de uma unidade.

Quando as unidades são **independentes** e **redundantes**, isto é, quando é necessário que **todas falhem** para haver falha no sistema, a função de distribuição acumulada do sistema (isto é, a probabilidade de ele falhar no tempo x) é o **produto** das funções de distribuição acumulada individuais das unidades (interseção):

$$F_S(x) = F_1(x) \times F_2(x) \ldots \times F_n(x)$$

E a probabilidade de a variável assumir um valor **maior que** $\boldsymbol{x}$, $P(X > x)$? Por exemplo, qual seria a probabilidade de o micro-organismo viver **mais que $\boldsymbol{x = 3}$ dias**?

Para isso, calculamos a probabilidade do evento **complementar**:

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

$$\boldsymbol{P(X \geq x) = e^{-\lambda.x}}$$

Para efetuar os cálculos da distribuição exponencial, é essencial que x e λ estejam na **mesma unidade de tempo**. Se a questão trabalhar com unidades de tempo diferentes, será necessário ajustar as medidas.

Por exemplo, vamos calcular a probabilidade de o micro-organismo durar menos de 6 **horas**, considerando que $\lambda = \frac{1}{12}$ **por dia**. Sabendo que há 24 horas em um dia, então 6 horas correspondem a:

$$x = \frac{6}{24} = \frac{1}{4}\, dia$$

E a probabilidade $P\left(X < \frac{1}{4}\right)$ é:

$$P\left(X < \frac{1}{4}\right) = 1 - e^{-\frac{1}{12}\times\frac{1}{4}} = 1 - e^{-\frac{1}{48}}$$

(2016 – EBSERH) Um estatístico ajustou um modelo de distribuição exponencial à variável aleatória correspondente ao tempo de falha T (tempo até falhar em anos) de um produto. O modelo tem a expressão f(t) = $0{,}2e^{-0{,}2t}$ t $\geq$ 0. Então, a probabilidade de o produto falhar dentro da garantia pretendida de 1 ano é

a) 0,818731

b) 0,821754

c) 0,803112

d) 0,181269

e) 0,196888

Comentários:

Trata-se de uma distribuição exponencial, com parâmetro $\lambda = 0{,}2$, em que se deseja calcular a probabilidade $P(X < 1)$:

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

$$P(X < 1) = 1 - e^{-0{,}2} \cong 0{,}181$$

Gabarito: D.

Esperança, Variância e Propriedade

A **esperança (ou média)** da distribuição exponencial é o inverso do parâmetro λ:

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

Por exemplo, sendo a vida útil de uma lâmpada uma variável exponencial com parâmetro $\lambda = \frac{1}{100} = 0{,}01$ por hora, então a média será de:

$$E(X) = \frac{1}{\lambda} = \frac{1}{0{,}01} = 100$$

A **variância** é dada por:

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

Nesse exemplo, a variância é de:

$$V(X) = \frac{1}{\lambda^2} = \frac{1}{0{,}01^2} = 10.000$$

E o **desvio padrão**, raiz quadrada da variância é:

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

$$\sigma = \frac{1}{\lambda}$$

Em outras palavras, o **desvio padrão** é igual à **média**!

A distribuição exponencial guarda uma relação muito especial com a **distribuição de Poisson**: a distribuição exponencial descreve o **tempo entre as ocorrências** de eventos sucessivos de uma distribuição de Poisson, com o **mesmo parâmetro**!

Por exemplo, se a chegada de pessoas em uma loja segue uma distribuição de **Poisson** com $\lambda_X = 2$ pessoas por hora, então o **tempo decorrido** entre a chegada de uma e de outra pessoa segue uma distribuição **exponencial** com parâmetro de $\lambda_Y = 2$ por hora.

Desse modo, a média do tempo decorrido entre as chegadas, medida em hora, é:

$$E(Y) = \frac{1}{\lambda} = \frac{1}{2}$$

Vale destacar também a seguinte **propriedade**: a distribuição exponencial é considerada "**sem memória**", ou seja, não importa o que já ocorreu, as probabilidades para o futuro permanecem a mesma:

$$P(X > t + s|X > s) = P(X > t)$$

Ou seja, **sabendo** que já se passou um tempo $X > s$ desde a ocorrência do último evento, a probabilidade de se passar um tempo $X > t$ **a mais**, até a ocorrência do próximo evento, ou seja, $X > t + s$ no total, é igual à probabilidade de se passar um tempo $X > t$, até a ocorrência de um evento.

Vamos considerar um exemplo numérico para facilitar o entendimento. Suponha uma variável exponencial que considere horas como unidade de tempo. De acordo com essa propriedade, sabendo que **já se passaram 3 horas** desde a ocorrência do último evento, a probabilidade de se passar **pelo menos 2 horas a mais** (totalizando 5 horas) é **igual** à probabilidade de se passar **pelo menos 2 horas** até a ocorrência do evento (não condicionada):

$$P(X > 2 + 3|X > 3) = P(X > 2)$$

INDO MAIS FUNDO!

Existe uma forma alternativa de parametrizar a f.d.p. de uma distribuição exponencial, substituindo $\boldsymbol{\lambda}$ por $\frac{\mathbf{1}}{\boldsymbol{\beta}}$:

$$f(x) = \begin{Bmatrix} \frac{1}{\beta}e^{-\frac{x}{\beta}}, & se\ x \geq 0 \\ 0, & se\ x < 0 \end{Bmatrix}$$

O parâmetro β, inverso do parâmetro taxa λ, pode ser interpretado como o parâmetro de escala da distribuição, medida em **unidade de tempo**.

No exemplo do inseto, teríamos $\beta = 12$ dias; e no da lâmpada, teríamos $\beta = 100$ horas.

Para essa parametrização, temos:

$$\boldsymbol{E(X) = \beta}, \qquad \boldsymbol{V(X) = \beta^2}$$

Ou seja, o valor do parâmetro β corresponde ao valor da **média** da distribuição.

ESQUEMATIZANDO

Distribuição Exponencial

Função Densidade de Probabilidade: $f(x) = \boldsymbol{\lambda}.e^{-\lambda x}, \;\; se\ x \geq 0$:

Cálculo da Probabilidade:

$P(\boldsymbol{a} < X < \boldsymbol{b}) = e^{-\lambda.\boldsymbol{a}} - e^{-\lambda.\boldsymbol{b}}$ $\qquad$ $P(X > \boldsymbol{a}) = e^{-\lambda.\boldsymbol{a}}$

Função de Distribuição Acumulada: $P(X < \boldsymbol{b}) = 1 - e^{-\lambda.\boldsymbol{b}}$

Esperança: $E(X) = \frac{1}{\lambda}$; $\qquad$ **Variância:** $V(X) = \frac{1}{\lambda^2}$

Propriedade: $P(X > t + s | X > s) = P(X > t)$

HORA DE PRATICAR!

(FGV/2014 – DPE-RJ) Constatou-se que o tempo de tramitação de um processo pelas instâncias do judiciário, até o arquivamento em definitivo, é uma variável aleatória contínua exponencial. Para os casos de processos em que quadros da Defensoria Pública atuam, o tempo médio de duração tem sido de 225 dias. Então a probabilidade de que um processo tenha duração inferior a um mês e meio (45 dias) é igual a:

a) $1 – e^{-5}$

b) e^{-5}

c) $1 – e^{-0,2}$

d) $e^{-0,2}$

e) $\left(C_{225}^{45}\right)^{-1}$

Comentários:

O enunciado informa que o tempo de tramitação de um processo tem distribuição exponencial com média de 225 dias, logo:

$$E(X) = \frac{1}{\lambda} = 225$$

$$\lambda = \frac{1}{225}$$

A probabilidade de um processo ter duração inferior a 45 dias é dada por:

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

$$P(X < 45) = 1 - e^{-\frac{45}{225}} = 1 - e^{-0,2}$$

Gabarito: C

(FCC/2015 – Analista Judiciário do TRE/RR) Em um determinado órgão público o tempo X, em horas, entre duas solicitações consecutivas, feitas pelo departamento de recursos humanos, pode ser considerado como tendo distribuição exponencial com média de 5 horas. Nessas condições, a probabilidade do tempo entre duas solicitações estar compreendido entre 2 horas e 6 horas é, em %, igual a

Dados: $e^{-0,2} = 0,819$; $e^{-0,4} = 0,670$; $e^{-1,2} = 0,301$.

a) 18,1

b) 63,1

c) 51,9

d) 36,9

e) 34,5

Comentários:

Sendo X uma variável com distribuição exponencial e média 5, temos:

$$E(X) = \frac{1}{\lambda} = 5$$

$$\lambda = \frac{1}{5}$$

Logo, a probabilidade de o tempo estar compreendido entre 2 horas e 6 horas é dado por:

$$P(\boldsymbol{a} < X < \boldsymbol{b}) = e^{-\lambda.\boldsymbol{a}} - e^{-\lambda.\boldsymbol{b}}$$

$$P(\mathbf{2} < X < \mathbf{6}) = e^{-\frac{2}{5}} - e^{-\frac{6}{5}} = e^{-0,4} - e^{-1,2}$$

Pelos dados fornecidos pelo enunciado, temos:

$$P(2 < X < 6) = 0,670 - 0,301 = 0,369 = 36,9\%$$

Gabarito: D

(FGV/2021 – FunSaúde/CE) Suponha que carros passem por um posto de observação em uma estrada remota de acordo com um processo Poisson, com taxa média de ocorrência igual a 2 carros por minuto.

Se um carro acaba de passar por esse posto, o tempo de espera, até que o próximo carro passe pelo posto, tem distribuição de probabilidades:

a) Cauchy $(\alpha = 1, \beta = 2)$

b) Beta $(\alpha = 1, \beta = 2)$

c) uniforme $(0, 2)$

d) exponencial $(\lambda = 2)$

e) normal $(\mu = 2, \sigma^2 = 1)$

Comentários:

O tempo decorrido entre duas ocorrências de um processo de Poisson segue distribuição exponencial com o mesmo parâmetro. Considerando que a passagem de carros segue distribuição de Poisson com parâmetro $\lambda = 2$, o tempo entre essas ocorrências segue distribuição exponencial com parâmetro $\lambda = 2$.

Gabarito: D

(CESPE/2016 – Analista de Controle Externo do TCE/PA) Se o tempo de espera por atendimento (T, em minutos) em determinada repartição pública segue uma distribuição exponencial com média igual a 30 minutos, então

P(T > 35 | T > 30) = P(T > 35)

Comentários:

Essa questão trabalha com a propriedade "sem memória" da distribuição exponencial:

$$P(X > t + s|X > s) = P(X > t)$$

Sendo X = T, s = 30 e t = 5, temos:

$$P(T > 35|T > 30) = P(T > 5)$$

Gabarito: Errado.

(FGV/2016 – IBGE) Um fabricante de equipamentos de informática, que conhece a distribuição do tempo de vida útil dos HDs externos, precisa avaliar os gastos com serviços de garantia. Essa distribuição é a exponencial com média β = 15 anos, sendo que os HDs já vendidos têm, por hipótese, 3 anos de uso, sem apresentar defeitos. Supondo que a garantia é de 12 anos, a probabilidade de que ele tenha que prestar assistência a um determinado HD entre os vendidos é

a) $1 – e^{-0,6}$;

b) $e^{-0,75}$;

c) $e^{-0,6}$;

d) $1 – e^{-0,75}$;

e) $e^{-0,25}$.

Comentários:

A probabilidade de o fabricante ter que prestar assistência corresponde à probabilidade de o HD durar **menos que 12 anos** e a probabilidade de ele **não** prestar assistência corresponde à probabilidade de o HD durar **mais que 12 anos**.

Sabendo que os HDs foram vendidos há 3 anos, a probabilidade de ele **não** prestar assistência, ou seja, de durar pelo menos 12 – 3 = 9 anos a mais, segue a propriedade sem memória da distribuição exponencial:

$$P(T > 12|T > 3) = P(T > 9)$$

E a probabilidade de o fabricante ter que **prestar assistência** pode ser calculada pela probabilidade **complementar**:

$$P(T < 12|T > 3) = 1 - P(T > 12|T > 3) = 1 - P(T > 9)$$
$$P(T < 12|T > 3) = P(T < 9)$$

Conhecendo a média da distribuição, β = 15 anos, podemos calcular o valor de λ (observe que foi considerada a parametrização alternativa da distribuição exponencial):

$$E(X) = \beta = \frac{1}{\lambda} = 15$$
$$\lambda = \frac{1}{15}$$

A probabilidade de o HD durar menos de 9 anos é, então, dada por:

$$P(T < 9) = 1 - e^{-\lambda .9} = 1 - e^{-\frac{9}{15}} = 1 - e^{-0,6}$$

Gabarito: A

Considerando duas ou mais variáveis

Agora vamos aprender um raciocínio diferente para resolver questões envolvendo a **seleção** aleatória de objetos que apresentam **distribuições exponenciais distintas**, utilizando o **Teorema da Probabilidade Total**!

Por exemplo, seja X uma variável com distribuição exponencial com parâmetro $\lambda_X = 0{,}01$ e Y uma variável com distribuição exponencial com parâmetro $\lambda_Y = 0{,}05$.

Suponha, ainda que a proporção de objetos que sigam a distribuição X seja P(X) = 60% = 0,6 e que os demais objetos sigam a distribuição Y, ou seja, P(Y) = 1 - 0,6 = 0,4.

Se selecionarmos um objeto ao acaso, a probabilidade de ele ter uma vida útil maior que z = 50, por exemplo, pode ser calculada pelo Teorema da Probabilidade Total.

Para isso, vamos chamar o **atributo desejado**, qual seja ter uma vida útil maior que z = 50, de **evento A.** Pelo **Teorema da Probabilidade Total**, temos:

$$P(A) = P(A|X) \times P(X) + P(A|Y) \times P(Y)$$

Nessa equação, $P(A|X)$ representa a probabilidade de o objeto ter uma vida útil maior que z = 50 (evento A), dado que segue a distribuição de X, enquanto $P(A|Y)$ representa a probabilidade de o objeto ter uma vida útil maior que z = 50 (evento A), dado que segue a distribuição de Y.

Essas probabilidades condicionais podem ser calculadas pelas fórmulas da **distribuição exponencial**:

$$P(A|X) = P(X > 50) = e^{-50\times 0{,}01} = e^{-0{,}5}$$

$$P(A|Y) = P(Y > 50) = e^{-50\times 0{,}05} = e^{-2{,}5}$$

Logo, conhecendo essas probabilidades condicionais e as probabilidades P(X) e P(Y), podemos calcular a probabilidade desejada:

$$P(A) = e^{-0,5} \times 0,6 + e^{-2,5} \times 0,4 \cong 0,61 \times 0,6 + 0,08 \times 0,4 \cong 0,397$$

Pontue-se que é possível aplicar esse mesmo raciocínio para **outras distribuições**.

A **soma de variáveis exponenciais independentes** segue distribuição **hipoexponencial** ou **distribuição generalizada de Erlang**. Ela é chamada de **hipo**exponencial porque o seu coeficiente de variação (CV) é sempre **menor que 1**.

Para a distribuição exponencial, o CV é **igual a 1**, pois o desvio padrão é igual à média.

Já, a distribuição **hiperexponencial** tem CV **maior que 1** e a sua função densidade de probabilidade corresponde à **soma das funções densidade** de variáveis **exponenciais** multiplicadas pelas respectivas **probabilidades**.

Considerando o exemplo da vida útil dos objetos com as mesmas probabilidades, teríamos:

$$f_{hiper}(x) = 0,01. e^{-0,01.x} \times 0,6 + 0,05. e^{-0,05.x} \times 0,4$$

(FCC/2013 – TRT/5ª Região) Uma empresa produz componentes de dois tipos: A e B. Sejam as variáveis aleatórias: X = tempo de vida do componente A, em horas e Y = tempo de vida do componente B, em horas. De um lote de 120 componentes do tipo A e 80 componentes do tipo B, retira-se ao acaso um componente. Sabendo-se que X tem distribuição exponencial com média de 1.000 horas e que Y tem distribuição exponencial com média de 700 horas, a probabilidade do componente selecionado ter duração inferior a 1.400 horas é

Dados: $e^{-1} = 0,37$; $e^{-1,4} = 0,25$; $e^{-2} = 0,14$

a) 0,569

b) 0,742

c) 0,618

d) 0,794

e) 0,634

Comentários:

Vamos chamar o atributo desejado, isto é, o fato de um componente ter duração inferior a 1400 horas de evento C. Pelo Teorema da Probabilidade Total, temos:

$$P(C) = P(C|A) \times P(A) + P(C|B) \times P(B)$$

Pelas informações do enunciado, sabemos que;

$$P(A) = \frac{n(A)}{n(U)} = \frac{120}{200} = 0{,}6$$

$$P(B) = \frac{n(B)}{n(U)} = \frac{80}{200} = 0{,}4$$

Sabendo que o componente A segue uma distribuição exponencial X, com média de 1000 horas, temos:

$$E(X) = \frac{1}{\lambda_X} = 1000$$

$$\lambda_X = \frac{1}{1000}$$

Sabendo que o componente B segue uma distribuição exponencial Y, com média de 700 horas, temos:

$$E(Y) = \frac{1}{\lambda_Y} = 700$$

$$\lambda_Y = \frac{1}{700}$$

Assim, podemos calcular as probabilidades condicionais:

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

$$P(C|A) = P(X < 1400) = 1 - e^{-\frac{1400}{1000}} = 1 - e^{-1{,}4} = 1 - 0{,}25 = 0{,}75$$

$$P(C|B) = P(Y < 1400) = 1 - e^{-\frac{1400}{700}} = 1 - e^{-2} = 1 - 0{,}14 = 0{,}86$$

Agora, podemos calcular o valor de um objeto aleatório durar menos de 1400 horas:

$$P(C) = 0{,}75 \times 0{,}6 + 0{,}86 \times 0{,}4 = 0{,}794$$

Gabarito: D

A distribuição exponencial é um caso **particular** da **distribuição gama**, cuja f.d.p. é:

$$f(y) = \begin{cases} \dfrac{\beta^{\alpha} y^{\alpha-1} e^{-\beta . y}}{\Gamma(\alpha)}, & se\ y \geq 0 \\ 0, & caso\ contrário \end{cases}$$

Em que $\Gamma(\alpha)$ é chamada de função gama.

A distribuição gama depende dos parâmetros $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$**.**

A distribuição gama se reduz à **distribuição exponencial** para $\boldsymbol{\alpha = 1}$, sendo $\boldsymbol{\beta = \lambda}$.

Observe essa relação nas fórmulas da média e variância da distribuição gama indicadas abaixo, em comparação com a distribuição exponencial X, em que $E(X) = \frac{1}{\lambda}$ e $V(X) = \frac{1}{\lambda^2}$:

$$E(Y) = \frac{\alpha}{\beta}, \qquad V(Y) = \frac{\alpha}{\beta^2}$$

Outra distribuição que apresenta determinada relação com a distribuição exponencial é a **distribuição de Weibull**, cuja f.d.p. é dada pela seguinte fórmula:

$$f(z) = \begin{cases} \dfrac{k}{\lambda}\left(\dfrac{z}{\lambda}\right)^{k-1} e^{-\left(\frac{z}{\lambda}\right)^k}, & se\ z \geq 0 \\ 0, & caso\ contrário \end{cases}$$

Podemos observar que essa distribuição depende dos parâmetros k e λ.

Compare a função de distribuição acumulada dessa distribuição, indicada abaixo, com a f.d.a. da distribuição exponencial X, em que $F(x) = 1 - e^{-\boldsymbol{\lambda} x}$:

$$F(Z) = 1 - e^{-\left(\frac{z}{\boldsymbol{\lambda}}\right)^{\boldsymbol{k}}}$$

Se Z segue **distribuição de Weibull**, então X, definida abaixo, segue **distribuição exponencial com $\boldsymbol{\lambda = 1}$**:

$$X = \left(\frac{Z}{\boldsymbol{\lambda}}\right)^{\boldsymbol{k}}$$

Essas duas distribuições pertencem à **família exponencial**.

Outra distribuição que também pertence a essa família é a **distribuição beta**, cuja f.d.p. é:

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha).\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & se\ 0 \leq x \leq 1 \\ 0, & caso\ contrário \end{cases}$$

Essa função depende dos parâmetros α e β, podendo ser escrita como:

$$f(x) = \frac{1}{\mathrm{B}(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

B é chamada de função beta, funcionando como uma constante de **normalização** para que a probabilidade de todo o Espaço Amostral seja igual a 1.

A média dessa distribuição é:

$$E(X) = \frac{\alpha}{\alpha+\beta}$$

(CESPE/2011 – EBC) Julgue o item subsequente, relativo à família exponencial de distribuições.

Tendo em vista que a distribuição exponencial é um caso particular da distribuição de Weibull, e considerando que a distribuição exponencial pertence à família exponencial, é correto concluir que a distribuição de Weibull também pertence à família exponencial.

Comentários:

A distribuição exponencial é um caso particular da distribuição **gama**. A exponencial também apresenta uma relação com a distribuição de Weibull, mas não pode ser considerada um caso particular desta. Todas essas distribuições pertencem à família exponencial.

Gabarito: Errado.

(CESPE/2014 – Analista Judiciário do TJ/SE) Nas estatísticas do Poder Judiciário, a taxa de congestionamento (X), que consiste em um indicador que permite medir a efetividade da movimentação processual de um tribunal, é uma variável aleatória contínua com função de densidade f(x) expressa por:

$f(x) = \begin{cases} \beta.x^8.(1-x)^2, se\ x \in [0,1] \\ 0, \qquad se\ x \notin [0,1] \end{cases}$, em que β é uma constante real

Com base nessas informações, julgue o próximo item.

O valor de β é superior a 450 e inferior a 500.

