Estatística I

Licenciatura em Gestão do Desporto (LGD)

2.º Ano/1.º Semestre

2025/2026

Aulas Teórico-Práticas N.ºs 21 e 22 (Semana 11)

Docente: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt

https://doity.com.br/estatistica-aplicada-a-nutricao

https://basiccode.com.br/produto/informatica-basica/

Conteúdos Programáticos

Aulas TP (Semanas 1 e 3)	Aulas TP (Semanas 3 a 6)	Aulas TP (Semanas 7 a 9)	Aulas TP (Semanas 10 a 12)
• **Capítulo 1**: Análise Descritiva • **Capítulo 2**: Probabilidades	• **Capítulo 3**: Variáveis Aleatórias Unidimensionais	• **Capítulo 4**: Variáveis Aleatórias Multidimensionais	• **Capítulo 5**: Variáveis Aleatórias Especiais

Material didático: Exercícios do Livro Murteira et al (2015), Formulário e Tabelas Estatísticas

Bibliografia: B. Murteira, C. Silva Ribeiro, J. Andrade e Silva, C. Pimenta e F. Pimenta; *Introdução à Estatística*, 2ª ed., Escolar Editora, 2015.

5. Variáveis aleatórias especiais

5.1. Variáveis aleatórias discretas

5.1.1. A distribuição uniforme discreta

5.1.2. A variável aleatória de Bernoulli

5.1.3. A variável aleatória binomial

5.1.4. A variável aleatória de Poisson

5.2. Variáveis aleatórias contínuas

5.2.1. A variável aleatória uniforme contínua

5.2.2. A variável aleatória normal

5.2.3. A variável aleatória exponencial

5.2.4. As variáveis aleatórias gama e chi-quadrado

5.3. O Teorema Limite Central

Distribuição Normal: Exercícios do Murteira et al (2015)

Variáveis Aleatórias Contínuas

1

39. Se X tem distribuição normal com média 6 e variância 25, calcule:

a) $P(6 < X \leq 12)$.

b) $P(0 \leq X < 8)$.

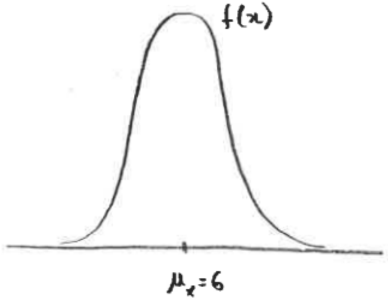
c) $P(X < -4)$.

d) $P(|X - 6| > 10)$.

e) O valor de k tal que $P(X > k) = 0.90$.

Exercício 39 (a)

$X \sim N(6, 25) \rightarrow \mu_x = 6 \; , \; \sigma_x^2 = 25$



(a)

$$P(6 < X \leq 12) \rightarrow \text{normalcdf}\,(6, 12, \overset{\mu}{\uparrow}{6}, \overset{\sigma}{\uparrow}{5}) \approx 0.38493$$

$$P(6 < X \leq 12) = P\left(\frac{6-6}{\sqrt{25}} < \underbrace{\frac{X - \mu_x}{\sigma_x}}_{Z \sim N(0,1)} \leq \frac{12-6}{\sqrt{25}}\right) = P(0 < Z \leq 1.2) = \underset{(F_Z)}{\Phi(1.2)} - \underset{(F_Z)}{\Phi(0)} =$$

$$= 0.8849 - 0.5 = 0.3849$$

Exercício 39 (b)

$P(0 \leq X < 8) \longrightarrow \text{normalcdf}(0, 8, 6, 5) \approx 0.54035$

$$P(0 \leq X < 8) = P\left(\frac{0-6}{\sqrt{25}} \leq \frac{X-\mu_X}{\sigma_X} < \frac{8-6}{\sqrt{25}}\right) = P(-1.2 \leq Z < 0.4) = \Phi(0.4) - \Phi(-1.2) =$$

$$= \Phi(0.4) - [1 - \Phi(1.2)] = 0.6554 - (1 - 0.8849) = 0.5403$$

Exercício 39 (c)