Comentários:

Observe que essa variável segue distribuição beta. O valor da constante β pode ser calculado, considerando-se que a probabilidade de todo o Espaço Amostral é 1. Para isso, precisamos integrar a f.d.p.:

$$F(x) = \int \beta.x^8.(1-x)^2.dx = \int \beta.x^8.(1-2x+x^2).dx = \int \beta.(x^8 - 2.x^9 + x^{10}).dx$$

$$F(x) = \int \beta.x^8.dx - \int 2.\beta.x^9 dx + \int \beta.x^{10} dx$$

Integrando em separado, temos:

$$\int \beta.x^8.dx = \beta\frac{x^9}{9}$$

$$\int 2.\beta.x^9 dx = 2.\beta\frac{x^{10}}{10}$$

$$\int \beta.x^{10} dx = \beta\frac{x^{11}}{11}$$

Juntando esses resultados, temos:

$$F(x) = \beta\frac{x^9}{9} - 2.\beta\frac{x^{10}}{10} + \beta\frac{x^{11}}{11}$$

Sabendo que x varia no intervalo [0,1], a diferença entre F(1) e F(0) corresponde à probabilidade de todo o Espaço Amostral:

$$F(1) - F(0) = \beta\frac{1^9}{9} - 2.\beta\frac{1^{10}}{10} + \beta\frac{1^{11}}{11} - \left(\beta\frac{0^9}{9} - 2.\beta\frac{0^{10}}{10} + \beta\frac{0^{11}}{11}\right) = 1$$

$$F(1) - F(0) = \frac{\beta}{9} - \frac{\beta}{5} + \frac{\beta}{11} = \frac{55.\beta - 99.\beta + 45.\beta}{495} = \frac{\beta}{495} = 1$$

$$\beta = 495$$

Logo, o valor de β é superior a 450 e inferior a 500.

Gabarito: Certo.

DISTRIBUIÇÃO NORMAL

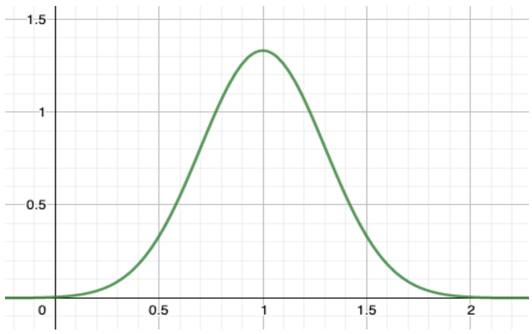
A distribuição **normal**, também chamada de **gaussiana**, é uma das distribuições contínuas mais importantes! A função densidade de probabilidade (f.d.p.) dessa distribuição é dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad x \in (-\infty, \infty)$$

Essa f.d.p. é bastante complicada, não é? Mas não se preocupe! Você não vai precisar integrar ou derivar!

Observe que essa função depende apenas dos **parâmetros μ** (média) e $\boldsymbol{\sigma^2}$ (variância), que são parâmetros **independentes**.

No gráfico abaixo, temos uma f.d.p. com distribuição normal. Observe que a curva apresenta um formato de **sino**, que é uma característica de todas as variáveis normais.



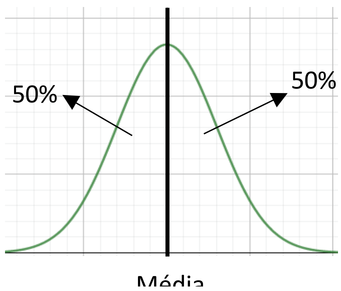
As distribuições normais são **simétricas**, ou seja, tem-se:

Média = Mediana = Moda

Logo, o valor de $\boldsymbol{\mu}$ divide a distribuição em **duas partes iguais**. Sabendo que a área total, sob toda a curva, corresponde à probabilidade de todo o Espaço Amostral e, portanto, a 100%, então:

$$\boxed{P(X > \mu) = P(X < \mu) = 50\%}$$

A probabilidade de um intervalo corresponde à **área** sob a f.d.p. limitada por esse intervalo. Assim, a igualdade acima pode ser ilustrada como no gráfico seguir:



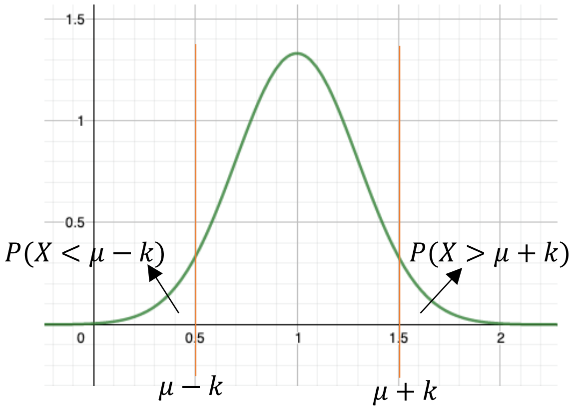
No exemplo do gráfico anterior, temos $\mu = 1$, logo:

$$P(X > 1) = P(X < 1) = 50\%$$

Mas a simetria não implica somente nisso. A partir da média, toda a distribuição de probabilidades para os valores superiores é igual à distribuição para os valores inferiores.

Assim, para qualquer k real, a probabilidade de a variável ser maior do que $\mu + k$ é igual à probabilidade de ser menor do que $\mu - k$:

$$\boxed{P(X > \mu + k) = P(X < \mu - k)}$$

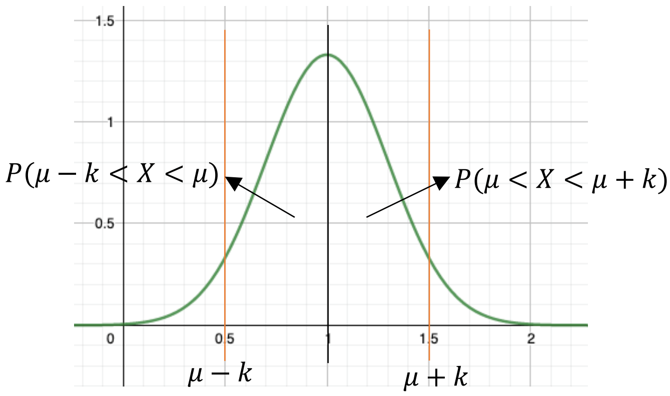


Em relação ao nosso exemplo, em que $\mu = 1$, temos:

- Para $k = 1$: $P(X > 2) = P(X < 0)$
- Para $k = 2$: $P(X > 3) = P(X < -1)$
- Para $k = 2{,}5$: $P(X > 3{,}5) = P(X < -1{,}5)$
- ...

Similarmente, as probabilidades associadas aos intervalos entre a média e esses limites $\mu + k$ e $\mu - k$ também são iguais, conforme equação abaixo e gráfico a seguir:

$$\boxed{P(\mu < X < \mu + k) = P(\mu - k < X < \mu)}$$



Em relação ao nosso exemplo em que $\mu = 1$, temos:

- Para $k = 1$: $P(1 < X < 2) = P(0 < X < 1)$
- Para $k = 2$: $P(1 < X < 3) = P(-1 < X < 1)$
- Para $k = 2{,}5$: $P(1 < X < 3{,}5) = P(-1{,}5 < X < 1)$
- ...

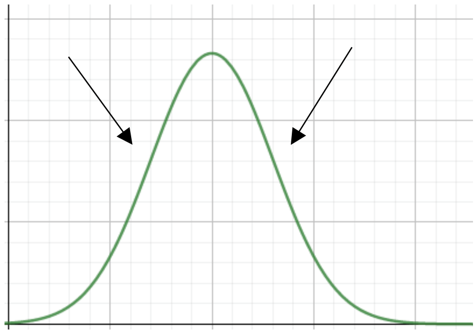
Podemos observar, ainda, que a curva normal apresenta **duas assíntotas**.

De modo geral, uma assíntota ocorre quando uma curva se **aproxima** cada vez mais a uma **reta**, porém **sem tocá-la**. A curva normal se **aproxima do eixo x** (eixo das abcissas) tanto para $x \to -\infty$, quanto para $x \to +\infty$. Por isso, dizemos que a curva normal é **duplamente assintótica**.

Além disso, existem **dois pontos de inflexão** na curva normal.

Pontos de inflexão são aqueles em que a **concavidade** da curva **muda**.

No início da curva normal, a concavidade está voltada para cima. No ponto (aproximado) indicado pela seta da esquerda, a **concavidade muda** para baixo, e no ponto (aproximado) indicado pela seta da direita, a **concavidade muda** novamente para cima.



Esses pontos de inflexão ocorrem precisamente a **1 desvio padrão** da média, ou seja, em $\boldsymbol{\mu - \sigma}$ e em $\boldsymbol{\mu + \sigma}$.

(FGV/2010 – SEAD-AP – *Adaptada*) Em relação à distribuição normal, julgue as afirmativas a seguir.

I – A função de densidade de probabilidade é simétrica em relação à média.

II – O valor da mediana é igual ao valor da média.

III – A média de uma variável aleatória com distribuição normal pode ser negativa.

Comentários:

Sabemos que a distribuição normal é simétrica em relação à média (logo, a afirmativa I está correta). Por ser simétrica, ela apresenta média = mediana (logo, a afirmativa II está correta).

A distribuição normal pode ter qualquer valor de média, inclusive negativa. Por exemplo, se estivermos tratando do lucro das empresas que vão à falência, provavelmente, a média será negativa. Logo, a afirmativa III está correta.

Resposta: Todas corretas.

Distribuição Normal Padrão

Para calcular os valores de probabilidade, temos uma **tabela** que relaciona os valores de intervalo da variável aos respectivos valores de probabilidade.

Essa tabela, inserida abaixo, se refere a uma distribuição normal $\boldsymbol{N(0,1)}$, isto é, com média $\boldsymbol{\mu = 0}$ e variância $\boldsymbol{\sigma^2 = 1}$, chamada de normal **padrão** ou **reduzida**, que denotamos por $\boldsymbol{Z}$.

Pelo gráfico anterior à tabela, deduzimos que os seus valores correspondem à probabilidade entre a média $\mu = 0$ e o valor de z indicado. Assim, os campos da tabela informam a probabilidade $P(0 < Z < z)$.

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817

Z (cont)	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

E quanto aos valores de z?

O valor de z começa a ser lido na primeira coluna (que apresenta as unidades e os décimos de z) e termina de ser lido na primeira linha (que apresenta os centésimos de z). Assim, a probabilidade $P(0 < Z < z)$ é o valor que está no campo, cuja linha corresponda à unidade e ao décimo de z e cuja coluna corresponda ao centésimo de z.

Por exemplo, para encontrar o valor de $P(0 < Z < 1{,}96)$, precisamos buscar o número que está na linha 1,9 e na coluna 0,06, conforme indicado abaixo. Podemos observar que $P(0 < Z < 1{,}96) = 0{,}475$.

Z	...	0,05	0,06	0,07	...
...	...	...	...	...	...
1,8	...	0,4678	0,4686	0,4693	...
1,9	...	0,4744	0,475	0,4756	...
2	...	0,4798	0,4803	0,4808	...
...	...	...	...	...	...

Também podemos fazer o caminho inverso, qual seja, encontrar o valor de z que corresponde à probabilidade desejada.

Vamos encontrar o valor de z tal que $P(0 < Z < z) = 0{,}40$, por exemplo. Para isso, devemos buscar o valor 0,40 nos campos da tabela. Como não consta exatamente esse valor, somente 0,3997 e 0,4015, optamos pelo valor mais próximo, isto é, 0,3997.

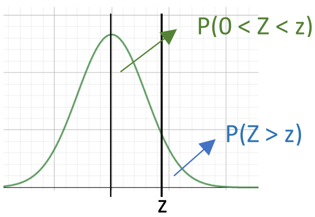
Este se encontra na linha 1,2 e na coluna 0,08, conforme indicado a seguir. Logo, concluímos que z = 1,28.

Z	...	0,07	0,08	0,09
...	...	0,3577	0,3599	0,3621
1,1	...	0,379	0,381	0,383
1,2	...	0,398	0,3997	0,4015
1,3	...	0,4147	0,4162	0,4177
...	...	...	...	...

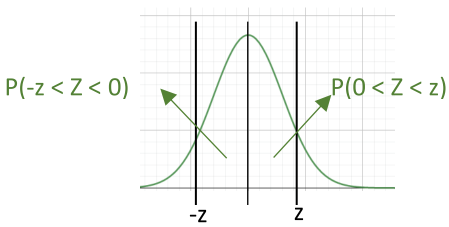
Para resolver questões envolvendo a tabela normal padrão é importante lembrar que essa distribuição é simétrica, com média $\mu = 0$.

TOME NOTA!

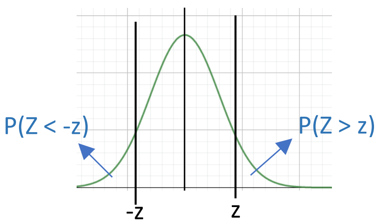
$$P(0 < Z < z) = 0,5 - P(Z > z)$$



$$P(-z < Z < 0) = P(0 < Z < z)$$

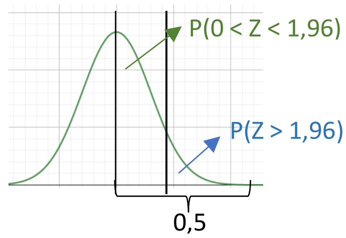


$$P(Z < -z) = P(Z > z)$$

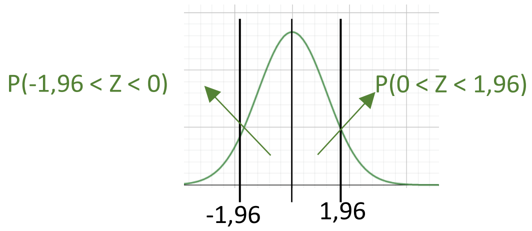


Supondo, por exemplo, z = 1,96, vimos que $P(0 < Z < 1{,}96) = 0{,}475$. Logo:

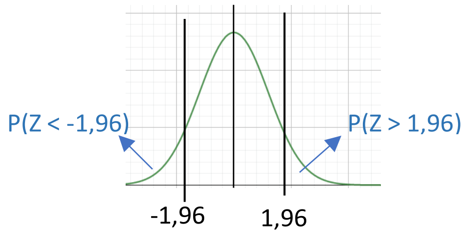
$$P(Z > 1,96) = 0,5 - P(0 < Z < 1,96) = 0{,}5 - 0{,}475 = 0{,}025$$



$$P(-1,96 < Z < 0) = P(0 < Z < 1,96) = 0{,}475$$

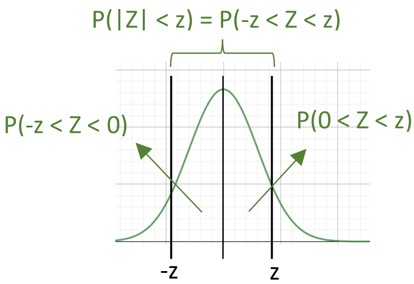


$$P(Z < -1,96) = P(Z > 1,96) = 0{,}5 - 0{,}475 = 0{,}025$$



A questão pode solicitar e/ou fornecer a probabilidade em módulo, da forma $P(|Z| < z)$ ou $P(|Z| > z)$. Para resolvê-las, é importante lembrar que a probabilidade $P(|Z| < z)$ corresponde a:

$$\boldsymbol{P(|Z| < z)} = P(-z < Z < z) = P(-z < Z < 0) + P(0 < Z < z)$$



Pela simetria da normal padrão, temos $P(-z < Z < 0) = P(0 < Z < z)$, logo:

$$\boldsymbol{P(|Z| < z) = 2 \times P(0 < Z < z)}$$

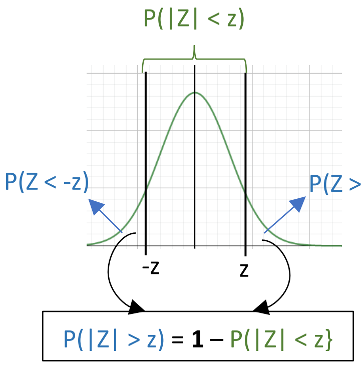
Supondo, por exemplo, z = 1,96, vimos que $P(0 < Z < 1{,}96) = 0{,}475$. Logo:

$$P(|Z| < 1.96) = 2 \times P(0 < Z < 1.96) = 2 \times 0{,}475 = 0{,}95$$

E a probabilidade $P(|Z| > z)$ pode ser calculada pela fórmula da probabilidade complementar:

$$\boldsymbol{P}(|\boldsymbol{Z}| > \boldsymbol{z}) = 1 - P(|Z| < z)$$

$$\boldsymbol{P(|Z| > z) = 1 - 2 \times P(0 < Z < z)}$$

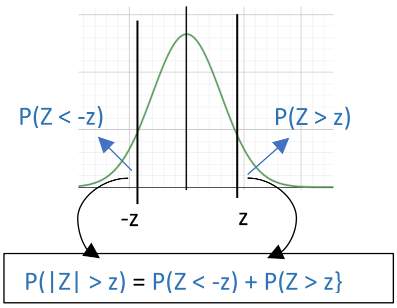


Supondo, por exemplo, z = 1,96, vimos que $P(0 < Z < 1{,}96) = 0{,}475$. Logo:

$$P(|Z| > z) = 1 - 2 \times P(0 < Z < z) = 1 - 2 \times 0{,}475 = 1 - 0{,}95 = 0{,}05$$

Ou, também podemos calcular $P(|Z| > z)$, aplicando-se o raciocínio análogo ao que fizemos anteriormente:

$$\boldsymbol{P}(|\boldsymbol{Z}| > \boldsymbol{z}) = P(Z < -z \cup Z > z) = P(Z < -z) + P(Z > z)$$



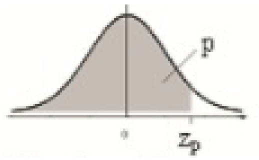
Pela simetria da normal padrão, temos $P(Z < -z) = P(Z > z)$, logo:

$$\boldsymbol{P}(|\boldsymbol{Z}| < \boldsymbol{z}) = 2 \times P(Z > z)$$

Para z = 1,96, em que $P(Z > 1{,}96) = 0{,}025$, temos:

$$P(|Z| > z) = 2 \times P(Z > z) = 2 \times 0{,}025 = 0{,}05$$

Existem, ainda, outros tipos de tabela para a distribuição normal padrão, que apresentam as probabilidades para outros intervalos, como por exemplo para a região indicada a seguir:



Esse tipo de tabela indica as probabilidades da forma $P(-\infty < Z < z)$, que correspondem à **função da distribuição normal acumulada**.

Transformação entre Distribuições Normais

Mas, e se a média da distribuição for diferente de zero e/ou a variância for diferente de 1?

Para isso, fazemos uma **transformação** de uma **distribuição normal qualquer** para a **distribuição normal padrão**, conforme fórmula indicada abaixo:

$$z = \frac{x-\mu}{\sigma}$$

Com essa transformação, encontramos os valores de **z** na distribuição normal padrão **associados** aos valores de **x** da distribuição normal de interesse, com média μ e desvio padrão σ.

Vamos supor uma distribuição normal com média $\mu = 1$ e desvio padrão $\sigma = 3$, em que estamos interessados no valor de $x = 7$. A transformação desse valor para a distribuição normal padrão é:

$$z = \frac{x-\mu}{\sigma} = \frac{7-1}{3} = 2$$

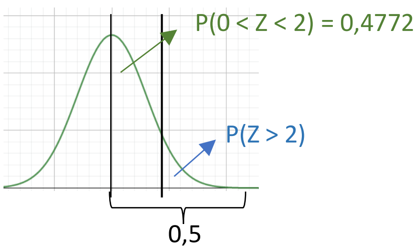
Isso significa que os **intervalos** associados a **z = 2** na distribuição **normal padrão** apresentam a **mesma probabilidade** daqueles associados a $\boldsymbol{x} = \mathbf{7}$ **na distribuição** $\boldsymbol{X}$, com média $\mu = 1$ e desvio padrão $\sigma = 3$.

Por exemplo:

$$\boldsymbol{P(X > 7) = P(Z > 2)}$$

Pela tabela da normal padrão, temos $P(0 < Z < 2) = 0{,}4772$, logo:

$$P(Z > 2) = 0{,}5 - 0{,}4772 = 0{,}0228$$



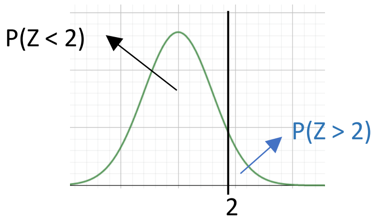
Portanto, $P(X > 7) = 0{,}0228 = 2{,}28\%$.

Analogamente, temos:

$$\boldsymbol{P(X < 7) = P(Z < 2)}$$

Sabemos que $P(Z > 2) = 0{,}5 - 0{,}4772 = 0{,}0228$, logo:

$$P(Z < 2) = 1 - P(Z > 2) = 1 - 0{,}0228 = 0{,}9772$$



Portanto, $P(X < 7) = 0{,}9772$.

Para encontrar intervalos envolvendo outros valores, por exemplo P(4 < X < 7), precisamos aplicar a transformação para ambos os valores. Para x = 4, temos:

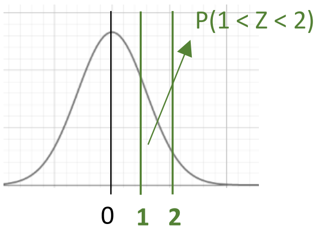
- $$z = \frac{x - \mu}{\sigma} = \frac{4 - 1}{3} = 1$$

Sabendo que a transformação para x = 7 é z = 2, então podemos concluir que:

P(4 < X < 7) = P(1 < Z < 2)

A probabilidade P(1 < Z < 2) pode ser calculada como:

P(1 < Z < 2) = P(0 < Z < 2) – P(0 < Z < 1)



Pela tabela da distribuição normal, observamos que P(0 < Z < 1) = 0,3413 e que P(0 < Z < 2)=0,4772, logo:

P(1 < Z < 2) = 0,4772 – 0,3413 = 0,1359

Assim, concluímos que P(4 < X < 7) = 0,1359 = 13,59%

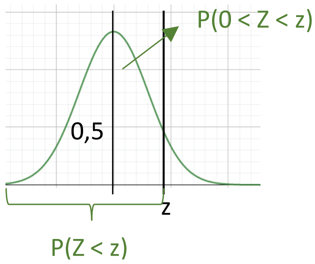
Também podemos fazer o **caminho inverso**, encontrando o **valor de x** em uma distribuição com média μ e desvio padrão σ, a partir de uma **probabilidade desejada**.