$P(X < -4) \rightarrow \text{normalcdf}(-10^{99}, -4, 6, 5) \approx 0.02275$

$$P(X < -4) = P\left(\frac{X - \mu_x}{\sigma_x} < \frac{-4-6}{\sqrt{25}}\right) = P(Z < -2) = \Phi(-2) = 1 - \Phi(2) = 1 - 0.9772 = 0.022$$

Exercício 39 (d)

$$P(|X-6|>10) = 1 - P(|X-6|\leq 10) = 1 - P(-10 \leq X-6 \leq 10) = 1 - \underbrace{P(-4 \leq X \leq 16)}_{\text{normalcdf}(-4,16,6,5)} \approx$$

$$\approx 1 - 0.9545 = 0.0455$$

$$P(|X-6|>10) = 1 - P(|X-6|\leq 10) = 1 - P(-10 \leq X - \overset{\mu_X}{\overset{\nearrow}{6}} \leq 10) = 1 - P\left(\frac{-10}{\sqrt{25}} \leq \frac{X-\mu_X}{\sigma_X} \leq \frac{10}{\sqrt{25}}\right) =$$

$$= 1 - P(-2 \leq Z \leq 2) = 1 - [\Phi(2) - \Phi(-2)] = 1 - \Phi(2) + [1 - \Phi(2)] =$$

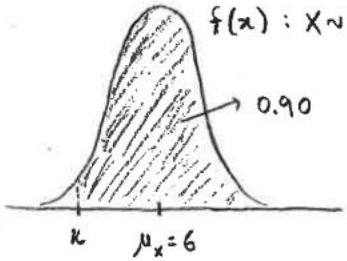
$$= 1 - \Phi(2) + 1 - \Phi(2) = 2 - 2\Phi(2) = 2[1-\Phi(2)] = 2(1-0.9772) =$$

$$= 2 \times 0.0228 = 0.0456$$

Exercício 39 (e)

$K: \; P(X > K) = 0.90$

$f(x): X \sim N(6, 25)$

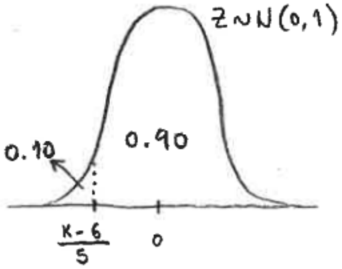


prob. à <u>esquerda</u>!

$P(X > K) = 0.90 \longrightarrow \text{invNorm}(0.10, 6, 5) \approx -0.4078$

Exercício 39 (e)

$$P(X > K) = 0.90 \Leftrightarrow P\left(\frac{X-\mu_X}{\sigma_X} > \frac{K-6}{5}\right) = 0.90 \Leftrightarrow P\left(Z > \frac{K-6}{5}\right) = 0.90$$



$$P\left(Z \leq \frac{K-6}{5}\right) = 0.10 \quad \text{(simetria} \Rightarrow$$

$$\Rightarrow P\left(Z \geq -\frac{K-6}{5}\right) = 0.10 \Leftrightarrow$$

$$\Leftrightarrow P\left(Z \geq \frac{6-K}{5}\right) = 0.10$$

Pela tabela 5 : $\frac{6-K}{5} = 1.282 \Leftrightarrow 6 - K = 1.282 \times 5 \Leftrightarrow -K = 1.282 \times 5 - 6 \Leftrightarrow$

$$\Leftrightarrow K = 6 - 1.282 \times 5 = -0.41$$

43. Considere que o tempo gasto numa visita à feira do livro é uma variável aleatória com distribuição normal, de média igual a duas horas. Suponha que apenas 2.5% dos visitantes permanecem mais de três horas.

a) Qual o desvio padrão da variável?

b) Sabendo que um visitante já chegou há uma hora, qual a probabilidade de se ir embora nos próximos 30 minutos?

c) Calcule a mediana e o intervalo interquartil de X, e interprete o seu significado.

d) Calcule a probabilidade de em 20 visitantes seleccionados ao acaso haver no máximo um que permaneça mais de três horas.

Exercício 43 (a)

X - v.a. tempo visita, em horas

$X \sim N(2, \sigma_x^2) \rightarrow \mu_x = 2 \; , \; \sigma_x^2 = ?$

Supor que 2.5% dos visitantes permanecem mais de 3 horas : $P(X > 3) = 0.025$

(a)

$\sigma_x = ?$

$$P(X>3) = 0.025 \Leftrightarrow P\left(\frac{X-\mu_x}{\sigma_x} > \frac{3-2}{\sigma_x}\right) = 0.025 \Leftrightarrow P\left(Z > \frac{1}{\sigma_x}\right) = 0.025$$

Pela tabela 5 : $\frac{1}{\sigma_x} = 1.96 \Leftrightarrow \sigma_x = \frac{1}{1.96} = 1.96^{-1} \approx 0.5102$

Exercício 43 (b)

$$P(X \leq 1.5 \mid X \geq 1) = \frac{P(X \leq 1.5 \cap X \geq 1)}{P(X \geq 1)} = \frac{P(1 \leq X \leq 1.5)}{P(X \geq 1)}$$

- $P(1 \leq X \leq 1.5) \rightarrow \text{normalcdf}(1, 1.5, 2, 0.5102) \approx 0.1385$
- $P(X \geq 1) \rightarrow \text{normalcdf}(1, 10^{99}, 2, 0.5102) \approx 0.975$

Logo, $P(X \leq 1.5 \mid X \geq 1) = \frac{0.1385}{0.975} \approx 0.1421$

Exercício 43 (c)

Dado que $X \sim N$ e a distribuição normal é simétrica, tem-se que: $med(x) = \mu_x = 2$.

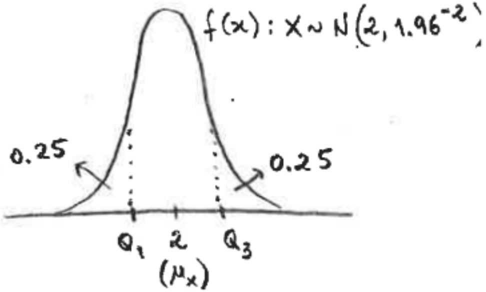
Isto significa que metade das visitas demora menos de 2 horas e a outra metade, mais de 2 horas.

Exercício 43 (c)

Intervalo interquartil de X : (Q_1, Q_3) ; onde Q_1 é o 1º quartil e Q_3 é o 3º quartil.

Logo, quer-se encontrar Q_1 e Q_3, tais que:

$$P(X \leq Q_1) = 0.25 \quad e \quad P(X \leq Q_3) = 0.75$$



1º Quartil

$$P(X \leq Q_1) = 0.25 \rightarrow \text{invNorm}(0.25, 2, 0.5102) \Rightarrow Q_1 \approx 1.656$$

Exercício 43 (c)

3º Quartil

$P(X \leq Q_3) = 0.75 \rightarrow \text{invNorm}(0.75, 2, 0.5102) \Rightarrow Q_3 \approx 2.344$

Logo, o intervalo interquartil de X é $(1.656, 2.344)$

Isto significa que 25% das visitas demora menos de 1.656 horas e outros 25% demora mais de 2.344 horas e metade das visitas tem duração entre estes dois valores.

Exercício 43 (d)

Num conjunto de 20 visitas haver, no máximo, 1 com duração superior a 3 horas.