Para isso, primeiro encontramos o valor de **z** correspondente a essa probabilidade desejada, utilizando a tabela da normal padrão. Em seguida, aplicamos a fórmula da transformação.

Por exemplo, podemos calcular o valor de x para o qual a probabilidade de $P(X < x) = 0{,}8$, para a distribuição normal com os mesmos parâmetros do exemplo anterior ($\mu = 1$ e $\sigma = 3$).

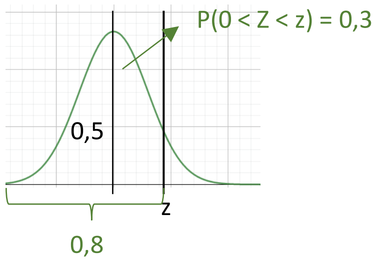
Considerando a simetria em torno de zero da normal padrão, temos que:

$$P(0 < Z < z) = P(Z < z) - 0{,}5$$



Assim, precisamos encontrar, na tabela normal padrão, o valor de z que corresponde a:

$$P(0 < Z < z) = 0{,}8 - 0{,}5 = 0{,}3$$



Pela tabela da distribuição normal padrão, observamos que esse valor é $z = 0{,}84$, pois P(0<Z<0,84) = 0,2995, que é o valor da tabela mais próximo de 0,3.

Substituindo os valores conhecidos na fórmula transformação ($\mu = 1$, $\sigma = 3$ e $z = 0{,}84$), podemos encontrar o valor de x, que delimita uma probabilidade $P(X < x) = 0{,}8$:

$$z = \frac{x - \mu}{\sigma}$$

$$0{,}84 = \frac{x - 1}{3}$$

$$x = 3 \times 0{,}84 + 1 = 3{,}52$$

Agora, vamos calcular as probabilidades associadas a intervalos genéricos, em função do **desvio padrão** σ.

Para calcular a probabilidade $\boldsymbol{P(\mu - \sigma < X < \mu + \sigma)}$, utilizamos a seguinte transformação, para $\boldsymbol{x = \mu + \sigma}$:

$$z = \frac{x - \mu}{\sigma} = z = \frac{\mu + \sigma - \mu}{\sigma} = 1$$

Para $\boldsymbol{x = \mu - \sigma}$, temos:

$$z = \frac{x - \mu}{\sigma} = z = \frac{\mu - \sigma - \mu}{\sigma} = -1$$

Portanto:

$$\boldsymbol{P(\mu - \sigma < X < \mu + \sigma)} = P(-1 < Z < 1) = \boldsymbol{P(-1 < Z < 0) + P(0 < Z < 1)}$$

De maneira geral, quando os intervalos são da forma $P(\mu - k < X < \mu + k)$, em que os extremos são **equidistantes** da média, basta fazermos a transformação para um dos extremos, pois o outro estará associado ao **mesmo valor de z**, porém multiplicado por –1.

Pela tabela da curva normal, temos $P(0 < Z < 1) = 0{,}3413$. Pela simetria da normal padrão, em torno da média 0, temos:

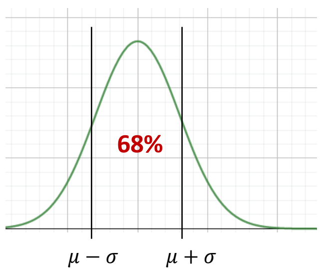
$$P(-1 < Z < 0) = P(0 < Z < 1) = 0{,}3413$$

Logo:

$$P(-1 < Z < 1) = P(-1 < Z < 0) + P(0 < Z < 1) = 0{,}3413 + 0{,}3413 = 0{,}6826$$

$$\boldsymbol{P(\mu - \sigma < X < \mu + \sigma) \cong 68\%}$$

Ou seja, a probabilidade de uma variável normal qualquer se afastar da média (para cima ou para baixo) em até **1 desvio padrão** é aproximadamente **68%**, conforme ilustrado a seguir.



Para $\boldsymbol{P(\mu - 2\sigma < X < \mu + 2\sigma)}$, temos:

$$z = \frac{x - \mu}{\sigma} = \frac{\mu + 2\sigma - \mu}{\sigma} = 2$$

Portanto:

$$\boldsymbol{P(\mu - 2\sigma < X < \mu + 2\sigma)} = P(-2 < Z < 2) = \boldsymbol{P(-2 < Z < 0)} + \boldsymbol{P(0 < Z < 2)}$$

Pela tabela, temos $P(0 < Z < 2) = 0{,}4772$. Considerando a simetria da normal padrão, temos:

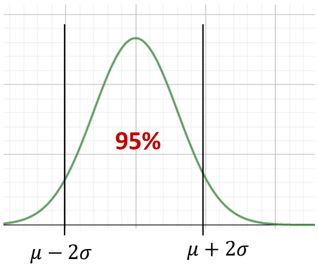
$$P(-2 < Z < 0) = P(0 < Z < 2) = 0{,}4772$$

E a probabilidade desejada é:

$$P(-2 < Z < 2) = P(-2 < Z < 0) + P(0 < Z < 2) = 0{,}4772 + 0{,}4772 = 0{,}9544$$

$$\boldsymbol{P(\mu - 2\sigma < X < \mu + 2\sigma) \cong 95\%}$$

Ou seja, a probabilidade de uma variável normal qualquer se afastar da média (para cima ou para baixo) em até **2 desvios padrão** é aproximadamente **95%**, como ilustrado abaixo.



Para $\boldsymbol{P(\mu - 3\sigma < X < \mu + 3\sigma)}$, obtemos, pela transformação, $z = 3$, logo:

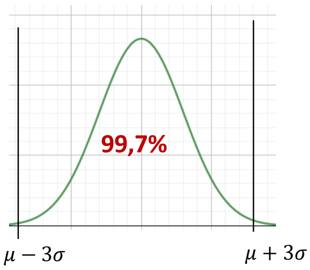
$$\boldsymbol{P(\mu - 3\sigma < X < \mu + 3\sigma)} = P(-3 < Z < 3) = \boldsymbol{P(-3 < Z < 0)} + \boldsymbol{P(0 < Z < 3)}$$

Pela tabela, temos $P(0 < Z < 3) = 0{,}4987$ e, pela simetria, $P(-3 < Z < 0) = 0{,}4987$, logo:

$$P(-3 < Z < 3) = P(-3 < Z < 0) + P(0 < Z < 3) = 0{,}4987 + 0{,}4987 = 0{,}9974$$

$$\boldsymbol{P(\mu - 3\sigma < X < \mu + 3\sigma) \cong 99{,}7\%}$$

Ou seja, a probabilidade de uma variável normal qualquer se afastar da média (para cima ou para baixo) em até **3 desvios padrão** é aproximadamente **99,7%**, ilustrado abaixo.



Essas probabilidades (**68%** para $\boldsymbol{\mu \pm \sigma}$; **95%** para $\boldsymbol{\mu \pm 2\sigma}$ e **99,7%** para $\boldsymbol{\mu \pm 3\sigma}$) compõem a chamada **Regra Empírica**. Algumas questões (não muitas) exigem que você memorize essas probabilidades.

Distribuição Normal: $N(\mu, \sigma^2)$

Simétrica, com formato de sino, definida em **toda a reta real**

Regra Empírica: 68% para $\boldsymbol{\mu \pm \sigma}$; **95%** para $\boldsymbol{\mu \pm 2\sigma}$ e **99,7%** para $\boldsymbol{\mu \pm 3\sigma}$

Transformação para a **Normal Padrão** N(0,1): $\boldsymbol{z = \frac{x-\mu}{\sigma}}$

(FGV/2010 – SEAD-AP – *Adaptada*) Em relação à distribuição normal, julgue as afirmativas a seguir:

I – Se X tem distribuição normal com média μ e variância σ^2 então a variável $Z = \frac{(X-\mu)}{\sigma^2}$ tem distribuição normal padrão.

II – A probabilidade de que uma variável Z que tenha distribuição normal padrão seja maior que 5 é aproximadamente igual a 0.

Comentários:

Em relação à afirmativa I, a transformação para a Normal Padrão é $Z = \frac{(X-\mu)}{\sigma}$, ou seja, a divisão é pelo desvio padrão, não pela variância. Por isso, a afirmativa I está incorreta.

Em relação à afirmativa II, a distribuição se concentra em 3 desvios-padrão para ambos os lados (mais de 99% se encontram nesse intervalo). De fato, as tabelas da normal padrão, em geral, fornecem valores até z=3,99 porque valores a probabilidade de Z ser maior que isso é praticamente nula. Logo, a afirmativa II está correta.

Resposta: I – incorreta, II – correta.

(VUNESP/2009 – CETESB) Para um determinado horário, considerando-se todos os dias de um período, ao se calcular a média de congestionamento de trânsito em km obtém-se o valor μ e desvio padrão σ. Considerando-se que os valores obtidos pela variável e suas respectivas probabilidades constituem uma distribuição normal, no intervalo de $(\mu - \sigma)$ até $(\mu + \sigma)$, a percentagem dos dados contidos é cerca de

a) 25%

b) 50%

c) 68%

d) 94%

e) 99%

Comentários:

Essa questão exige o conhecimento da **Regra Empírica**. Vimos que o intervalo de $(\mu - \sigma)$ até $(\mu + \sigma)$ concentra 68% da distribuição normal.

Gabarito: C

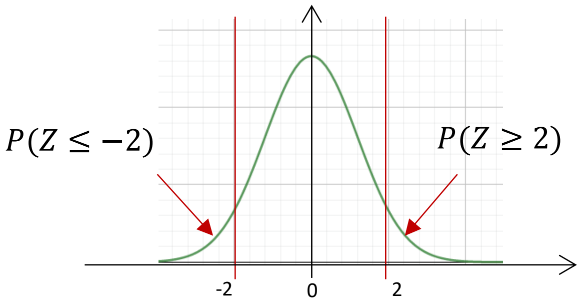
(CESPE/2016 – Analista da FUNPRESP-JUD) A simetria de Z implica que P(Z ≥ 2) = 1 – P(Z ≤ -2).

Comentários:

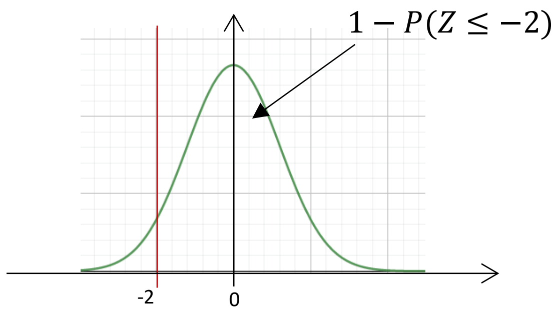
A simetria da curva normal com média igual a zero implica na seguinte relação entre P(Z ≥ 2) e P(Z ≤ -2):

$$P(Z \geq 2) = P(Z \leq -2)$$

Essa relação está ilustrada no gráfico abaixo.



A relação descrita no enunciado iguala a probabilidade $P(Z \geq 2)$ à probabilidade $1 - P(Z \leq -2)$. Esta corresponde a toda a região indicada abaixo:



Podemos observar que $P(Z \geq 2)$ é bem menor que 50%, enquanto $1 - P(Z \leq -2)$ é bem maior que 50%, ou seja:

$$P(Z \geq 2) \neq 1 - P(Z \leq -2)$$

Logo, o item está errado.

Gabarito: Errado.

(FGV/2022 – EPE) O salário médio dos funcionários de uma empresa é normalmente distribuído com média de R$ 2.500,00 e desvio padrão de R$ 1.500,00. A empresa divide os funcionários em 5 classes, a saber: M, N, O, P e Q, onde "M" é a classe com melhor salário e "Q" a classe com menor salário.

Se apenas 5% dos funcionários dessa empresa estão na classe "M", o menor valor do salário do funcionário para ele pertencer à classe "M" é

[Considere que P(Z ≤ 1,64) = 0,95.]

a) 3900,00

b) 4170,00

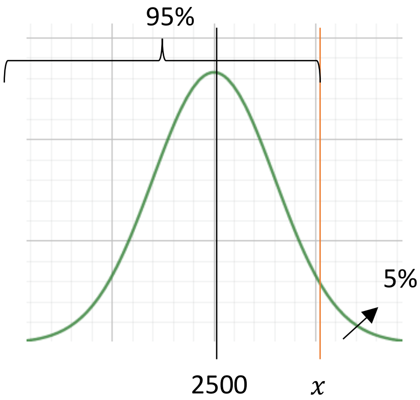
c) 4960,00

d) 5160,00

e) 5350,00

Comentários:

O enunciado informa que os salários seguem distribuição normal com média $\mu = \text{R\$ } 2.500{,}00$ e desvio padrão $\sigma = \text{R\$}1.500{,}00$; e pede o valor do menor salário da classe M, associada aos 5% melhores salários, conforme ilustrado a seguir:



Sabendo que 5% (ou 0,05) da distribuição é maior do que o valor buscado, então 95% (ou 0,95) da distribuição é **menor** e o enunciado informa justamente que P(Z ≤ 1,64) = 0,95.

Assim, devemos aplicar a transformação para z = 1,64, sabendo que a média é 2500 e o desvio padrão é 1500:

$$z = \frac{x - \mu}{\sigma}$$

$$1{,}64 = \frac{x - 2500}{1500}$$

$$x - 2500 = 1{,}64 \times 1500 = 2460$$

$$x = 4960$$

Gabarito: C

(VUNESP/2014 – EMPLASA) O tempo de vida da população de um determinado país tem distribuição normal com a média igual a 68 anos e o desvio padrão igual a 11.

Considere os valores da tabela e a fórmula $Z = \frac{X-\mu}{\sigma}$:

Z	Distribuição normal reduzida
0,5	0,1915
1,0	0,3413
1,5	0,4332
2,0	0,4772
2,5	0,4938

A probabilidade de uma pessoa viver mais do que 90 anos é de

a) 15,87%

b) 6,68%

c) 4,82%

d) 3,36%

e) 2,28%

Comentários:

Sendo $\mu = 68$ e $\sigma = 11$, temos a seguinte transformação para $x = 90$:

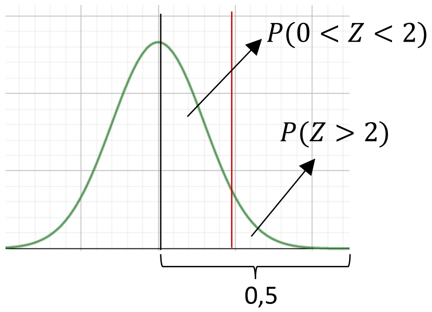
$$z = \frac{x-\mu}{\sigma} = \frac{90-68}{11} = \frac{22}{11} = 2$$

Pela tabela, que apresenta a probabilidade $P(0 < Z < z)$, temos:

$$P(0 < Z < 2) = 0{,}4772$$

Logo:

$$P(Z > 2) = 0{,}5 - P(0 < Z < 2) = 0{,}5 - 0{,}4772 = 0{,}0228$$



Gabarito: E.

(FCC/2015 – Auditor Fiscal da SEFAZ/PI) Se Z tem distribuição normal padrão, então: P(Z < 0,4) = 0,655; P(Z < 1,2) = 0,885; P(Z < 1,6) = 0,945; P(Z < 1,8) = 0,964; P(Z < 2) = 0,977.

O efeito do medicamento A é o de baixar a pressão arterial de indivíduos hipertensos. O tempo, em minutos, decorrido entre a tomada do remédio e a diminuição da pressão é uma variável aleatória X com distribuição normal, tendo média μ e desvio padrão σ. Se o valor de μ é de 56 min e o valor de σ é de 10 min, a probabilidade de X estar compreendido entre 52 min e 74 min é igual a

a) 30,9%

b) 56,0%

c) 61,9%

d) 52,4%

e) 64,5%

Comentários:

A probabilidade de X estar entre 52 min e 74 min pode ser calculada a partir das transformações para x = 74 min e para x = 52 min, considerando a média de 56 min e desvio padrão de 10 min.

Para x = 74, temos:

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{74 - 56}{10} = \frac{18}{10} = 1{,}8$$

A transformação para x = 52 é:

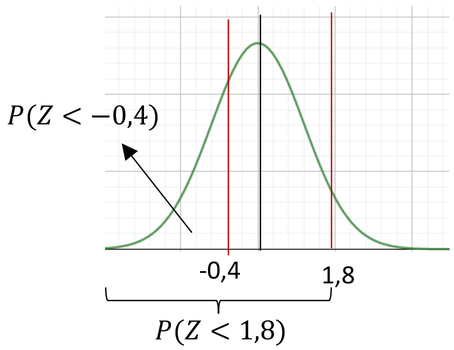
$$z = \frac{52 - 56}{10} = \frac{-4}{10} = -0{,}4$$

Então, a probabilidade de X estar compreendido entre 52 min e 74 min corresponde a:

$$P(52 < X < 74) = P(-0{,}4 < Z < 1{,}8)$$

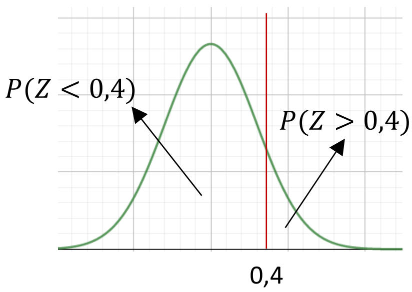
Essa probabilidade pode ser calculada como:

$$P(-0{,}4 < Z < 1{,}8) = P(Z < 1{,}8) - P(Z < -0{,}4)$$



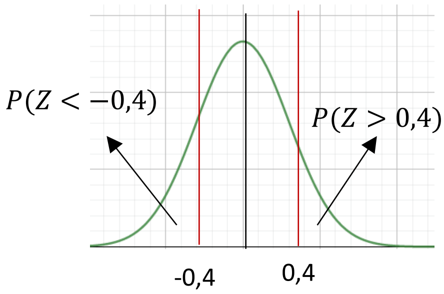
Pela tabela observamos que P(Z < 1,8) = 0,964. Além disso, temos P(Z < 0,4) = 0,655, logo, o seu complementar é:

P(Z > 0,4) = 1 – 0,655 = 0,345



Pela simetria da normal padrão, temos:

P(Z < -0,4) = P(Z > 0,4) = 0,345



Inserindo esses valores na equação acima, temos:

$$P(52 < X < 74) = P(Z < 1{,}8) - P(Z < -0{,}4) = 0{,}964 - 0{,}345 = 0{,}619$$

Gabarito: C

SOMA DE VARIÁVEIS E O TEOREMA CENTRAL DO LIMITE

Nesta seção, veremos a **soma** de variáveis com distribuição normal e, também, com outras distribuições. Ao final, veremos como aproximar uma distribuição binomial a uma normal.

Soma de Variáveis com Distribuição Normal

A **soma** de variáveis **independentes** com distribuição normal também segue uma distribuição **normal**, cuja média corresponde à **soma das médias** das variáveis e a variância corresponde à **soma das variâncias**.

> Se $X_1, X_2, \ldots, X_n$ são variáveis **independentes** que seguem distribuição **normal**, então a soma $Y = X_1 + X_2 + \cdots + X_n$ também segue distribuição **normal** com média e variância dadas respectivamente por:
>
> $$E(Y) = E(X_1) + E(X_2) + \cdots + E(X_n)$$
>
> $$V(Y) = V(X_1) + V(X_2) + \cdots + V(X_n)$$

Ademais, a **diferença** entre duas variáveis **independentes** com distribuição normal também segue uma distribuição **normal**, cuja média e variância podem ser calculadas pelas propriedades dessas medidas.

> Se X_1 e X_2 são variáveis **independentes** com distribuição **normal**, então a diferença $Y = X_1 - X_2$ segue distribuição **normal** com média e variância dadas respectivamente por:
>
> $$E(Y) = E(X_1) - E(X_2)$$
>
> $$V(Y) = V(X_1) + V(X_2)$$

Por exemplo, sendo $Y = X_1 - X_2$, em que X_1 e X_2 são variáveis normais independentes, com médias $E(X_1) = 3$ e $E(X_2) = 5$, e variâncias $V(X_1) = 4$ e $V(X_2) = 1$, então, Y terá **distribuição normal**, com parâmetros:

$$E(Y) = E(X_1) - E(X_2) = 3 - 5 = -2$$

$$V(Y) = V(X_1) + V(X_2) = 4 + 1 = 5$$

Pontue-se que a **soma**, a **subtração**, a **multiplicação** ou a **divisão** de uma distribuição normal por uma **constante** real **também** segue **distribuição normal**, cuja média e variância podem ser calculadas pelas propriedades de esperança e variância.

Por exemplo, sendo X uma variável normal com média E(X) = 3 e variância V(X) = 4, então a variável Y = 2X – 6 terá **distribuição normal**, com média e variância dadas por:

$$E(Y) = E(2X - 6) = 2.E(X) - 6 = 2 \times 3 - 6 = 0$$

$$V(Y) = V(2X - 6) = 2^2.V(X) = 4 \times 4 = 16$$

(2020 – FADESP) A variável aleatória X tem distribuição normal com média $\mu = 2$ e variância $\sigma^2 = 9$. Seja Y uma variável aleatória definida por Y = 2X + 1. Nestas condições, pode-se afirmar que Y tem distribuição

a) normal com média $\mu = 2$ e variância $\sigma^2 = 30$.

b) qui-quadrado com média $\mu = 5$ e variância $\sigma^2 = 36$.

c) normal com média $\mu = 5$ e variância $\sigma^2 = 9$.

d) normal com média $\mu = 5$ e variância $\sigma^2 = 36$.

Comentários:

Vimos que a soma de variáveis normais segue uma distribuição normal, mesmo quando somadas a uma constante. Logo, a variável Y possui distribuição normal, com média:

$$E(Y) = E(2X + 1) = 2.E(X) + 1$$

Sendo E(X) = 2, temos:

$$E(Y) = 2.2 + 1 = 5$$

A variância é:

$$V(Y) = V(2X + 1) = 4V(X)$$

Sendo V(X) = 36, então:

$$V(Y) = 4.9 = 36.$$

Gabarito: D.

(2018 – Petrobras) As variáveis aleatórias X e Y são independentes. A variável X segue uma distribuição Normal com média 4 e variância 16, e a Y segue uma distribuição Normal com média 9 e variância 1. A distribuição de X - Y é Normal com

a) média -5 e variância 15.

b) média -5 e variância 17.

c) média 5 e variância 15.

d) média 5 e variância 17.

e) média 13 e variância 15.

Comentários:

Vimos que a diferença de variáveis normais segue uma distribuição normal, com média:

$$\mu = \mu_X - \mu_Y = 4 - 9 = -5$$

E variância:

$$\sigma^2 = \sigma_X^2 + \sigma_Y^2 = 16 + 1 = 17$$

Gabarito: B.

(FCC/2015 – Analista do CNMP – *Adaptada*) Se Z tem distribuição normal padrão, então: P(Z < 0,5) = 0,591; P(Z < 1) = 0,841; P(Z < 1,15) = 0,8951; P(Z < 1,17) = 0,879; P(Z < 1,2) = 0,885; P(Z < 1,4) = 0,919; P(Z < 1,64) = 0,95; P(Z < 2) = 0,977; P(Z < 2,06) = 0,98; P(Z < 2,4) = 0,997.

Sejam X_1, X_2, ... , X_n variáveis aleatórias independentes e identicamente distribuídas com distribuição normal padrão. Seja a variável aleatória $Y = X_1 + X_2 + \cdots + X_n$.

Considerando essa informação, julgue a seguinte afirmativa.

Para n = 4, P(− 2 < Y < 1) = 0,432.

Comentários:

A soma de n = 4 variáveis aleatórias independentes com distribuição normal padrão N(0,1) segue distribuição normal com média e variância dadas por:

$$E(Y) = E(X_1) + E(X_2) + E(X_3) + E(X_4) = 0 + 0 + 0 + 0 = 0$$

$$V(Y) = V(X_1) + V(X_2) + V(X_3) + V(X_4) = 1 + 1 + 1 + 1 = 4$$

Logo, o desvio padrão é:

$$\sigma = \sqrt{V(Y)} = 2$$

Para calcular a probabilidade desejada, utilizaremos a fórmula da transformação para a normal padrão. O valor de z para y = 1 é:

$$z = \frac{y - \mu}{\sigma}$$

$$z = \frac{1 - 0}{2} = 0{,}5$$

E o valor de z para y = -2 é:

$$z = \frac{-2 - 0}{2} = -1$$

Logo, a probabilidade desejada corresponde a:

P(− 2 < Y < 1) = P(− 1 < Z < 0,5) = P(Z < 0,5) – P(Z < -1)

Pela tabela fornecida, observamos que P(Z < 0,5) = 0,591. Temos também que P(Z < 1) = 0,841, logo o seu complementar é:

P(Z > 1) = 1 – 0,841 = 0,159

Pela simetria da normal padrão, temos:

P(Z < -1) = P(Z > 1) = 0,159

Assim, a probabilidade P(-2 < Y < 1) corresponde a:

P(Z < 0,5) – P(Z < -1) = 0,591 – 0,159 = 0,432

Resposta: Certo.

Teorema Central do Limite

Um dos motivos pelos quais a **Distribuição Normal** (ou de **Gauss**) é tão importante em Estatística decorre do **Teorema Central do Limite** (TLC), que trata da soma de variáveis que seguem uma distribuição **qualquer**:

*Para variáveis aleatórias $X_1, X_2, \dots, X_n$ **independentes e identicamente distribuídas** (i.i.d), a distribuição da soma $X_1 + X_2 + \cdots + X_n$ **tende** a uma distribuição **normal**, à medida em que **n cresce**, cuja média e variância são dadas por:*

$$E(Y) = E(X_1) + E(X_2) + \cdots + E(X_n) = n.E(X)$$

$$V(Y) = V(X_1) + V(X_2) + \cdots + V(X_n) = n.V(X)$$

Note que as variáveis podem apresentar **qualquer** distribuição, mesmo assim, a sua soma seguirá **aproximadamente** uma distribuição **normal**.

A convergência de que trata o TLC é uma **convergência em distribuição**. A convergência em distribuição de uma sequência X_n a uma variável X significa que **as funções de distribuição acumuladas** de X_n e de X convergem.

As funções densidade de probabilidade não necessariamente convergem, ou seja, **não** se trata de uma convergência ponto a ponto. A convergência em distribuição é o tipo mais **fraco** de convergência, dentre os usuais.

Vamos supor que haja 100 variáveis **independentes** X_i, com i = 1, 2, ..., 100, todas com média **$E(X_i) = 3$** e variância **$V(X_i) = 4$**. Assim, sendo $Y = X_1 + X_2 + ... + X_{100}$, então a média e variância de Y serão:

$$E(Y) = 100 \times E(X_i) = 100 \times 3 = 300$$

$$V(Y) = 100 \times V(X_i) = 100 \times 4 = 400$$

Ainda que as variáveis **não** sejam **identicamente distribuídas**, isto é, ainda que apresentem distribuições **distintas**, mesmo assim, a sua soma terá, **aproximadamente**, uma distribuição normal, desde que as variáveis sejam **independentes** e tenham variâncias similares.

Nesse caso, a esperança e variância serão:

$$E(Y) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

$$V(Y) = V(X_1) + V(X_2) + \cdots + V(X_n)$$

Atenção: devemos somar as **variâncias**, **não** os desvios padrão!

(FGV/2017 – IBGE) Sejam $X_1, X_2, X_3, ..., X_{64}$ variáveis aleatórias discretas, com distribuição Binomial, todas com p = 0,25 e n = 12. Também são conhecidos valores da função distribuição acumulada da normal-padrão, mais especificamente: $\phi(2) = 0{,}977$, $\phi(1{,}5) = 0{,}933$, $\phi(1{,}25) = 0{,}894$.

No caso da extração de uma amostra (n = 64), a probabilidade de que a soma dos valores seja superior a 207 é igual a:

a) 0,023

b) 0,046

c) 0,067

d) 0,106

e) 0,134

Comentários:

Pelo Teorema Central do Limite, a soma de N variáveis independentes identicamente distribuídas X_i, cada uma com média E(X) e variância V(X), é uma variável Y, com distribuição aproximadamente normal, com média e variância dadas por:

$$E(Y) = N.E(X)$$

$$V(Y) = N.V(X)$$

Sendo X uma variável binomial com n = 12 e p = 0,25 (logo, q = 1 – p = 0,75), então:

$$E(X) = n.p = 12 \times 0{,}25 = 3$$

$$V(X) = n.p.q = 12 \times 0{,}25 \times 0{,}75 = 9/4$$

Assim, a distribuição da soma de N = 64 variáveis X terá média e variância dadas por:

$$E(Y) = N.E(X) = 64 \times 3 = 192$$

$$V(Y) = N.V(X) = 64 \times \frac{9}{4} = 144$$

O desvio padrão de Y é, portanto:

$$\sigma_Y = \sqrt{V(X)} = \sqrt{144} = 12$$

Assim, a probabilidade $P(Y > 207)$ pode ser calculada pela seguinte transformação:

$$z = \frac{y - \mu}{\sigma} = \frac{207 - 192}{12} = \frac{15}{12} = 1{,}25$$

O enunciado informa que $P(-\infty < Z < 1{,}25) = 0{,}894$. Logo:

$$P(Z > 1{,}25) = 1 - P(-\infty < Z < 1{,}25) = 1 - 0{,}894 = 0{,}106$$

Gabarito: D.

Aproximação da Binomial pela Normal

Como consequência do **Teorema Central do Limite**, podemos **aproximar** uma distribuição **binomial** a uma distribuição **normal** Y, quando o número de ensaios $\boldsymbol{n}$ for grande.

Nesse caso, a média e a variância da distribuição são, respectivamente:

$$E(Y) = n.p$$

$$V(Y) = n.p.(1 - p)$$

Assim, podemos aproximar a probabilidade de uma variável **binomial** X apresentar valores dentro de um intervalo $[a, b]$, ou seja, $P(a \leq X \leq b)$, à probabilidade de uma variável **normal** Y apresentar valores no referido intervalo:

$$P(a \leq X \leq b) \cong P(a < Y < b)$$

EXEMPLIFICANDO

Considerando uma variável binomial X, com $n = 50$ e $p = 0{,}5$, vamos calcular a probabilidade associada ao intervalo $[20, 30]$.

Pela distribuição binomial, teríamos que calcular:

$$P(20 \leq X \leq 30) = P(X = 20) + P(X = 21) + \cdots + P(X = 30)$$

$$P(20 \leq X \leq 30) = \binom{50}{20} 0{,}5^{20}. 0{,}5^{30} + \binom{50}{21} 0{,}5^{21}. 0{,}5^{29} + \cdots + \binom{50}{30} 0{,}5^{30}. 0{,}5^{20}$$

Cansativo, não é? O resultado dessa conta é aproximadamente 0,8811 (*confia em mim*).

Aproximando essa distribuição a uma variável Y que segue uma distribuição **normal**, com média $\mu_Y = n.p = 25$ e variância $\sigma_Y^2 = n.p.(1-p) = 12{,}5$ (desvio padrão $\sigma_Y \cong 3{,}5355$), temos a seguinte transformação para $x = 30$:

$$z = \frac{x-\mu}{\sigma} \cong \frac{30-25}{3{,}5255} \cong 1{,}41$$

Pela tabela normal, podemos observar que $P(0 < Z < 1{,}41) = 0{,}4207$.

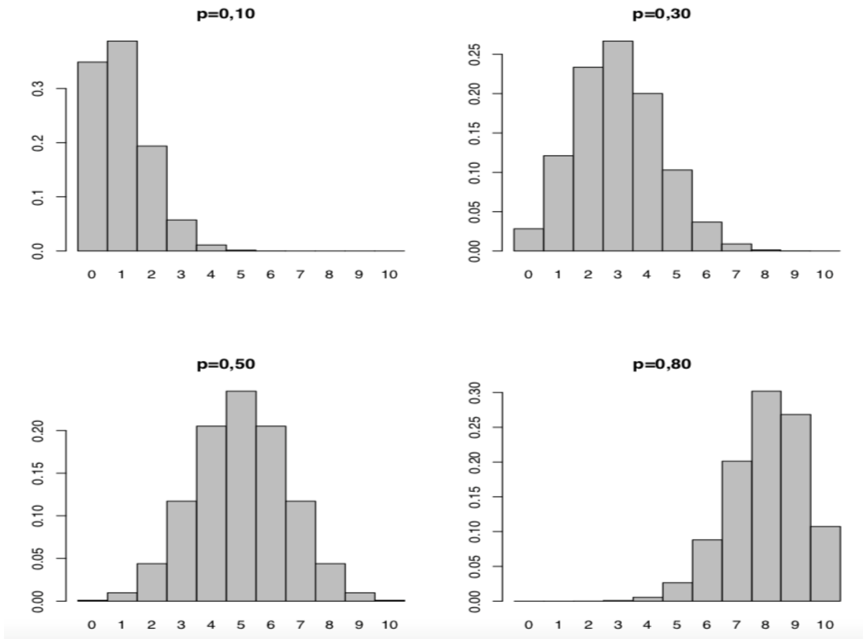
Considerando que o intervalo [20, 30] é simétrico em relação à média $\mu_Y = 25$, então:

$$P(20 \leq X \leq 30) \cong P(-1{,}41 < Z < 1{,}41) = 2 \times 0{,}4207 = 0{,}8414$$

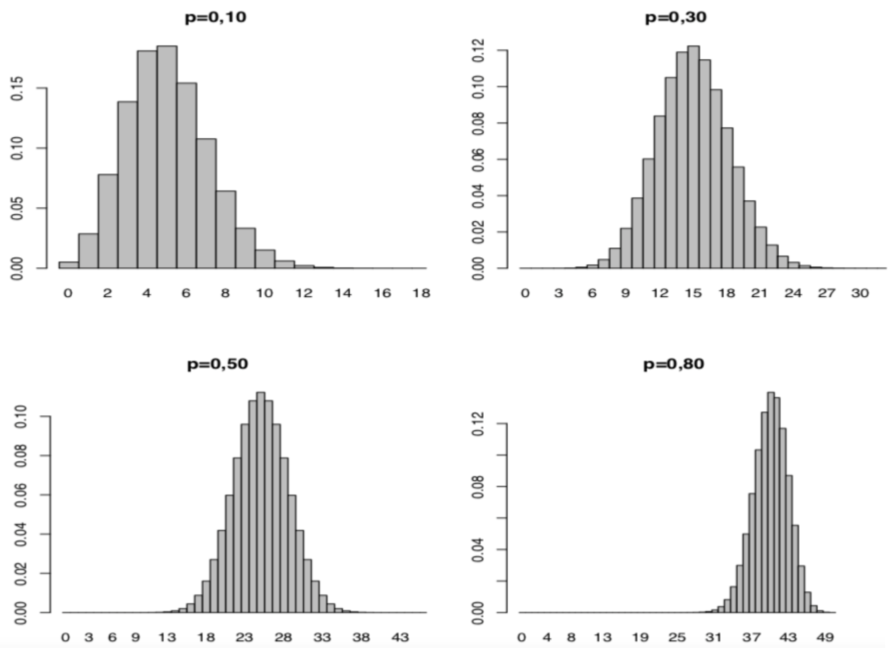
O número exato de ensaios necessários que permite essa aproximação é controverso (e depende do grau de precisão desejado). O que se sabe é que quanto **maior n** e quanto **mais próximo de $^1/_2$ é o valor de p**, mais a distribuição binomial se aproxima da normal.

A seguir, apresento os histogramas[1] de distribuições binomiais, para $n = 10$. Note que o histograma para $p = 0{,}5$ é bem próximo de uma normal, mesmo para um número relativamente pequeno de ensaios (10). Por outro lado, o histograma para $p = 0{,}1$ está bem distante de uma normal.

[1] Histogramas obtidos das aulas de Bacharelado em Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA/USP), disponível em https://www.ime.usp.br/~yambar/MAE0219/Aula%208%20Teorema%20do%20Limite%20Central/Aula%208-Teorema%20do%20Limite%20Central.pdf



Vejamos agora os histogramas para $n = 50$:



Nesse caso, todos os histogramas são próximos de uma curva normal, mesmo para $p = 0{,}1$.

Existe uma regra empírica que exige que o **produto** entre $\boldsymbol{n}$ e a **menor probabilidade entre $\boldsymbol{p}$ e $\boldsymbol{q}$** deve ser **maior que 5**, ou, para uma melhor aproximação, maior que 15.

Na prática, a questão deve fornecer elementos que permitirão concluir que a distribuição binomial pode ser aproximada a uma normal.

Correção de Continuidade

Por se tratar de uma distribuição **discreta** sendo aproximada a uma distribuição **contínua**, para melhorar a precisão dos resultados, é necessário introduzir uma **correção de continuidade**.

Para isso, devemos **aumentar** o intervalo da binomial da forma $[a, b]$, isto é, com os **extremos incluídos**, em **0,5** unidade para cada extremo. Em outras palavras, acrescentamos 0,5 unidade ao extremo superior e subtraímos 0,5 unidade do extremo inferior.

Assim, o intervalo da distribuição normal correspondente será $(a - 0{,}5, b + 0{,}5)$:

$$P(a \leq X \leq b) \cong P(a - 0{,}5 < Y < b + 0{,}5)$$

Por exemplo, o intervalo [2, 4] em uma distribuição binomial (discreta) corresponde ao seguinte intervalo na distribuição normal (contínua):

$$P(2 \leq X \leq 4) \cong P(2 - 0{,}5 < Y < 4 + 0{,}5)$$

$$P(2 \leq X \leq 4) \cong P(1{,}5 < Y < 4{,}5)$$

Se o intervalo desejado **não incluir** um dos extremos, ou ambos, então primeiro **ajustamos** o intervalo para incluí-lo(s).

Por exemplo, vamos calcular a probabilidade de obter **mais de 3** sucessos e **menos de 8** sucessos, o que corresponde ao intervalo (3, 8), isto é, **sem** os extremos.

Para utilizar a aproximação da binomial à normal, primeiro **ajustamos o intervalo** para que os extremos sejam **incluídos**.

Ora, o evento **mais de 3** sucessos e **menos de 8** sucessos, em uma distribuição binomial (discreta), equivale ao evento **4 ou mais** sucessos e **7 ou menos** sucessos, o que corresponde ao intervalo [4, 7], isto é, com os extremos incluídos.

$$P(3 < X < 8) = P(4 \leq X \leq 7)$$

Após o ajuste do intervalo, podemos aplicar a **correção de continuidade**, somando-se 0,5 unidade ao extremo superior e subtraindo 0,5 unidade do extremo inferior:

$$P(4 \leq X \leq 7) \cong P(4 - 0{,}5 < Y < 7 + 0{,}5)$$

$$P(4 \leq X \leq 7) \cong P(3{,}5 < Y < 7{,}5)$$

EXEMPLIFICANDO

Vamos, então, refazer o nosso exemplo com a correção de continuidade:

$$P(20 \leq X \leq 30) \cong P(19{,}5 < Y < 30{,}5)$$

Para $x = 30{,}5$, com $\mu = 25$ e $\sigma \cong 3{,}5355$, temos a seguinte transformação:

$$z = \frac{x-\mu}{\sigma} \cong \frac{30{,}5-25}{3{,}5255} \cong 1{,}56$$

Pela tabela normal, podemos observar que $P(0 < Z < 1{,}56) = 0{,}4406$. Logo:

$$P(20 \leq X \leq 30) \cong P(-1{,}56 < Z < 1{,}56) = 2 \times 0{,}4406 = 0{,}8812$$

Note que esse resultado é extremamente próximo do resultado exato, qual seja, 0,8811!

Logo, para resolver questões envolvendo a aproximação de uma distribuição binomial X a uma distribuição normal Y, precisamos:

i. Calcular a média e a variância da distribuição, $E(Y) = E(X) = n.p$ e $V(Y) = V(X) = n.p.q$
ii. Aplicar a correção de continuidade $P(a \leq X \leq b) \cong P(a - 0{,}5 < Y < b + 0{,}5)$
iii. Utilizar a transformação para a normal padrão $z = \frac{x-\mu}{\sigma}$
iv. Consultar a tabela da normal padrão para encontrar os valores de probabilidade

HORA DE PRATICAR!

(CESPE/2013 – TRT 17ª Região) No que se refere a distribuições discretas, julgue o seguinte item.

A aproximação da distribuição binomial pela normal não se aplica com base no teorema limite central, visto que a binomial não se relaciona com uma soma de variáveis aleatórias.

Comentários:

O Teorema Central do Limite garante que, para n suficientemente grande, a distribuição binomial pode sim ser aproximada pela distribuição Normal.

Pontue-se, ainda, que a distribuição binomial é uma soma de variáveis de Bernoulli.

Gabarito: Errado.

(FGV/2017 – MPE/BA) A probabilidade de que uma decisão de 1ª instância da Justiça Federal do Paraná seja reformada pelo Tribunal Superior da 4ª Região é de 0,20. No momento 100 recursos aguardam por uma decisão dos Srs. Desembargadores daquele Tribunal.

São informados alguns valores da distribuição acumulada da normal-padrão:

Ø(1) = 0,87, Ø(1,28)=0,90 e Ø(2) = 98

Sem usar o ajuste de continuidade, a probabilidade de que mais de 24 decisões sejam reformadas é:

a) 13%

b) 10%

c) 8%

d) 5%

e) 2%

Comentários:

A questão trabalha com a aproximação da binomial à normal. Sabendo que a probabilidade de reforma de decisão é p = 0,20 e que há n = 100 recursos, então a média da distribuição é:

$$\mu = n \times p = 100 \times 0{,}2 = 20$$

Sendo a probabilidade de não reforma de decisão complementar q = 1 – p = 0,9, a variância é:

$$V(X) = n.p.q = 100 \times 0{,}2 \times 0{,}8 = 16$$

E o desvio padrão:

$$\sigma = \sqrt{V(X)} = \sqrt{16} = 4$$

Considerando que não deve ser utilizada a correção de continuidade, então a transformação de x = 24 para a normal padrão é:

$$z = \frac{x-\mu}{\sigma} = \frac{24-20}{4} = \frac{4}{4} = 1$$

Pelas informações fornecidas, observamos que a probabilidade acumulada é $P(Z \leq 1) = 0{,}87$. Logo, a probabilidade de haver mais decisões reformadas é complementar:

$$P(Z > 1) = 1 - P(Z \leq 1) = 1 - 0{,}87 = 0{,}13 = 13\%$$

Gabarito: A.

DISTRIBUIÇÃO QUI-QUADRADO

A distribuição qui-quadrado resulta da **soma** de distribuições **normais reduzidas** (ou padrão) **independentes**, elevadas ao **quadrado**.

Em outras palavras, a distribuição qui-quadrado $\mathcal{X}_k^2$ corresponde à soma dos quadrados de k variáveis com distribuição normal padrão, $Z_i \sim N(0,1)$:

$$\mathcal{X}_k^2 = \sum_{i=1}^{k} (Z_i)^2$$

Dizemos que a distribuição apresenta ***k* graus de liberdade**, sendo ***k*** o número de variáveis **normais** que compõem a distribuição qui-quadrado. O número de graus de liberdade é o **único parâmetro** da distribuição.