Seja, Y - v.a. nº visitas com mais de 3 horas, num conjunto de 20

Então, $Y \sim B(20, \theta)$ $\theta = P(X>3) = 0.025 \Rightarrow Y \sim B(20, 0.025)$

Quer-se: $P(Y \leq 1) \approx 0.912 \rightarrow$ binomcdf $(20, 0.025, 1)$

Distribuição Exponencial

Variáveis Aleatórias Contínuas

2

Distribuição Exponencial

- A distribuição exponencial (ou exponencial negativa) tem a sua génese associada ao processo de Poisson, mas também é utilizada fora deste contexto;
- Se uma sucessão de eventos constitui um processo de Poisson de intensidade λ e se se inicia a contagem no instante 0, o **tempo de espera** pela chegada da primeira ocorrência é uma variável aleatória, com distribuição exponencial de parâmetro λ (dedução no livro).
- Diz-se que a v.a. X tem distribuição exponencial de parâmetro λ quando a sua função densidade é da forma,

e = 2,71828 (Número de Euler)

$$f(x|\lambda) = \begin{cases} 0 & x \leq 0, \\ \lambda e^{-\lambda x} & x > 0. \end{cases}$$

A função densidade de probabilidade de uma variável aleatória com **Distribuição Exponencial** pode ser escrita destas duas formas.

Formulário

- **EXPONENCIAL** $X \sim \text{Ex}(\lambda)$, $(\lambda > 0)$; $X \sim \text{Ex}(\lambda) \Leftrightarrow X \sim G(1, \lambda)$ Distribuição Gama

$f(x|\lambda) = \lambda e^{-\lambda x}$, $x > 0$ $E(X) = \frac{1}{\lambda}$; $\text{Var}(X) = \frac{1}{\lambda^2}$; $M_X(s) = \frac{\lambda}{\lambda - s}$, $s < \lambda$; $\gamma_1 = 2$; $\gamma_2 = 9$

Propriedades:

- $X_i \sim \text{Ex}(\lambda)$ $(i = 1, 2, ..., k)$ independentes $\Rightarrow$ $\sum_{i=1}^{k} X_i \sim G(k, \lambda)$ e $\min X_i \sim \text{Ex}(k\lambda)$

Exponencial vs Poisson

Observações:

- Existe uma relação muito importante entre a distribuição exponencial e a distribuição de Poisson, que surge muitas vezes na prática. Estando a observar a ocorrência de certos acontecimentos em intervalos de tempo, pretendemos caracterizar T o tempo ao fim do qual se verifica a primeira ocorrência.

Teorema

Seja X uma v.a. de Poisson de parâmetro λ. Seja T a v.a. que designa o tempo de espera pela ocorrência do primeiro acontecimento, então T tem distribuição exponencial, $T \frown Exp[\beta)$, de parâmetro $\beta = 1/\lambda$.

moduloI_aula2.pdf

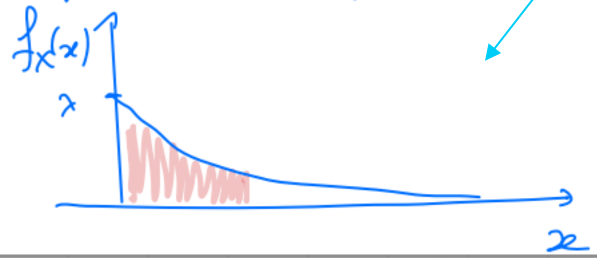
Distribuição Exponencial

Representação gráfica da função densidade de probabilidade de uma variável aleatória com **Distribuição Exponencial** de parâmetro λ.

ii) $X \sim \text{Exponencial}(\lambda)$ [lê-se: X tem distribuição exponencial de parâmetro $\lambda > 0$)

$$f_X(x) = \lambda e^{-\lambda x}, \; x \geq 0$$

$$E(X) = \frac{1}{\lambda}, \; Var(X) = \frac{1}{\lambda^2}$$



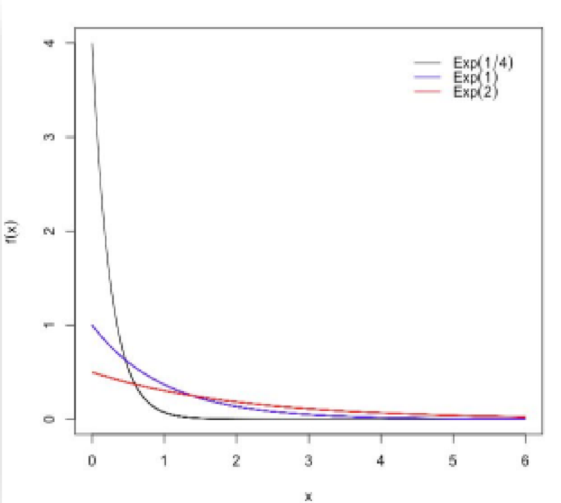
Note-se que $1/\lambda$ representa o valor médio de $X \sim \text{Exp}(\lambda)$ e λ representa uma frequência média.