Em particular, para $k = 1$, isto é, havendo uma única variável normal padrão elevada ao quadrado Z^2, temos uma distribuição qui-quadrado com **um** grau de liberdade, $\mathcal{X}_1^2$.

Pontue-se que, se as variáveis normais X_i **não forem reduzidas**, ou seja, se apresentarem média μ_i e desvio padrão σ_i, é necessário aplicar a **transformação** para a normal reduzida, antes de elevar ao quadrado, para formar a distribuição qui-quadrado padrão:

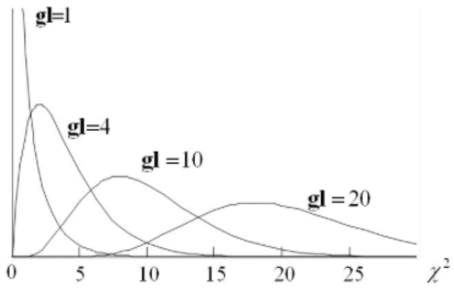
$$\mathcal{X}_k^2 = \sum_{i=1}^{k} (Z_i)^2 = \sum_{i=1}^{k} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2$$

FIQUE ATENTO!

Quando distribuições normais com média $\mu \neq 0$ **não padronizadas** são utilizadas para formar uma distribuição qui-quadrada, dizemos que a distribuição resultante é **não central**, com **parâmetro de não centralidade** igual à **média** μ das variáveis normais.

Para a distribuição qui-quadrado, os diferentes graus de liberdade reproduzem funções densidade de probabilidade distintas, conforme ilustrado a seguir[1].

[1] Gráfico do professor Ivan Balducci da Universidade Estadual Paulista, FOSJC/UNESP, disponível em https://slideplayer.com.br/slide/45581/.



Podemos observar que as variáveis com distribuição qui-quadrado são **positivas** (afinal correspondem à soma do **quadrado** de variáveis) e **assimétricas à direita**.

Observamos, ainda, que quanto **maior o grau de liberdade**, isto é, quanto mais variáveis normais ao quadrado são somadas, mais **simétrica** será a distribuição.

Inclusive, pelo **Teorema Central do Limite**, a distribuição qui-quadrado se **aproxima a uma distribuição normal**, à medida que o número de graus de liberdade $\boldsymbol{k}$ aumenta.

A **soma** de variáveis com distribuição qui-quadrado com $k_1, k_2, \ldots, k_n$ graus de liberdade, segue uma distribuição **qui-quadrado** com $k_1 + k_2 + \cdots + k_n$ graus de liberdade.

A **média** da distribuição qui-quadrado é igual ao número de graus de liberdade $\boldsymbol{k}$; e a **variância** é igual ao dobro do número de graus de liberdade $\boldsymbol{2k}$:

$$\boldsymbol{E(\mathcal{X}_k^2) = k}$$

$$\boldsymbol{V(\mathcal{X}_k^2) = 2k}$$

A distribuição qui-quadrado também possui uma **tabela** que associa uma probabilidade $P(\mathcal{X}_k^2 < x)$ ou $P(\mathcal{X}_k^2 > x)$ a um valor positivo x, de acordo com o grau de liberdade k.

Pelo fato de a distribuição **não** ser **simétrica**, a distribuição acima da média é diferente daquela abaixo da média. Por isso, a tabela apresenta os valores de probabilidade para **toda** a distribuição.

A tabela a seguir apresenta os valores para $P(\mathcal{X}_k^2 < x)$, indicados na primeira linha. A primeira (assim como a última) coluna apresenta os graus de liberdade k, que a tabela indica como n.

Por exemplo, a mediana para uma distribuição com 5 graus de liberdade, isto é, o valor de x para o qual $P(\mathcal{X}_5^2 < x) = 50\%$, conforme se observa na linha $n = 5$ e na coluna $P = 0{,}5$, é $x = 4{,}351$.

	$P(\chi_n^2 \le x)$													
n	0,005	0,01	0,025	0,05	0,1	0,25	0,5	0,75	0,9	0,95	0,975	0,99	0,995	
1	3,93E-05	0,000157	0,000982	0,003932	0,016	0,102	0,455	1,323	2,706	3,841	5,024	6,635	7,879	1
2	0,010	0,020	0,051	0,103	0,211	0,575	1,386	2,773	4,605	5,991	7,378	9,210	10,597	2
3	0,072	0,115	0,216	0,352	0,584	1,213	2,366	4,108	6,251	7,815	9,348	11,345	12,838	3
4	0,207	0,297	0,484	0,711	1,064	1,923	3,357	5,385	7,779	9,488	11,143	13,277	14,860	4
5	0,412	0,554	0,831	1,145	1,610	2,675	4,351	6,626	9,236	11,070	12,832	15,086	16,750	5
6	0,676	0,872	1,237	1,635	2,204	3,455	5,348	7,841	10,645	12,592	14,449	16,812	18,548	6
7	0,989	1,239	1,690	2,167	2,833	4,255	6,346	9,037	12,017	14,067	16,013	18,475	20,278	7
8	1,344	1,647	2,180	2,733	3,490	5,071	7,344	10,219	13,362	15,507	17,535	20,090	21,955	8
9	1,735	2,088	2,700	3,325	4,168	5,899	8,343	11,389	14,684	16,919	19,023	21,666	23,589	9
10	2,156	2,558	3,247	3,940	4,865	6,737	9,342	12,549	15,987	18,307	20,483	23,209	25,188	10
11	2,603	3,053	3,816	4,575	5,578	7,584	10,341	13,701	17,275	19,675	21,920	24,725	26,757	11
12	3,074	3,571	4,404	5,226	6,304	8,438	11,340	14,845	18,549	21,026	23,337	26,217	28,300	12
13	3,565	4,107	5,009	5,892	7,041	9,299	12,340	15,984	19,812	22,362	24,736	27,688	29,819	13
14	4,075	4,660	5,629	6,571	7,790	10,165	13,339	17,117	21,064	23,685	26,119	29,141	31,319	14
15	4,601	5,229	6,262	7,261	8,547	11,037	14,339	18,245	22,307	24,996	27,488	30,578	32,801	15
16	5,142	5,812	6,908	7,962	9,312	11,912	15,338	19,369	23,542	26,296	28,845	32,000	34,267	16
17	5,697	6,408	7,564	8,672	10,085	12,792	16,338	20,489	24,769	27,587	30,191	33,409	35,718	17
18	6,265	7,015	8,231	9,390	10,865	13,675	17,338	21,605	25,989	28,869	31,526	34,805	37,156	18
19	6,844	7,633	8,907	10,117	11,651	14,562	18,338	22,718	27,204	30,144	32,852	36,191	38,582	19
20	7,434	8,260	9,591	10,851	12,443	15,452	19,337	23,828	28,412	31,410	34,170	37,566	39,997	20
21	8,034	8,897	10,283	11,591	13,240	16,344	20,337	24,935	29,615	32,671	35,479	38,932	41,401	21
22	8,643	9,542	10,982	12,338	14,041	17,240	21,337	26,039	30,813	33,924	36,781	40,289	42,796	22
23	9,260	10,196	11,689	13,091	14,848	18,137	22,337	27,141	32,007	35,172	38,076	41,638	44,181	23
24	9,886	10,856	12,401	13,848	15,659	19,037	23,337	28,241	33,196	36,415	39,364	42,980	45,558	24
25	10,520	11,524	13,120	14,611	16,473	19,939	24,337	29,339	34,382	37,652	40,646	44,314	46,928	25
26	11,160	12,198	13,844	15,379	17,292	20,843	25,336	30,435	35,563	38,885	41,923	45,642	48,290	26
27	11,808	12,878	14,573	16,151	18,114	21,749	26,336	31,528	36,741	40,113	43,195	46,963	49,645	27
28	12,461	13,565	15,308	16,928	18,939	22,657	27,336	32,620	37,916	41,337	44,461	48,278	50,994	28
29	13,121	14,256	16,047	17,708	19,768	23,567	28,336	33,711	39,087	42,557	45,722	49,588	52,335	29
30	13,787	14,953	16,791	18,493	20,599	24,478	29,336	34,800	40,256	43,773	46,979	50,892	53,672	30
40	20,707	22,164	24,433	26,509	29,051	33,660	39,335	45,616	51,805	55,758	59,342	63,691	66,766	40
50	27,991	29,707	32,357	34,764	37,689	42,942	49,335	56,334	63,167	67,505	71,420	76,154	79,490	50
60	35,534	37,485	40,482	43,188	46,459	52,294	59,335	66,981	74,397	79,082	83,298	88,379	91,952	60
70	43,275	45,442	48,758	51,739	55,329	61,698	69,334	77,577	85,527	90,531	95,023	100,425	104,215	70
80	51,172	53,540	57,153	60,391	64,278	71,145	79,334	88,130	96,578	101,879	106,629	112,329	116,321	80
90	59,196	61,754	65,647	69,126	73,291	80,625	89,334	98,650	107,565	113,145	118,136	124,116	128,299	90
100	67,328	70,065	74,222	77,929	82,358	90,133	99,334	109,141	118,498	124,342	129,561	135,807	140,170	100

INDO MAIS FUNDO!

A distribuição qui-quadrado é um **caso particular** da distribuição **gama**, cuja f.d.p. é:

$$f(x) = \begin{Bmatrix} \dfrac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}, & se\ x \ge 0 \\ 0, & caso\ contrário \end{Bmatrix}$$

A f.d.p. da distribuição **qui-quadrado** é obtida com $\boldsymbol{\alpha = \frac{k}{2}}$ e $\boldsymbol{\beta = \frac{1}{2}}$.

Por exemplo, para $\boldsymbol{k = 2}$, temos uma distribuição **qui-quadrado** com 2 graus de liberdade, o que corresponde a uma distribuição **gama** com $\boldsymbol{\alpha = 1}$ e $\boldsymbol{\beta = \frac{1}{2}}$.

Nessa situação, em que $\alpha = 1$, a distribuição gama se reduz a uma distribuição **exponencial**, com $\beta = \lambda$. Então, uma distribuição **qui-quadrado** com **2 graus de liberdade** é também uma distribuição **exponencial**, com $\boldsymbol{\lambda = \beta = \frac{1}{2}}$.

A **média** dessa distribuição é $E(X) = \frac{1}{\lambda} = \mathbf{2}$ ou $E(\mathcal{X}_k^2) = k = \mathbf{2}$.

ESQUEMATIZANDO

Distribuição Qui-quadrado com $\boldsymbol{k}$ **graus de liberdade**
Soma de $\boldsymbol{k}$ variáveis **normais padrão** elevadas ao quadrado: $\mathcal{X}_k^2 = \sum_{i=1}^{k}(Z_i)^2$
Assimétrica à direita, com $x \geq 0$
Esperança: $\boldsymbol{E(\mathcal{X}_k^2) = k}$; **Variância**: $\boldsymbol{V(\mathcal{X}_k^2) = 2k}$

HORA DE PRATICAR!

(FCC/2015 – Analista do CNMP – *Adaptada*) Sejam X_1, X_2, ... , X_n variáveis aleatórias independentes e identicamente distribuídas com distribuição normal padrão e seja a variável aleatória $W = X_1^2 + X_2^2 + \cdots + X_n^2$. Considerando essa informação, julgue a seguinte afirmativa:

A variável W tem distribuição qui-quadrado com (n − 1) graus de liberdade.

Comentários:

Sendo $W = X_1^2 + X_2^2 + \cdots + X_n^2$, então W apresenta distribuição qui-quadrado com k = **n** graus de liberdade (não, n – 1).

Resposta: Errado.

(FGV/2021 – FunSaúde/CE) Se X_1, X_2,...X_n são n variáveis aleatórias independentes e identicamente distribuídas com distribuição N(μ,σ²), então a variável

$$Q = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2}$$

Tem a seguinte distribuição de probabilidades:

a) normal com média $n\mu$ e variância $n\sigma^2$.

b) qui-quadrado com n graus de liberdade

c) qui-quadrado com n-1 graus de liberdade.

d) t-Student com n graus de liberdade.

e) t-Student com n-1 graus de liberdade

Comentários:

A variável Q indicada no enunciado corresponde à soma de n variáveis normais, reduzidas à normal padrão, $Z_i = \frac{X_i - \mu_i}{\sigma_i}$, elevadas ao quadrado.

Assim, a variável Q apresenta distribuição qui-quadrado com k = n graus de liberdade.

Gabarito: B

(CESPE/2018 – Agente de Polícia Federal) Uma amostra aleatória simples $Y_1, Y_2, ..., Y_{25}$ foi retirada de uma distribuição normal com média nula e variância σ^2, desconhecida. Considerando que $P(x^2 \leq 13) = P(x^2 > 41) = 0,025$, em que x^2 representa a distribuição qui-quadrado com 25 graus de liberdade, e que $S^2 = \sum_{i=1}^{25} Y_i^2$, julgue o item a seguir.

A variância da distribuição X^2 com 25 graus de liberdade é superior a 40.

Comentários:

A variância da distribuição qui-quadrado, com k = 25 graus de liberdade, é:

$$V(X) = 2k = 2 \times 25 = 50$$

Assim, a variância é superior a 40.

Gabarito: Certo.

(CESPE/2018 – Analista Administrativo – EBSERH) Considerando que X e Y sejam variáveis aleatórios mutuamente independentes que seguem distribuição normal padrão, julgue o próximo item.

A soma dos quadrados $Q = X^2 + Y^2$ segue uma distribuição exponencial com média igual a 2.

Comentários:

A soma de k variáveis independentes com distribuição normal padrão elevadas ao quadrado segue distribuição **qui-quadrado**, com k graus de liberdade. Nesse caso, temos uma variável com distribuição qui-quadrado com k = 2. A média dessa distribuição é:

$$E(X) = k = 2$$

Especialmente para k = 2, a distribuição qui-quadrado corresponde a uma distribuição exponencial, com $\lambda = \frac{1}{2}$.

Gabarito: Certo.

DISTRIBUIÇÃO T DE STUDENT

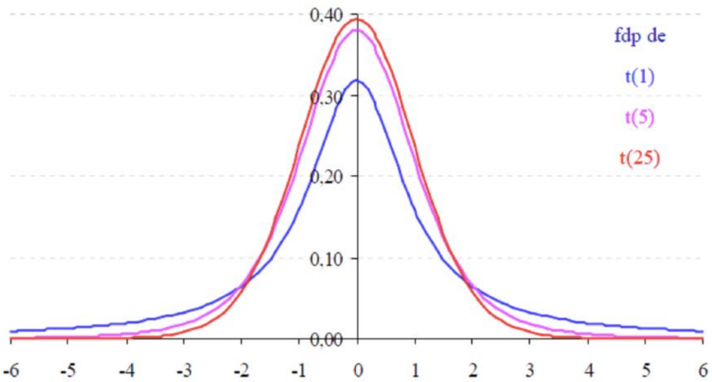
A variável T com distribuição **t de Student** (ou **t-Student**) é definida como:

$$T = \frac{Z}{\sqrt{\frac{\chi_k^2}{k}}}$$

em que Z é uma variável **normal padrão**, χ_k^2 é uma variável com distribuição **qui-quadrado**, com k graus de liberdade, sendo Z e χ_k^2 **independentes**.

A variável T apresenta **k graus de liberdade**, que também é o **único parâmetro** da distribuição.

O gráfico a seguir[1] apresenta as funções densidade de probabilidade para distribuições t-Student com 1, 5 e 25 graus de liberdade.



Observa-se que, assim como a normal padrão, a distribuição t-Student é **simétrica**, com média $\mu = 0$ e formato de **sino**, assumindo valores em toda a reta real, $(-\infty, \infty)$.

Porém, em comparação com a normal, a curva **t-Student** é mais **baixa** na região **central** e mais **larga** nos **extremos**. Ou seja, a distribuição t-Student apresenta **maior variabilidade**.

A **variância** da distribuição de t-Student é dada por:

$$V(T) = \frac{k}{k-2}, \ para\ k > 2$$

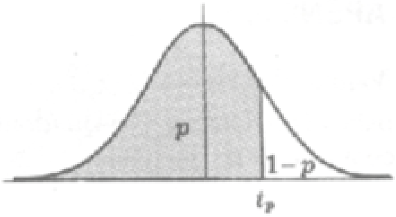
Pontue-se que a variância **não** é definida um grau de liberdade menor ou igual a 2 ($k \leq 2$).

[1] Obtido na apresentação das aulas da Prof. Tarciana Liberal, da Universidade Federal da Paraíba, disponível em http://www.de.ufpb.br/~tarciana/Probabilidade2/Aula15.pdf

A **variância da distribuição** de t-Student é sempre **maior que 1**, para qualquer valor de k, pois a razão $\frac{k}{k-2}$ é sempre um valor **superior a 1**. Porém, à medida que k aumenta, essa razão se **reduz**, **aproximando-se de 1**.

Assim, à medida que o grau de liberdade ***k* aumenta**, a distribuição de t-Student se **aproxima** de uma curva **normal**, com **variância igual a 1** (outra consequência do **Teorema Central do Limite**).

A distribuição t-Student também apresenta uma **tabela** que associa aos valores de probabilidade os intervalos de valores de t, de acordo com o grau de liberdade, conforme indicado a seguir. Observe pelo gráfico anterior a essa tabela, que os seus valores de probabilidade são da forma $P(T < t)$.



ν	$t_{0,55}$	$t_{0,60}$	$t_{0,70}$	$t_{0,75}$	$t_{0,80}$	$t_{0,90}$	$t_{0,95}$	$t_{0,975}$	$t_{0,99}$	$t_{0,995}$
1	,158	,325	,727	1,000	1,376	3,08	6,31	12,71	31,82	63,66
2	,142	,289	,617	,816	1,061	1,89	2,92	4,30	6,96	9,92
3	,137	,277	,584	,765	,978	1,64	2,35	3,18	4,54	5,84
4	,134	,271	,569	,741	,941	1,53	2,13	2,78	3,75	4,60
5	,132	,267	,559	,727	,920	1,48	2,02	2,57	3,36	4,03
6	,131	,265	,553	,718	,906	1,44	1,94	2,45	3,14	3,71
7	,130	,263	,549	,711	,896	1,42	1,90	2,36	3,00	3,50
8	,130	,262	,546	,706	,889	1,40	1,86	2,31	2,90	3,36
9	,129	,261	,543	,703	,883	1,38	1,83	2,26	2,82	3,25
10	,129	,260	,542	,700	,879	1,37	1,81	2,23	2,76	3,17
11	,129	,260	,540	,697	,876	1,36	1,80	2,20	2,72	3,11
12	,128	,259	,539	,695	,873	1,36	1,78	2,18	2,68	3,06
13	,128	,259	,538	,694	,870	1,35	1,77	2,16	2,65	3,01
14	,128	,258	,537	,692	,868	1,34	1,76	2,14	2,62	2,98
15	,128	,258	,536	,691	,866	1,34	1,75	2,13	2,60	2,95
16	,128	,258	,535	,690	,865	1,34	1,75	2,12	2,58	2,92
17	,128	,257	,534	,689	,863	1,33	1,74	2,11	2,57	2,90
18	,127	,257	,534	,688	,862	1,33	1,73	2,10	2,55	2,88
19	,127	,257	,533	,688	,861	1,33	1,73	2,09	2,54	2,86
20	,127	,257	,533	,687	,860	1,32	1,72	2,09	2,53	2,84
21	,127	,257	,532	,686	,859	1,32	1,72	2,08	2,52	2,83
22	,127	,256	,532	,686	,858	1,32	1,72	2,07	2,51	2,82
23	,127	,256	,532	,685	,858	1,32	1,71	2,07	2,50	2,81
24	,127	,256	,531	,685	,857	1,32	1,71	2,06	2,49	2,80
25	,127	,256	,531	,684	,856	1,32	1,71	2,06	2,48	2,79
26	,127	,256	,531	,684	,856	1,32	1,71	2,06	2,48	2,78
27	,127	,256	,531	,684	,855	1,31	1,70	2,05	2,47	2,77
28	,127	,256	,530	,683	,855	1,31	1,70	2,05	2,47	2,76
29	,127	,256	,530	,683	,854	1,31	1,70	2,04	2,46	2,76
30	,127	,256	,530	,683	,854	1,31	1,70	2,04	2,46	2,75
40	,126	,255	,529	,681	,851	1,30	1,68	2,02	2,42	2,70
60	,126	,254	,527	,679	,848	1,30	1,67	2,00	2,39	2,66
120	,126	,254	,526	,677	,845	1,29	1,66	1,98	2,36	2,62
∞	,126	,253	,524	,674	,842	1,28	1,645	1,96	2,33	2,58

A primeira coluna apresenta os graus de liberdade, que a tabela chama de v, e a primeira linha apresenta os valores de probabilidade acumulada, até os valores de t indicados nos campos. Por exemplo, para 5 graus de liberdade (5ª linha), o valor que delimita uma probabilidade $P(T < t) = 0{,}9$ (6ª coluna), é $t = 1{,}48$.

Lembre-se que a distribuição é **simétrica**, com média, mediana e moda iguais a $\mu = 0$. Assim:

$$P(T < -t) = P(T > t)$$

$$P(T < 0) = P(T > 0) = 0{,}5$$

Por isso, são apresentados apenas valores de probabilidades $P(T < t)$ maiores que 50%, uma vez que à esquerda da média, os valores são **simétricos**.

Assim como para a curva normal padrão, há outras formas de apresentar a tabela de t-Student.

INDO MAIS FUNDO!

Ressalte-se que a distribuição t-Student para $k = 1$ consiste na **razão entre duas variáveis normais padrão** independentes:

$$T = \frac{Z_1}{\sqrt{\frac{\chi_1^2}{1}}} = \frac{Z_1}{\sqrt{Z_2^2}} = \frac{Z_1}{Z_2}$$

Nesse caso específico, temos uma distribuição de **Cauchy**, que corresponde à **razão entre duas variáveis normais** independentes.

Essa distribuição também é **simétrica** e **mais achatada** em relação à normal. Porém, ela é uma distribuição "**patológica**", pois **não** possui **média** ou **variância** definidas.

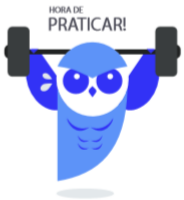
ESQUEMATIZANDO

Distribuição t-Student, definida para todos os valores reais

$$T = \frac{Z}{\sqrt{\frac{\chi_k^2}{k}}}$$

Simétrica, com $\mu = 0$ e formato de **sino mais achatado** do que a normal

Variância: $V(X) = \frac{k}{k-2},\ para\ k > 2$



(2017 – TRF 2ª Região – *Adaptada*) Sobre a distribuição t-Student, julgue a afirmativa a seguir.

Considere uma variável aleatória Z com distribuição normal padrão e uma outra variável aleatória V com distribuição qui-quadrado e v graus de liberdade. Se Z e V forem independentes, então a variável aleatória $T = \frac{Z}{\sqrt{\frac{V}{v}}}$ tem distribuição t de *Student* com v graus de liberdade.

Comentários:

Vimos que a distribuição t-Student é definida como:

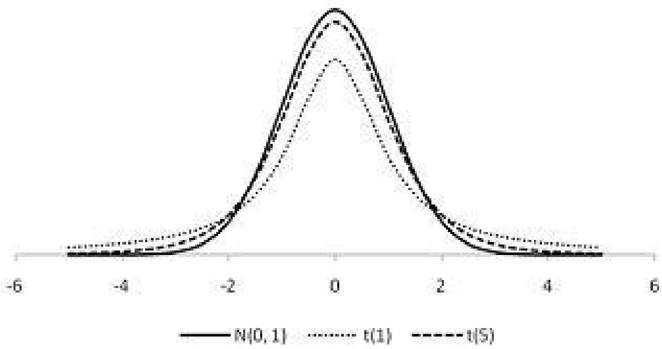
$$T = \frac{Z}{\sqrt{\frac{\chi_k^2}{k}}}$$

Então, sendo V uma distribuição qui-quadrado com v graus de liberdade, temos:

$$T = \frac{Z}{\sqrt{\frac{V}{v}}}$$

Resposta: Certo.

(CESPE/2011 – Analista Judiciário do TJ/ES)



O gráfico acima mostra a função de densidade da distribuição normal padrão N(0, 1) e t(1) e t(5), que representam, respectivamente, as densidades da distribuição t de Student com 1 e 5 graus de liberdade.

Com base nesse gráfico, julgue o próximo item.

A distribuição N(0, 1) possui variância unitária, a t(5) possui variância igual a 5/3, e a variância da distribuição t-Student com 1 grau de liberdade é indefinida.

Comentários:

Vamos analisar as três afirmações.

Em relação à primeira afirmação, de fato, a distribuição normal padrão N(0,1) possui variância igual a 1 (unitária).

Em relação à segunda afirmação, a distribuição de t-Student com k = 5 graus de liberdade é:

$$V(X) = \frac{k}{k-2} = \frac{5}{3}$$

Em relação à terceira afirmação, a distribuição de t-Student com k = 1 grau de liberdade **não** tem variância definida.

Gabarito: Certo.

(FGV/2010 – FIOCRUZ) Duas variáveis aleatórias independentes X e Y são tais que X tem distribuição Normal com média 0 e variância 4 e Y pode ser escrita como $Y = Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2$, em que os Z_i são independentes e identicamente distribuídos com distribuição normal padrão, i = 1, 2, 3, 4. Nesse caso, a seguinte variável tem distribuição t- Student

a) XY

b) $XY^{-0,5}$

c) $X^{-1}Y$

d) $4XY^{-0,5}$

e) 2XY

Comentários:

A variável t-Student é definida como:

$$T = \frac{Z}{\sqrt{\frac{\mathcal{X}_k^2}{k}}}$$

Sendo X uma variável normal com média $\mu = 0$ e variância V(X) = 4, para transformá-la na normal padrão precisamos dividir pelo desvio padrão $\sigma = \sqrt{V(X)} = 2$:

$$Z = \frac{X-\mu}{\sigma} = \frac{X-0}{2} = \frac{X}{2}$$

Sendo Y a soma de 4 variáveis normais padrão independentes elevadas ao quadrado, então Y é uma variável qui-quadrado $\mathcal{X}_k^2$ com k = 4 graus de liberdade.

Sendo as variáveis independentes, então a seguinte razão apresenta distribuição de t-Student:

$$T = \frac{Z}{\sqrt{\frac{\mathcal{X}_k^2}{k}}} = \frac{\frac{X}{2}}{\sqrt{\frac{Y}{4}}} = \frac{\frac{X}{2}}{\frac{\sqrt{Y}}{2}} = \frac{X}{2} \times \frac{2}{\sqrt{Y}} = \frac{X}{\sqrt{Y}}$$

Essa razão pode ser representada como:

$$T = \frac{X}{Y^{\frac{1}{2}}} = X \times Y^{-\frac{1}{2}} = X \times Y^{-0,5}$$

Gabarito: B

(CESPE/2020 – Analista Judiciário do TJ/PA) Se X e Y são variáveis aleatórias normais independentes, tais que X ~ N(0,1) e Y ~ N(0,1), a razão $\frac{X}{Y}$ segue uma distribuição

a) de Cauchy

b) de Pareto

c) de Weibull

d) t de Student com 2 graus de liberdade

e) normal padrão

Comentários:

A razão entre variáveis independentes com distribuição normal segue uma distribuição de Cauchy, que corresponde à razão entre duas variáveis normais.

Note que X e Y são variáveis normais padrão. Nesse caso, temos também uma distribuição t de Student, porém com 1 grau de liberdade, e não 2, como descrito na alternativa D.

Gabarito: A.

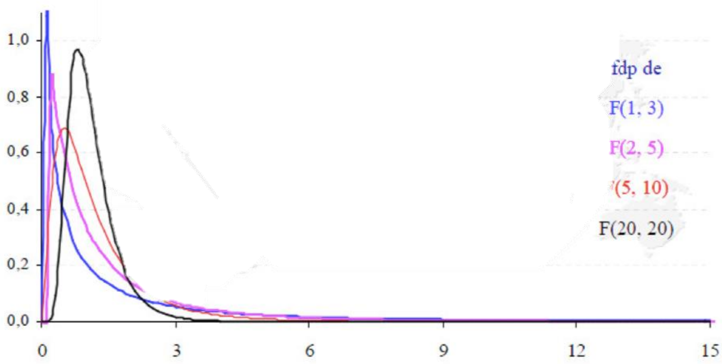
DISTRIBUIÇÃO F DE SNEDECOR

A distribuição F de Snedecor (ou de **Fisher**, ou de Fisher-Snedecor) é definida da seguinte forma, em que $\mathcal{X}_{k_1}^2$ e $\mathcal{X}_{k_2}^2$ são variáveis **independentes** com distribuição **qui-quadrado**, com k_1 e k_2 graus de liberdade, respectivamente:

$$F_{k_1,k_2} = \frac{\mathcal{X}_{k_1}^2 / k_1}{\mathcal{X}_{k_2}^2 / k_2}$$

A variável F_{k_1,k_2} apresenta $\boldsymbol{k_1}$ **graus de liberdade no numerador** e $\boldsymbol{k_2}$ **graus de liberdade no denominador**, os quais são os **parâmetros** da distribuição.

O gráfico abaixo[1] apresenta as f.d.p. de distribuições F de Snedecor, para diferentes parâmetros.



Observa-se que a variável assume apenas valores **positivos**, ou seja, no intervalo $(0, \infty)$ (afinal, resulta da divisão entre duas variáveis positivas). Além disso, observamos que a distribuição é **assimétrica à direita**, porém seu formato varia de acordo com os seus parâmetros.

A **média** dessa distribuição é dada por:

$$\mu = \frac{k_2}{k_2 - 2}, \ para\ k_2 > 2$$

Assim, para a distribuição F de Snedecor, apenas o **grau de liberdade do denominador** $\boldsymbol{k_2}$ influencia no cálculo da média.

Além disso, a média é $\boldsymbol{\mu > 1}$, sendo **reduzida** conforme k_2 aumenta. (Essa é a mesma fórmula da **variância da t-Student**!)

[1] Obtido na apresentação das aulas da Prof. Tarciana Liberal, da Universidade Federal da Paraíba, disponível em http://www.de.ufpb.br/~tarciana/Probabilidade2/Aula16.pdf

O **quadrado** de uma distribuição de **t-Student** com k graus de liberdade corresponde a uma distribuição **F de Snedecor** com 1 grau de liberdade no numerador e k graus de liberdade no denominador:

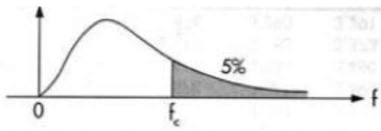
$$T^2 = \frac{Z^2}{\frac{\chi_k^2}{k}} = F_{1,k}$$

Para a distribuição F de Snedecor, também existe uma **tabela** para que possamos associar probabilidades aos intervalos de valores.

Porém, como a distribuição depende de **2 parâmetros**, sendo um associado às linhas e o outro, às colunas, a tabela é apresentada para valores **fixos** de **probabilidade**.

A tabela a seguir apresenta os valores de f que delimitam uma probabilidade $P(F > f) = 0{,}05$.

Tabela VI — Distribuição F
Corpo da tabela dá os valores f_c tais que $P(F > f_c) = 0{,}05$.



Graus de liberdade do denominador de F: ν_2	Grau de liberdade do numerador de F: ν_1																					
	1	2	3	4	5	6	7	8	9	10	12	14	15	16	18	20	24	30	40	60	120	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	245,4	245,9	246,5	247,3	248,0	249,1	250,1	251,1	252,2	253,3	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,42	19,43	19,43	19,44	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,72	8,70	8,69	8,67	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,87	5,86	5,84	5,82	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,64	4,62	4,60	4,58	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,96	3,94	3,92	3,90	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,53	3,51	3,49	3,47	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,24	3,22	3,20	3,17	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,03	3,01	2,99	2,96	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,87	2,85	2,83	2,80	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,74	2,72	2,70	2,67	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,64	2,62	2,60	2,57	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,55	2,53	2,52	2,48	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,48	2,46	2,44	2,41	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,42	2,40	2,39	2,35	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,37	2,35	2,33	2,30	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,34	2,31	2,29	2,26	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,29	2,27	2,25	2,22	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,26	2,23	2,22	2,18	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,22	2,20	2,18	2,15	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,20	2,18	2,16	2,12	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,17	2,15	2,13	2,10	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,15	2,13	2,11	2,08	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,13	2,11	2,09	2,05	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,11	2,09	2,07	2,04	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,09	2,07	2,05	2,02	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,08	2,06	2,04	2,00	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,06	2,04	2,02	1,99	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,05	2,03	2,01	1,97	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,04	2,01	1,99	1,96	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,95	1,92	1,90	1,87	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,86	1,84	1,81	1,78	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,77	1,75	1,72	1,69	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,69	1,67	1,63	1,60	1,57	1,52	1,46	1,39	1,32	1,22	1,00

Com essa tabela, podemos calcular o valor de f que delimita uma probabilidade $P(F > f) = 0{,}05$ para uma distribuição com 5 graus de liberdade no numerador (indicados nas linhas) e 10 graus de liberdade no denominador (indicado nas colunas), por exemplo.

Pela 5ª coluna e 10ª linha da tabela, podemos observar que $f = 3{,}33$.

Por outro lado, se precisarmos calcular o valor associado a uma outra probabilidade, $P(F > f) = 0{,}1$, por exemplo, precisaríamos de **outra tabela**.

Distribuição F-Snedecor

Razão entre variáveis qui-quadrado, divididas pelos respectivos graus de liberdade:

$$F_{k_1,k_2} = \frac{\chi^2_{k_1}/k_1}{\chi^2_{k_2}/k_2}$$

Assimétrica à direita, com $x \geq 0$

Média: $\mu = \frac{k_2}{k_2 - 2},\ para\ k_2 > 2$

(2017 – TRF 2ª Região – Adaptada) Sobre a distribuição F, julgue a afirmativa a seguir.

Se uma variável aleatória Y segue a distribuição F, então seus valores são não negativos e a forma da sua distribuição é controlada pelos graus de liberdade.

Comentários:

Vimos que a distribuição F assume apenas valores não negativos, e depende dos graus de liberdade no numerador e no denominador.

Resposta: Certo.

(FCC/2015 – Analista do CNMP) Sejam X_1, X_2, ... , X_n variáveis aleatórias independentes e identicamente distribuídas com distribuição normal padrão. Sejam as variáveis aleatórias:

$$Y = X_1 + X_2 + \cdots + X_n; \qquad W = X_1^2 + X_2^2 + \cdots + X_n^2; \qquad V = \frac{Y/2}{W/n}$$

Com base nessas informações, julgue a afirmativa a seguir:

A variável V tem distribuição F de Snedecor com graus de liberdade 2 e n.

Comentários:

A distribuição F de Snedecor é dada por:

$$F_{k_1,k_2} = \frac{\mathcal{X}_{k_1}^2/k_1}{\mathcal{X}_{k_2}^2/k_2}$$

O enunciado informa que:

$$V = \frac{Y/2}{W/n}$$

Em que Y é a soma de variáveis normais, ou seja, Y segue distribuição normal, em vez de qui-quadrado, então V **não** apresenta distribuição F de Snedecor.

Resposta: Errado.

(CESPE/2010 – Estatístico do MS) A média de uma distribuição F de Snedecor depende de dois parâmetros: o número de graus de liberdade do denominador e o número de graus de liberdade do numerador.

Comentários:

A média da distribuição F de Snedecor é dada por:

$$\mu = \frac{k_2}{k_2 - 2}, \qquad para\ k_2 > 2$$

Logo, a média depende apenas do número de graus de liberdade do **denominador**.

Gabarito: Errado.

RESUMO DA AULA

Distribuição Uniforme: $f(x) = \frac{1}{(b-a)},\ a \leq x \leq b$:

⇒ Probabilidade de um **intervalo**: $P(m < X < n) = \frac{(n-m)}{(b-a)}$

⇒ **Esperança:** $E(X) = \frac{b+a}{2}$ e **Variância**: $V(X) = \frac{(b-a)^2}{12}$

Distribuição Exponencial, $f(x) = \lambda. e^{-\lambda x}$:

⇒ Probabilidade de um **intervalo**: $P(a < X < b) = e^{-\lambda.a} - e^{-\lambda b}$

⇒ Função de **Distribuição Acumulada**: $P(X < b) = 1 - e^{-\lambda b}$

⇒ **Esperança:** $E(X) = \frac{1}{\lambda}$ e **Variância**: $V(X) = \frac{1}{\lambda^2}$

⇒ **Propriedade**: $P(X > t + s|X > s) = P(X > t)$

Distribuição Normal $N(\mu, \sigma^2)$: com média μ e variância σ^2

⇒ **Simétrica** com formato de **sino,** definida em toda a reta real

⇒ Transformação para a normal **padrão** $N(0,1)$: $z = \frac{x-\mu}{\sigma}$

⇒ **Teorema Central do Limite:** Soma de muitas variáveis tende a seguir distribuição normal

Distribuição Qui-quadrado: Positiva; **assimétrica**

$$\mathcal{X}_k^2 = \sum_{i=1}^{k} (Z_i)^2$$

⇒ **Esperança:** $E(X) = k$ e **Variância**: $V(X) = 2k$

Distribuição t-Student: Simétrica com $\mu = 0$ e formato de sino mais "**achatado**" do que a normal

$$T = \frac{Z}{\sqrt{\frac{\mathcal{X}_k^2}{k}}}$$

⇒ **Variância**: $V(X) = \frac{k}{k-2},\ para\ k > 2$

Distribuição Fisher-Snedecor: Positiva; **assimétrica**

$$F_{k_1,k_2} = \frac{\mathcal{X}_{k_1}^2 / k_1}{\mathcal{X}_{k_2}^2 / k_2}$$

⇒ **Esperança**: $E(X) = \frac{k_2}{k_2-2},\ para\ k_2 > 2$

QUESTÕES COMENTADAS – CESGRANRIO

Exponencial

1. (CESGRANRIO/2014 – EPE) Um equipamento eletrônico tem vida útil média de 10 anos. Supondo que a vida útil do equipamento segue o comportamento de uma variável aleatória com distribuição exponencial, qual é a probabilidade de esse equipamento ter vida útil acima de 12 anos?

a) exp(-120)

b) exp(-12)

c) exp(-1,2)

d) exp(-0,12)

e) exp(-0,012)

Comentários:

O enunciado informa que a variável segue distribuição exponencial com média E(X) = 10 anos. Então, o parâmetro da distribuição é:

$$E(X) = \frac{1}{\lambda} = 10$$

$$\lambda = \frac{1}{10} = 0{,}1$$

Logo, a probabilidade de a vida útil ser maior que x = 12 anos é:

$$P(X > x) = e^{-\lambda.x}$$

$$P(X > 12) = e^{-0{,}1\times 12} = e^{-1{,}2} = \exp\{-1{,}2\}$$

Gabarito: C

2. (CESGRANRIO/2010 – IBGE) O intervalo de tempo entre a chegada de dois navios a um porto, em horas, segue distribuição exponencial com média 1. Se acaba de chegar um navio, qual a probabilidade aproximada de que leve mais de uma hora até a chegada do próximo?

a) 0,37

b) 0,5

c) 0,63

d) 0,75

e) 0,9

Comentários:

O enunciado informa que a variável segue distribuição exponencial com média E(X) = 1 hora. Então, o parâmetro da distribuição é:

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

$$\lambda = 1$$

Logo, a probabilidade de o próximo navio levar mais de x = 1 hora é:

$$P(X > x) = e^{-\lambda.x}$$

$$P(X > 1) = e^{-1\times 1} = e^{-1}$$

Pela propriedade da função exponencial, quando o expoente é um número negativo, invertemos a base:

$$P(X > 1) = e^{-1} = \frac{1}{e^1} = \frac{1}{e}$$

Sabendo que $e \cong 2{,}718$, temos:

$$P(X > 1) = \frac{1}{2{,}718} \cong 0{,}368 \cong 37\%$$

Gabarito: A

3. (CESGRANRIO/2014 – EPE) Um componente tem a vida útil (em horas) regida pela distribuição exponencial com média θ horas. Qual a probabilidade de um dado componente atender à demanda de θ horas?

a) 1/2

b) $e^{-\theta^2}$

c) $e^{-\theta}$

d) $e^{-1/2}$

e) e^{-1}

Comentários:

O enunciado informa que a variável segue distribuição exponencial com média E(X) = θ horas. Então, o parâmetro da distribuição é:

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

$$\lambda = \frac{1}{\theta}$$

Logo, a probabilidade de o componente durar mais de x = θ horas é:

$$P(X > x) = e^{-\lambda . x}$$

$$P(X > \theta) = e^{-\frac{1}{\theta}\times\theta} = e^{-1}$$

Gabarito: E

4. (CESGRANRIO/2010 – Petrobrás) As ocorrências diárias de situações de emergência em uma instalação industrial são aleatórias e usualmente consideradas independentes umas das outras. Dessa forma, o modelo mais adequado para a simulação dos instantes de ocorrências é a Distribuição de Poisson e, consequentemente, os intervalos entre as ocorrências obedecem à Distribuição Exponencial. Na prática, observa-se que o tempo dedicado por um engenheiro à solução de cada emergência é bem modelado também pela Distribuição Exponencial. Esses são alguns dos motivos para que, em simulação desses processos de atendimento, o tempo (T) entre ocorrências e o tempo (T) de tratamento das mesmas sejam modelados por Distribuições Exponenciais que, entre outros aspectos, têm a propriedade denominada "ausência de memória" que (para quaisquer t > 0 e a > 0) é traduzida por:

a) P(T > t + a | T > a) = P(T > t)

b) Valor Esperado de T = variância de T ($\mu = \sigma^2$)

c) [Valor Esperado de T]^2 = variância de T $\mu^2 = \sigma^2$)

d) P(0 < T < a) > P(t < T < t + a)

e) P(0 < T < a) = P(t < T < t + a)

Comentários:

Essa questão trabalha com a propriedade sem memória da distribuição exponencial, que nos informa que a probabilidade de a variável superar t + a, no total, dado que superou a, é igual à probabilidade de a variável superar t:

$$P(T > t + a|T > a) = P(T > t)$$

Gabarito: A

5. (CESGRANRIO/2013 – BNDES) O tempo de ligações telefônicas segue uma distribuição de probabilidade exponencial com média de 3 minutos. Um sujeito chega a um telefone público e descobre que a pessoa à sua frente está na ligação há pelo menos dois minutos.

Qual é a probabilidade de essa ligação durar pelo menos cinco minutos no total?

a) e^{-1}

b) e^{-2}

c) e^{-3}

d) $1 - e^{-3}$

e) $1 - e^{-5}$

Comentários:

Essa questão trabalha com a propriedade sem memória da distribuição exponencial:

$$P(X > t + s|X > s) = P(X > t)$$

Essa propriedade nos afirma que a probabilidade de a ligação durar mais de t + s = 5 minutos, dado que durou mais de s = 2 minutos é igual à probabilidade de a ligação durar mais de t = 3 minutos.

Para isso, o enunciado informa que a média da distribuição é $E(X) = 3$. Logo, o parâmetro da distribuição é:

$$E(X) = \frac{1}{\lambda} = 3$$

$$\lambda = \frac{1}{3}$$

E a probabilidade de a ligação durar mais de t = 3 minutos é dada por:

$$P(X > t) = e^{-\lambda . t}$$

$$P(X > 3) = e^{-\frac{1}{3}\times 3} = e^{-1}$$

Que é igual à probabilidade de a ligação durar pelo menos 5 minutos no total, sabendo que já durou pelo menos 2 minutos.

Gabarito: A

6. (CESGRANRIO/2014 – EPE) Uma companhia possui dois geradores elétricos. O tempo até a falha de cada gerador se comporta segundo uma distribuição exponencial, com média de 10 anos. A companhia passa a usar o segundo gerador tão logo o primeiro em funcionamento falhe.

Qual é a variância do tempo total em que os dois geradores produziram energia?

a) 200

b) 150

c) 100

d) 20

e) 10

Comentários:

O enunciado informa que a variável segue distribuição exponencial com média E(X) = 10 anos. Então, o parâmetro da distribuição é:

$$E(X) = \frac{1}{\lambda} = 10$$

$$\lambda = \frac{1}{10} = 0{,}1$$

Portanto, a variância da distribuição é:

$$V(X) = \frac{1}{\lambda^2} = \frac{1}{0{,}1^2} = \frac{1}{0{,}01} = 100$$

Considerando que os geradores são independentes, a variância da soma é a soma das variâncias:

$$V(X_1 + X_2) = V(X_1) + V(X_2) = 100 + 100 = 200$$

Gabarito: A

QUESTÕES COMENTADAS – CESGRANRIO

Distribuição Normal

1. (CESGRANRIO/2010 – Petrobras – *Adaptada*) A distribuição normal é uma das mais utilizadas para inferência estatística da probabilidade de ocorrência de diversos fenômenos em engenharia. Nela, qualquer variável aleatória normal X é convertida em uma variável normal padronizada Z, tal que Z = $\frac{X-\mu}{\hat{\sigma}}$, onde $\hat{\sigma}$ é o desvio padrão e μ é a média aritmética. Na distribuição normal padronizada,

a) a variável Z é contínua e representa o número de desvios em relação à média.

b) a variável Z possui média não nula e desvio padrão igual a 1.

c) um ponto da curva da distribuição indica a probabilidade da variável aleatória assumir um valor real entre 0 e 1.

d) valores idênticos acima e abaixo da média têm probabilidades complementares, pois a curva é simétrica.

Comentários:

A questão trata da distribuição normal padrão Z ~ N(0,1).

Em relação à alternativa A, a variável Z é contínua e representa os valores de $\frac{X-\mu}{\hat{\sigma}}$, ou seja, dos desvios em relação à média μ (relativizados pelo desvio padrão). A alternativa A está correta.

Em relação à alternativa B, a variável Z possui média **nula** e desvio padrão igual a 1. A alternativa B está incorreta pois diz que a média é **não nula**.

Em relação à alternativa C, um ponto qualquer da distribuição normal padrão indica a probabilidade de a variável assumir um valor na reta real $(-\infty, \infty)$, e não entre 0 e 1, logo, a alternativa C está incorreta.

Em relação à alternativa D, a curva normal padrão é simétrica em torno de $\mu = 0$, logo valores idênticos acima e abaixo da média possuem as **mesmas** probabilidades. Por exemplo, P(Z < -1) = P(Z > 1). A alternativa D está incorreta.

Gabarito: A

2. (CESGRANRIO/2006 – EPE) Se uma distribuição segue um padrão normal, é correto afirmar que:

a) 98% dos números estão a dois desvios padrão da média.

b) 95% dos números estão a 1,5 desvio padrão da média.

c) 95% dos números estão a um desvio padrão da média.

d) 86% dos números estão a um desvio padrão da média.

e) 68% dos números estão a um desvio padrão da média.

Comentários:

A questão trabalha com a regra empírica da curva normal. Segundo essa regra, aproximadamente 68% dos números estão a 1 desvio padrão da média; 95% dos números estão a 2 desvios padrão da média; e 99,7% dos números estão a 3 desvios padrão da média.

Gabarito: E

3. (CESGRANRIO/2012 – Petrobras) Considere que, no verão, o valor das vendas de determinada loja, nas sextas-feiras que sejam dias úteis, tem uma distribuição de probabilidades normal, com média igual a R$ 10.000,00 e desvio padrão de R$ 2.000,00.

Com essa hipótese, a probabilidade de que, em uma dessas sextas-feiras úteis do verão, o valor das vendas seja inferior a R$ 8.000,00 é

a) menor que 20%

b) igual a 20%

c) maior que 20% e menor que 25%

d) igual a 25%

e) maior que 25%

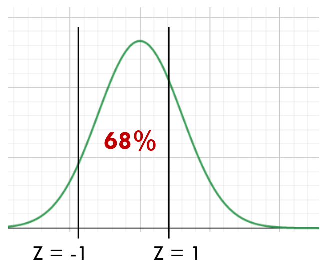
Comentários:

O enunciado informa que a variável segue distribuição normal com média $\mu = 10.000$ reais e desvio padrão $\sigma = 2000$ reais. Então, para calcular a probabilidade de o valor das vendas ser inferior a x = 8000 reais, usamos a seguinte transformação para a curva normal padrão:

$$z = \frac{x - \mu}{\sigma} = \frac{8000 - 10.000}{2000} = \frac{-2000}{2000} = -1$$

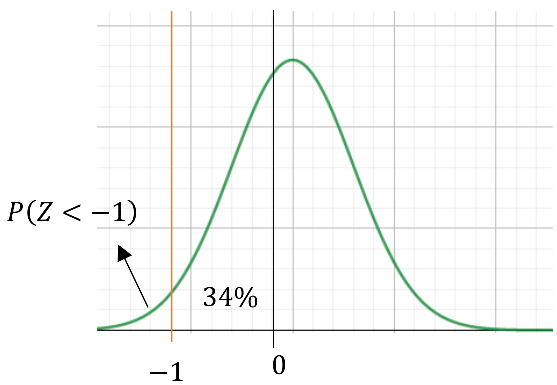
Pela regra empírica, sabemos que:

$$P(-1 < Z < 1) \cong 68\%$$



Pela simetria da normal padrão, a probabilidade P(-1 < Z < 0) é a metade desse valor:

$$P(-1 < Z < 0) \cong \frac{68\%}{2} = 34\%$$



Sabendo que a área à esquerda da média tem probabilidade P(Z < 0) = 50%, então a probabilidade P(Z < -1) é:

$$P(Z < -1) \cong 50\% - 34\% = 16\%$$

Que é menor que 20%.

Gabarito: A

4. (CESGRANRIO/2010 – BACEN) Estima-se que os retornos de um determinado mercado tenham distribuição normal, com média 20% e desvio padrão 10%. A probabilidade de perdas financeiras é de, aproximadamente,

a) 1%

b) 2,5%

c) 5%

d) 10%

e) 20%

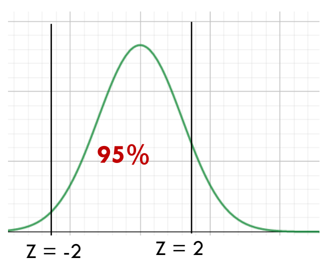
Comentários:

O enunciado informa que a variável segue distribuição normal com média $\mu = 20\% = 0{,}2$ e desvio padrão $\sigma = 10\% = 0{,}1$. Então, para calcular a probabilidade de haver perdas financeiras (retorno menor que x = 0%), usamos a seguinte transformação para a curva normal padrão:

$$z = \frac{x - \mu}{\sigma} = \frac{0 - 0{,}2}{0{,}1} = \frac{-0{,}2}{0{,}1} = -2$$

Pela regra empírica, sabemos que:

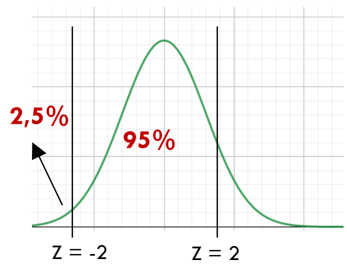
$$P(-2 < Z < 2) \cong 95\%$$



Então, a probabilidade associada aos valores externos ao intervalo (-2; 2) é:

100% - 95% = 5%

Pela simetria da normal padrão, a probabilidade P(Z < -2) é a metade desse valor:



$$P(Z < -2) \cong \frac{5\%}{2} = 2{,}5\%$$

Que é a probabilidade de perdas financeiras.

Gabarito: B

5. (CESGRANRIO/2012 – Innova) Suponha que o tempo de vida de baterias de celular tenha distribuição normal com média de 120 minutos e variância de 100 minutos. Qual é a probabilidade aproximada de uma bateria durar menos que 100 minutos?

a) 0,15%

b) 2,5%

c) 5%

d) 10%

e) 16%

Comentários:

O enunciado informa que a variável segue distribuição normal com média $\mu = 120$ e variância $\sigma^2 = 100$. Logo, o desvio padrão, raiz quadrada da variância é:

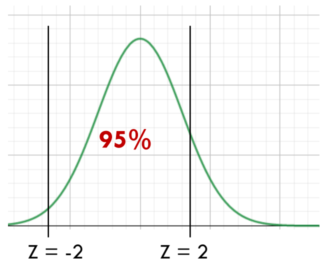
$$\sigma = \sqrt{\sigma^2} = \sqrt{100} = 10$$

Então, para calcular a probabilidade de a bateria durar menos de x = 100 minutos, usamos a seguinte transformação para a curva normal padrão:

$$z = \frac{x - \mu}{\sigma} = \frac{100 - 120}{10} = \frac{-20}{10} = -2$$

Pela regra empírica, sabemos que:

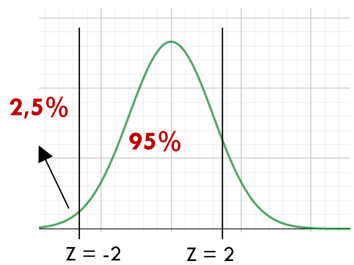
$$P(-2 < Z < 2) \cong 95\%$$



Então, a probabilidade associada aos valores externos ao intervalo (-2; 2) é:

100% - 95% = 5%

Pela simetria da normal padrão, a probabilidade P(Z < -2) é a metade desse valor:



$$P(Z < -2) \cong \frac{5\%}{2} = 2{,}5\%$$

Que é a probabilidade de a bateria durar menos de 100 minutos.

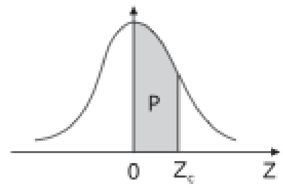
Gabarito: B

Para as próximas questões, utilize a tabela normal a seguir, constante nas respectivas provas da CESGRANRIO.

Distribuição Normal Padrão

$Z \sim N(0,1)$

Corpo da tabela dá a probabilidade p, tal que $p = P(0<Z<Z_c)$



parte inteira e primeira decimal de Z_c	Segunda decimal de Z_c 0	1	2	3	4	5	6	7	8	9	parte inteira e primeira decimal de Z_c
	p = 0										
0,0	00000	00399	00798	01197	01595	01994	02392	02790	03188	03586	0,0
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535	0,1
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409	0,2
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173	0,3
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793	0,4
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240	0,5
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490	0,6
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524	0,7
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327	0,8
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891	0,9
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214	1,0
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298	1,1
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147	1,2
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774	1,3
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189	1,4
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408	1,5
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449	1,6
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327	1,7
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062	1,8
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670	1,9
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169	2,0
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574	2,1
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899	2,2
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158	2,3
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361	2,4
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520	2,5
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643	2,6
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736	2,7
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807	2,8
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861	2,9
3,0	49865	49869	49874	49878	49882	49886	49889	49893	49897	49900	3,0
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929	3,1
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950	3,2
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965	3,3
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976	3,4
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983	3,5
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989	3,6
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49992	3,7
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995	3,8
3,9	49995	49995	49996	49996	49996	49996	49996	49996	49997	49997	3,9
4,0	49997	49997	49997	49997	49997	49997	49998	49998	49998	49998	4,0
4,5	49999	50000	50000	50000	50000	50000	50000	50000	50000	50000	4,5

6. (CESGRANRIO/2010 – IBGE) Seja H a variável aleatória que representa as alturas dos cidadãos de certo país. Sabe-se que H tem distribuição normal com média 1,70 m e desvio padrão 0,04 m. A probabilidade de que um cidadão desse país tenha mais do que 1,75 m de altura é, aproximadamente,

a) 9,9%

b) 10,6%

c) 22,2%

d) 39,4%

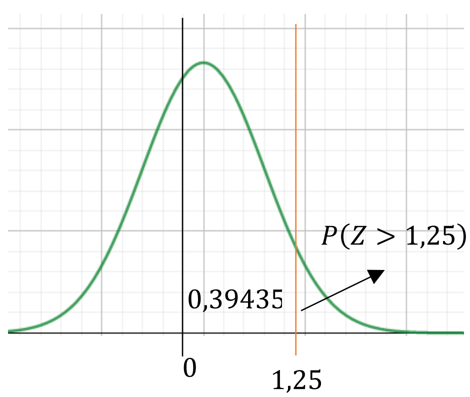
e) 40,6%

Comentários:

O enunciado informa que a variável segue distribuição normal com média $\mu = 1{,}70$ e desvio padrão $\sigma = 0{,}04$. Então, para calcular a probabilidade de um cidadão ter mais de x = 1,75 m de altura, usamos a seguinte transformação para a curva normal padrão:

$$z = \frac{x - \mu}{\sigma} = \frac{1{,}75 - 1{,}7}{0{,}04} = \frac{0{,}05}{0{,}04} = 1{,}25$$

Pela tabela normal, observamos que P(0 < Z < 1,25) = 0,39435, como ilustrado a seguir.



Sabendo que a área à direita da média tem probabilidade P(Z > 0) = 0,5, então a probabilidade P(Z > 1,25) é:

$$P(Z > 1{,}25) = 0{,}5 - 0{,}39435 = 0{,}10565 \cong 10{,}6\%$$

Gabarito: B

7. (CESGRANRIO/2006 – EPE) Suponha os pesos dos pacotes de arroz normalmente distribuídos com média 1kg e desvio padrão 20g. Escolhendo um pacote ao acaso, qual é a probabilidade de ele pesar mais de 1 030g?

a) 13,4%

b) 11,6%

c) 10,0%

d) 8,4%

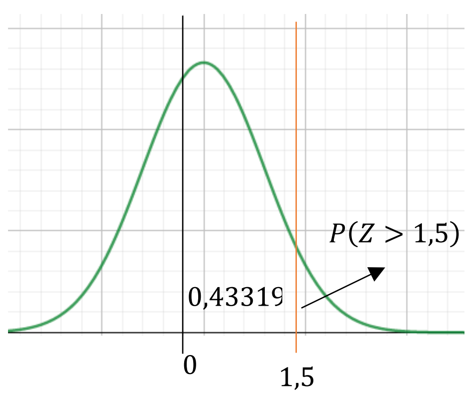
e) 6,7%

Comentários:

O enunciado informa que a variável segue distribuição normal com média $\mu = 1kg = 1000g$ e desvio padrão $\sigma = 20g$. Então, para calcular a probabilidade de um pacote de arroz pesar mais de x = 1030g, usamos a seguinte transformação para a curva normal padrão:

$$z = \frac{x - \mu}{\sigma} = \frac{1030 - 1000}{20} = \frac{30}{20} = 1,5$$

Pela tabela normal, observamos que P(0 < Z < 1,5) = 0,43319, como ilustrado a seguir.



Sabendo que a área à direita da média tem probabilidade P(Z > 0) = 0,5, então a probabilidade P(Z > 1,5) é:

$$P(Z > 1,5) = 0,5 - 0,43319 = 0,06681 \cong 6,7\%$$

Gabarito: E

8. (CESGRANRIO/2012 – Petrobras) O peso de lotes produzidos por uma certa indústria segue uma distribuição normal, com média de 10 kg e desvio padrão de 0,2 kg. Em um lote dessa indústria, selecionado aleatoriamente, qual a probabilidade de o peso do lote não se afastar por mais de 1% do peso médio?

a) 50%

b) 38,30%

c) 17,42%

d) 7,96%

e) 0%

Comentários:

O enunciado pede a probabilidade de o peso de um lote selecionado aleatoriamente não se afaste da média em mais de 1% = 0,01. Sabendo que a média é de 10kg, a distância máxima desejada é:

$$10 \times 0{,}01 = 0{,}1$$

Então, buscamos a probabilidade de o lote estar no intervalo:

$$P(9{,}9 < X < 10{,}1)$$

Sabendo que a média é $\mu = 10$ e desvio padrão $\sigma = 0{,}2$, calculamos a probabilidade de o lote ser superior x = 10,1, pela seguinte transformação para a curva normal padrão:

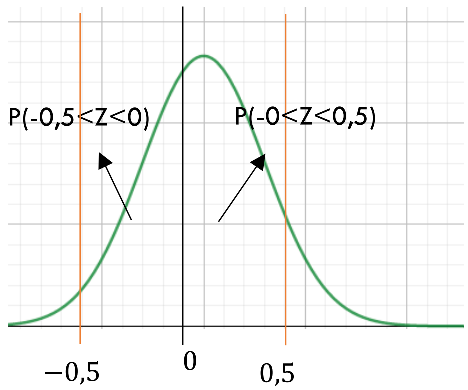
$$z = \frac{x - \mu}{\sigma} = \frac{10{,}1 - 10}{0{,}2} = \frac{0{,}1}{0{,}2} = 0{,}5$$

E a transformação para x = 9,9 é:

$$z = \frac{x - \mu}{\sigma} = \frac{9{,}9 - 10}{0{,}2} = \frac{-0{,}1}{0{,}2} = -0{,}5$$

Logo, a probabilidade buscada é igual a:

$$P(9{,}9 < X < 10{,}1) = P(-0{,}5 < Z < 0{,}5) = P(-0{,}5 < Z < 0) + P(0 < Z < 0{,}5)$$



Pela tabela normal, observamos que P(0 < Z < 0,5) = 0,19146. Pela simetria, temos:

$$P(0 < Z < 0,5) = P(-0,5 < Z < 0) = 0,19146$$

Logo:

$$P(-0,5 < Z < 0,5) = 0,19146 + 0,19146 = 0,38292 \cong 38,3\%$$

Gabarito: B

9. (CESGRANRIO/2016 – IBGE) O tempo de atendimento (em minutos) para chamadas de emergência do SAMU no Rio de Janeiro segue uma distribuição normal com média 12 e variância 25.

Qual a probabilidade de que o tempo de atendimento para uma dada chamada exceda a 20 minutos?

a) 1,79%

b) 2,87%

c) 5,48%

d) 25,50%

e) 37,45%

Comentários:

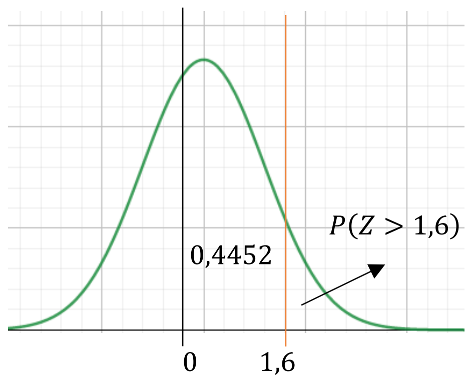
O enunciado informa que a variável segue distribuição normal com média $\mu = 12$ e variância $\sigma^2 = 25$. O desvio padrão é, então:

$$\sigma = \sqrt{\sigma^2} = \sqrt{25} = 5$$

Portanto, para calcular a probabilidade de uma chamada exceder x = 20 minutos, usamos a seguinte transformação para a curva normal padrão:

$$z = \frac{x - \mu}{\sigma} = \frac{20 - 12}{5} = \frac{8}{5} = 1{,}6$$

Pela tabela normal, observamos que P(0 < Z < 1,6) = 0,4452, como ilustrado a seguir.



Sabendo que a área à direita da média tem probabilidade P(Z > 0) = 0,5, então a probabilidade P(Z > 1,6) é:

$$P(Z > 1{,}6) = 0{,}5 - 0{,}4452 = 0{,}0548 = 5{,}48\%$$

Gabarito: C

10. (CESGRANRIO/2012 – Petrobras) Os salários de técnicos de uma empresa se distribuem normalmente com média de R$ 3.200,00 e desvio padrão de R$ 800,00.

Selecionando-se aleatoriamente dois salários de técnicos dessa empresa, qual a probabilidade de pelo menos um deles ser superior a R$ 3.880,00?

a) 72,25%

b) 15,86%

c) 3,91%

d) 19,77%

e) 35,63%

Comentários:

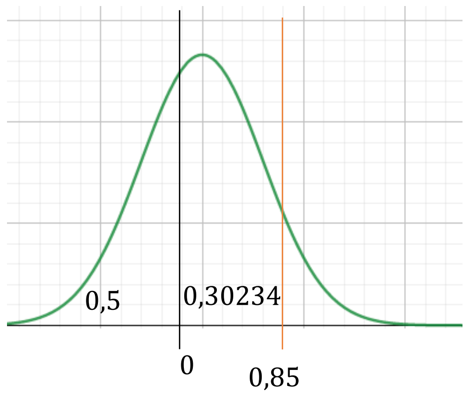
A probabilidade de pelo menos um, dentre os dois salários selecionados aleatoriamente, ser superior a R$ 3.880, pode ser calculada pelo complementar da probabilidade de ambos os salários serem inferiores a esse valor:

$$P(pelo\ menos\ um\ superior)\ =\ 1 - P(ambos\ inferiores)$$

Para calcular a probabilidade de ambos os salários serem inferiores a tal valor, precisamos primeiro calcular a probabilidade de um salário ser inferior. Para isso, o enunciado informa que os salários seguem distribuição normal com média $\mu = 3.200$ e desvio padrão $\sigma = 800$. Portanto, para calcular a probabilidade de um salário ser inferior a x = 3.880 minutos, usamos a seguinte transformação para a curva normal padrão:

$$z = \frac{x-\mu}{\sigma} = \frac{3880 - 3200}{800} = \frac{680}{800} = 0{,}85$$

Pela tabela normal, observamos que P(0 < Z < 0,85) = 0,30234, como ilustrado a seguir.



Sabendo que a área à esquerda da média tem probabilidade P(Z < 0) = 0,5, então a probabilidade P(Z < 0,85) é:

$$P(Z < 0{,}85) = 0{,}5 + 0{,}30234 = 0{,}80234$$

Então, a probabilidade de os dois salários selecionados aleatoriamente serem inferiores ao valor (interseção dos eventos) é o produto das probabilidades:

$$P(ambos\ inferiores) = 0{,}80234 \times 0{,}80234 \cong 0{,}6437$$

E a probabilidade de ter pelo menos um salário superior ao citado valor é o complementar:

$$P(pelo\ menos\ um\ superior) \cong 1 - 0{,}6437 = 0{,}3563 = 35{,}63\%$$

Gabarito: E

11. (CESGRANRIO/2007 – TCE/RO) O gasto médio dos clientes de um posto de gasolina é uma variável aleatória normal com média R$ 100,00 e desvio padrão R$ 25,00. Os 10% dos que mais consomem recebem um tratamento VIP, incluindo lavagem de carroceria, calibragem nos pneus e verificação do óleo e da água. Quanto você precisa gastar nesse posto de gasolina, em reais, para obter tratamento VIP?

a) 158,00

b) 149,00

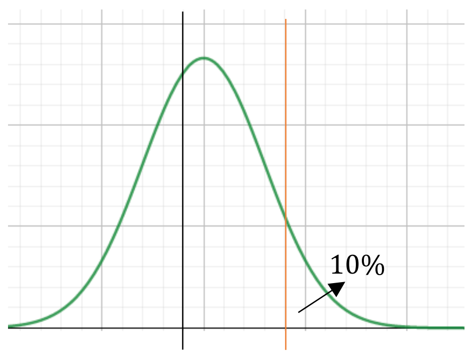
c) 141,00

d) 132,00

e) 128,00

Comentários:

O enunciado informa que o gasto dos clientes segue distribuição normal e que os 10% que mais consomem recebem tratamento VIP, como ilustrado a seguir:



Ou seja, para sabermos o valor do gasto necessário para obter tratamento VIP, primeiro precisamos do valor de z tal que P(Z > z) = 0,1. Sabendo que a área à direita da média tem probabilidade P(Z > 0) = 0,5, então precisamos do valor de z tal que:

$$P(0 < Z < z) = 0{,}5 - 0{,}1 = 0{,}4$$

Pela tabela, observamos que z = 1,28, pois P(Z < 1,28) = 0,39973 ≅ 0,4.

Considerando que a média dos gastos é $\mu = 100$ e que o desvio padrão é $\sigma = 25$, utilizamos a seguinte transformação para a normal padrão:

$$z = \frac{x - \mu}{\sigma}$$

$$1{,}28 = \frac{x - 100}{25}$$

$$x - 100 = 32$$

$$x = 132$$

Gabarito: D

12. (CESGRANRIO/2010 – IBGE) Suponha que as notas dos candidatos de um concurso público, em uma certa prova, sigam distribuição normal com média 7 e desvio padrão 1. A relação candidato/vaga é de 40 para 1. A nota mínima necessária para aprovação nessa prova é

a) 8,65

b) 8,96

c) 9,37

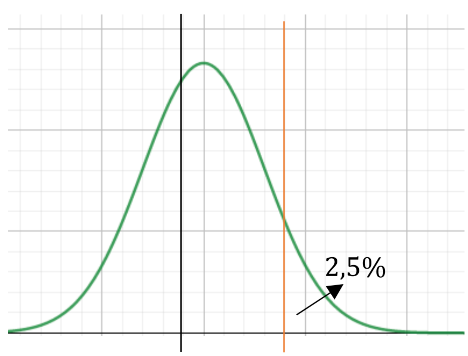
d) 9,58

e) 9,75

Comentários:

O enunciado informa que as notas em um concurso seguem distribuição normal e que a relação candidato-vaga é de 40 para 1. Então, para que o candidato passe nesse concurso, ele deverá tirar a melhor nota dentre 40 candidatos, ou seja, encaixar-se na seguinte proporção de melhores notas:

$$\frac{1}{40} = 2{,}5\% = 0{,}025$$



Ou seja, precisamos do valor de z tal que P(Z > z) = 0,025. Sabendo que a área à direita da média tem probabilidade P(Z > 0) = 0,5, então precisamos do valor de z tal que:

$$P(0 < Z < z) = 0{,}5 - 0{,}025 = 0{,}475$$

Pela tabela, observamos que z = 1,96, pois P(Z < 1,96) = 0,475.

Considerando que a média das notas é $\mu = 7$ e que o desvio padrão é $\sigma = 1$, utilizamos a seguinte transformação para a normal padrão:

$$z = \frac{x - \mu}{\sigma}$$

$$1{,}96 = \frac{x - 7}{1}$$

$$x - 7 = 1{,}96$$

$$x = 8{,}96$$

Gabarito: B

QUESTÕES COMENTADAS – CESGRANRIO

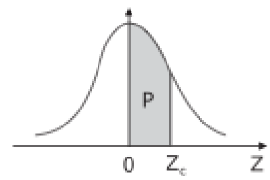
Soma de Variáveis e o Teorema Central do Limite

Para a próxima questão, utilize a tabela normal a seguir, constante na respectiva prova da CESGRANRIO.

Distribuição Normal Padrão

$Z \sim N(0,1)$

Corpo da tabela dá a probabilidade p, tal que $p = P(0<Z<Z_c)$



parte inteira e primeira decimal de Z_c	Segunda decimal de Z_c 0	1	2	3	4	5	6	7	8	9	parte inteira e primeira decimal de Z_c
	p = 0										
0,0	00000	00399	00798	01197	01595	01994	02392	02790	03188	03586	0,0
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535	0,1
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409	0,2
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173	0,3
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793	0,4
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240	0,5
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490	0,6
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524	0,7
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327	0,8
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891	0,9
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214	1,0
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298	1,1
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147	1,2
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774	1,3
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189	1,4
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408	1,5
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449	1,6
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327	1,7
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062	1,8
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670	1,9
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169	2,0
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574	2,1
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899	2,2
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158	2,3
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361	2,4
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520	2,5
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643	2,6
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736	2,7
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807	2,8
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861	2,9
3,0	49865	49869	49874	49878	49882	49886	49889	49893	49897	49900	3,0
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929	3,1
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950	3,2
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965	3,3
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976	3,4
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983	3,5
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989	3,6
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49992	3,7
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995	3,8
3,9	49995	49995	49996	49996	49996	49996	49996	49996	49997	49997	3,9
4,0	49997	49997	49997	49997	49997	49997	49998	49998	49998	49998	4,0
4,5	49999	50000	50000	50000	50000	50000	50000	50000	50000	50000	4,5

1. (CESGRANRIO/2011 – BNDES) Um produto tem massa normalmente distribuída com média 60 gramas e desvio padrão 8 gramas e é agrupado por dúzias. A probabilidade de a massa de uma dúzia ser superior a 750 gramas é, aproximadamente, de

a) 86%

b) 50%

c) 38%

d) 32%

e) 14%

Comentários:

O enunciado informa que a massa de um produto segue distribuição normal, com média $\mu_0 = 60$ e desvio padrão $\sigma_0 = 8$. Sabendo que o produto é agrupado por dúzias (12), então a média (esperança) do grupo é:

$$E(Dúzia) = 12 \times E(Produto)$$

$$\mu = 12 \times 60 = 720$$

Para calcular o desvio padrão da dúzia, primeiro calculamos a variância. Sabendo que a variância é o quadrado do desvio padrão, então a variância é 8^2 e a variância da dúzia é:

$$V(Dúzia) = 12 \times V(Produto) = 12 \times 8^2$$

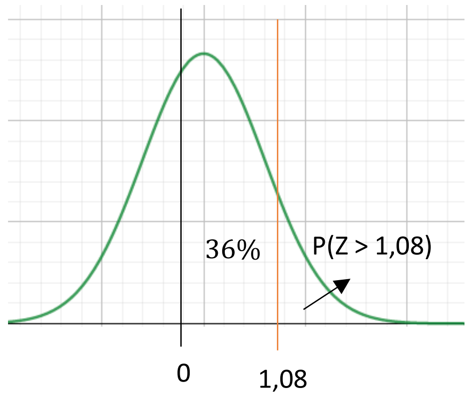
E o desvio padrão da dúzia é a raiz quadrada:

$$\sigma = \sqrt{12 \times 8^2} = 8 \times \sqrt{12} \cong 27{,}7$$

Conhecendo a média e o desvio padrão da dúzia, utilizamos a seguinte transformação para x = 750:

$$z = \frac{x - \mu}{\sigma} \cong \frac{750 - 720}{27{,}7} = \frac{30}{27{,}7} \cong 1{,}08$$

Pela tabela da distribuição normal padrão, observamos que P(0 < Z < 1,08) = 0,36 = 36%, como ilustrado a seguir:



Sabendo que a área à direita da média tem probabilidade P(Z > 0) = 0,5 = 50%, então a probabilidade desejada é:

$$P(Z > 1{,}08) = 50\% - 36\% = 14\%$$

Gabarito: E

LISTA DE QUESTÕES – CESGRANRIO

Exponencial

1. (CESGRANRIO/2014 – EPE) Um equipamento eletrônico tem vida útil média de 10 anos. Supondo que a vida útil do equipamento segue o comportamento de uma variável aleatória com distribuição exponencial, qual é a probabilidade de esse equipamento ter vida útil acima de 12 anos?

a) exp(-120)

b) exp(-12)

c) exp(-1,2)

d) exp(-0,12)

e) exp(-0,012)

2. (CESGRANRIO/2010 – IBGE) O intervalo de tempo entre a chegada de dois navios a um porto, em horas, segue distribuição exponencial com média 1. Se acaba de chegar um navio, qual a probabilidade aproximada de que leve mais de uma hora até a chegada do próximo?

a) 0,37

b) 0,5

c) 0,63

d) 0,75

e) 0,9

3. (CESGRANRIO/2014 – EPE) Um componente tem a vida útil (em horas) regida pela distribuição exponencial com média θ horas. Qual a probabilidade de um dado componente atender à demanda de θ horas?

a) 1/2

b) $e^{-\theta^2}$

c) $e^{-\theta}$

d) $e^{-1/2}$

e) e^{-1}

4. (CESGRANRIO/2010 – Petrobrás) As ocorrências diárias de situações de emergência em uma instalação industrial são aleatórias e usualmente consideradas independentes umas das outras. Dessa forma, o modelo mais adequado para a simulação dos instantes de ocorrências é a Distribuição de Poisson e, consequentemente, os intervalos entre as ocorrências obedecem à Distribuição Exponencial. Na prática, observa-se que o tempo dedicado por um engenheiro à solução de cada emergência é bem modelado também pela Distribuição Exponencial. Esses são alguns dos motivos para que, em simulação desses processos de atendimento, o tempo (T) entre ocorrências e o tempo (T) de tratamento das mesmas sejam modelados por Distribuições Exponenciais que, entre outros aspectos, têm a propriedade denominada "ausência de memória" que (para quaisquer t > 0 e a > 0) é traduzida por:

a) P(T > t + a | T > a) = P(T > t)

b) Valor Esperado de T = variância de T ($\mu = \sigma^2$)

c) [Valor Esperado de T]^2 = variância de T $\mu^2 = \sigma^2$)

d) P(0 < T < a) > P(t < T < t + a)

e) P(0 < T < a) = P(t < T < t + a)

5. (CESGRANRIO/2013 – BNDES) O tempo de ligações telefônicas segue uma distribuição de probabilidade exponencial com média de 3 minutos. Um sujeito chega a um telefone público e descobre que a pessoa à sua frente está na ligação há pelo menos dois minutos.

Qual é a probabilidade de essa ligação durar pelo menos cinco minutos no total?

a) e^{-1}

b) e^{-2}

c) e^{-3}

d) $1 - e^{-3}$

e) $1 - e^{-5}$

6. (CESGRANRIO/2014 – EPE) Uma companhia possui dois geradores elétricos. O tempo até a falha de cada gerador se comporta segundo uma distribuição exponencial, com média de 10 anos. A companhia passa a usar o segundo gerador tão logo o primeiro em funcionamento falhe.

Qual é a variância do tempo total em que os dois geradores produziram energia?

a) 200

b) 150

c) 100

d) 20

e) 10

GABARITO

1. LETRA C
2. LETRA A
3. LETRA E
4. LETRA A
5. LETRA A
6. LETRA A

LISTA DE QUESTÕES – CESGRANRIO

Distribuição Normal

1. (CESGRANRIO/2010 – Petrobras – *Adaptada*) A distribuição normal é uma das mais utilizadas para inferência estatística da probabilidade de ocorrência de diversos fenômenos em engenharia. Nela, qualquer variável aleatória normal X é convertida em uma variável normal padronizada Z, tal que Z = $\frac{X-\mu}{\hat{\sigma}}$, onde $\hat{\sigma}$ é o desvio padrão e μ é a média aritmética. Na distribuição normal padronizada,

a) a variável Z é contínua e representa o número de desvios em relação à média.

b) a variável Z possui média não nula e desvio padrão igual a 1.

c) um ponto da curva da distribuição indica a probabilidade da variável aleatória assumir um valor real entre 0 e 1.

d) valores idênticos acima e abaixo da média têm probabilidades complementares, pois a curva é simétrica.

2. (CESGRANRIO/2006 – EPE) Se uma distribuição segue um padrão normal, é correto afirmar que:

a) 98% dos números estão a dois desvios padrão da média.

b) 95% dos números estão a 1,5 desvio padrão da média.

c) 95% dos números estão a um desvio padrão da média.

d) 86% dos números estão a um desvio padrão da média.

e) 68% dos números estão a um desvio padrão da média.

3. (CESGRANRIO/2012 – Petrobras) Considere que, no verão, o valor das vendas de determinada loja, nas sextas-feiras que sejam dias úteis, tem uma distribuição de probabilidades normal, com média igual a R$ 10.000,00 e desvio padrão de R$ 2.000,00. Com essa hipótese, a probabilidade de que, em uma dessas sextas-feiras úteis do verão, o valor das vendas seja inferior a R$ 8.000,00 é

a) menor que 20%

b) igual a 20%

c) maior que 20% e menor que 25%

d) igual a 25%

e) maior que 25%

4. (CESGRANRIO/2010 – BACEN) Estima-se que os retornos de um determinado mercado tenham distribuição normal, com média 20% e desvio padrão 10%. A probabilidade de perdas financeiras é de, aproximadamente,

a) 1%

b) 2,5%

c) 5%

d) 10%

e) 20%

5. (CESGRANRIO/2012 – Innova) Suponha que o tempo de vida de baterias de celular tenha distribuição normal com média de 120 minutos e variância de 100 minutos. Qual é a probabilidade aproximada de uma bateria durar menos que 100 minutos?

a) 0,15%

b) 2,5%

c) 5%

d) 10%

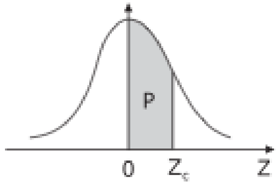
e) 16%

Para as próximas questões, utilize a tabela normal a seguir, constante nas respectivas provas da CESGRANRIO.

Distribuição Normal Padrão

$Z \sim N(0,1)$

Corpo da tabela dá a probabilidade p, tal que $p = P(0<Z<Z_c)$



parte inteira e primeira decimal de Z_c	Segunda decimal de Z_c: 0	1	2	3	4	5	6	7	8	9	parte inteira e primeira decimal de Z_c
	p = 0										
0,0	00000	00399	00798	01197	01595	01994	02392	02790	03188	03586	0,0
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535	0,1
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409	0,2
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173	0,3
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793	0,4
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240	0,5
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490	0,6
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524	0,7
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327	0,8
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891	0,9
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214	1,0
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298	1,1
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147	1,2
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774	1,3
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189	1,4
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408	1,5
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449	1,6
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327	1,7
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062	1,8
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670	1,9
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169	2,0
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574	2,1
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899	2,2
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158	2,3
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361	2,4
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520	2,5
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643	2,6
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736	2,7
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807	2,8
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861	2,9
3,0	49865	49869	49874	49878	49882	49886	49889	49893	49897	49900	3,0
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929	3,1
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950	3,2
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965	3,3
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976	3,4
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983	3,5
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989	3,6
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49992	3,7
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995	3,8
3,9	49995	49995	49996	49996	49996	49996	49996	49996	49997	49997	3,9
4,0	49997	49997	49997	49997	49997	49997	49998	49998	49998	49998	4,0
4,5	49999	50000	50000	50000	50000	50000	50000	50000	50000	50000	4,5

6. (CESGRANRIO/2010 – IBGE) Seja H a variável aleatória que representa as alturas dos cidadãos de certo país. Sabe-se que H tem distribuição normal com média 1,70 m e desvio padrão 0,04 m. A probabilidade de que um cidadão desse país tenha mais do que 1,75 m de altura é, aproximadamente,

a) 9,9%

b) 10,6%

c) 22,2%

d) 39,4%

e) 40,6%

7. (CESGRANRIO/2006 – EPE) Suponha os pesos dos pacotes de arroz normalmente distribuídos com média 1kg e desvio padrão 20g. Escolhendo um pacote ao acaso, qual é a probabilidade de ele pesar mais de 1 030g?

a) 13,4%

b) 11,6%

c) 10,0%

d) 8,4%

e) 6,7%

8. (CESGRANRIO/2012 – Petrobras) O peso de lotes produzidos por uma certa indústria segue uma distribuição normal, com média de 10 kg e desvio padrão de 0,2 kg. Em um lote dessa indústria, selecionado aleatoriamente, qual a probabilidade de o peso do lote não se afastar por mais de 1% do peso médio?

a) 50%

b) 38,30%

c) 17,42%

d) 7,96%

e) 0%

9. (CESGRANRIO/2016 – IBGE) O tempo de atendimento (em minutos) para chamadas de emergência do SAMU no Rio de Janeiro segue uma distribuição normal com média 12 e variância 25. Qual a probabilidade de que o tempo de atendimento para uma dada chamada exceda a 20 minutos?

a) 1,79%

b) 2,87%

c) 5,48%

d) 25,50%

e) 37,45%

10. (CESGRANRIO/2012 – Petrobras) Os salários de técnicos de uma empresa se distribuem normalmente com média de R$ 3.200,00 e desvio padrão de R$ 800,00. Selecionando-se aleatoriamente dois salários de técnicos dessa empresa, qual a probabilidade de pelo menos um deles ser superior a R$ 3.880,00?

a) 72,25%

b) 15,86%

c) 3,91%

d) 19,77%

e) 35,63%

11. (CESGRANRIO/2007 – TCE/RO) O gasto médio dos clientes de um posto de gasolina é uma variável aleatória normal com média R$ 100,00 e desvio padrão R$ 25,00. Os 10% dos que mais consomem recebem um tratamento VIP, incluindo lavagem de carroceria, calibragem nos pneus e verificação do óleo e da água. Quanto você precisa gastar nesse posto de gasolina, em reais, para obter tratamento VIP?

a) 158,00

b) 149,00

c) 141,00

d) 132,00

e) 128,00

12. (CESGRANRIO/2010 – IBGE) Suponha que as notas dos candidatos de um concurso público, em uma certa prova, sigam distribuição normal com média 7 e desvio padrão 1. A relação candidato/vaga é de 40 para 1. A nota mínima necessária para aprovação nessa prova é

a) 8,65

b) 8,96

c) 9,37

d) 9,58

e) 9,75

GABARITO

1. LETRA A
2. LETRA E
3. LETRA A
4. LETRA B
5. LETRA B
6. LETRA B
7. LETRA E
8. LETRA B
9. LETRA C
10. LETRA E
11. LETRA D
12. LETRA B

LISTA DE QUESTÕES – CESGRANRIO

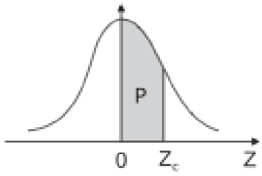
Soma de Variáveis e o Teorema Central do Limite

Para a próxima questão, utilize a tabela normal a seguir, constante na respectiva prova da CESGRANRIO.

Distribuição Normal Padrão

$Z \sim N(0,1)$

Corpo da tabela dá a probabilidade p, tal que $p = P(0<Z<Z_c)$



parte inteira e primeira decimal de Z_c	Segunda decimal de Z_c										parte inteira e primeira decimal de Z_c
	0	1	2	3	4	5	6	7	8	9	
	p = 0										
0,0	00000	00399	00798	01197	01595	01994	02392	02790	03188	03586	0,0
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535	0,1
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409	0,2
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173	0,3
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793	0,4
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240	0,5
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490	0,6
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524	0,7
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327	0,8
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891	0,9
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214	1,0
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298	1,1
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147	1,2
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774	1,3
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189	1,4
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408	1,5
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449	1,6
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327	1,7
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062	1,8
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670	1,9
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169	2,0
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574	2,1
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899	2,2
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158	2,3
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361	2,4
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520	2,5
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643	2,6
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736	2,7
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807	2,8
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861	2,9
3,0	49865	49869	49874	49878	49882	49886	49889	49893	49897	49900	3,0
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929	3,1
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950	3,2
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965	3,3
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976	3,4
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983	3,5
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989	3,6
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49992	3,7
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995	3,8
3,9	49995	49995	49996	49996	49996	49996	49996	49996	49997	49997	3,9
4,0	49997	49997	49997	49997	49997	49997	49998	49998	49998	49998	4,0
4,5	49999	50000	50000	50000	50000	50000	50000	50000	50000	50000	4,5

1. (CESGRANRIO/2011 – BNDES) Um produto tem massa normalmente distribuída com média 60 gramas e desvio padrão 8 gramas e é agrupado por dúzias. A probabilidade de a massa de uma dúzia ser superior a 750 gramas é, aproximadamente, de

a) 86%

b) 50%

c) 38%

d) 32%

e) 14%

GABARITO

1. LETRA E

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.