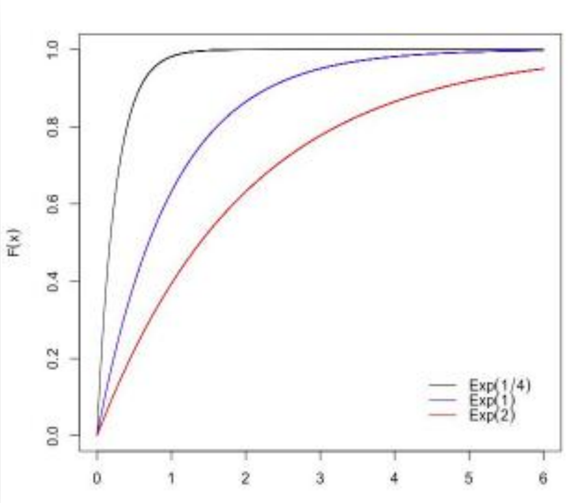
$$F_X(x) = \begin{cases} 0 & x < 0 \\ \int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

Função de distribuição de uma variável aleatória com **Distribuição Exponencial** de parâmetro λ.

Slides Professora Claúdia Nunes

Distribuição Exponencial





Representação gráfica da função densidade de probabilidade f(x) e da função de distribuição de uma variável aleatória com **Distribuição Exponencial** para vários valores do parâmetro λ.

Distribuição Exponencial

- A partir da fgm vem $E(X)=\frac{1}{\lambda}$, $\mathrm{Var}(X)=\frac{1}{\lambda^2}$, $\mathrm{CV}=1$, $\gamma_1=2$, $\gamma_2=9$.
- Moda: **Não existe** (a função densidade é decrescente com x e o domínio é aberto).
- Mediana: $\mu_e=(\ln 2)/\lambda$. Note-se que $\mu_e<\mu$ (como seria de esperar numa distribuição com assimetria positiva)

- É habitual classificar as distribuições simétricas comparando-as com a normal:
  - Distribuições *leptokurtica* $(\gamma_2>3)\rightarrow$ caudas mais "espessas" (com zona central mais "pontiaguda") que a distribuição normal.
  - Distribuições *platikurtica* $(\gamma_2<3)\rightarrow$ caudas mais "finas" (com zona central mais "achatada") que a distribuição normal.
  - Distribuição *mesokurtica* $(\gamma_2=3)$

https://fenix.iseg.ulisboa.pt/downloadFile/281608120794416/

Distribuição Exponencial: Propriedades

Duas propriedades importantes:

- "Falta de memória": $X \sim \text{Ex}(\lambda) \Rightarrow P(X > x + h \mid X > x) = P(X > h)$.
- $X_1, X_2, \ldots, X_k$ v.a. **independentes** então, $X_i \sim \text{Ex}(\lambda) \Rightarrow Y = \min_i X_i \sim \text{Ex}(k\lambda)$.

https://fenix.iseg.ulisboa.pt/downloadFile/281608120794416/

Distribuição Exponencial: Exemplo

Exemplo 5.20 (alterado) – Suponha-se que o tempo de vida útil de um dado componente é uma v.a. *X* com distribuição exponencial com média igual a 600 horas.

Se $E(X) = 1/\lambda = 600$, então tem-se $\lambda = 1/600$

- Probabilidade de o componente durar mais de 700 horas

$$P(X > 700) = \int_{700}^{+\infty} \frac{1}{600} e^{-x/600} dx = e^{-7/6} \approx 0.31.$$

https://fenix.iseg.ulisboa.pt/downloadFile/281608120794416/

Na **Distribuição Exponencial** tem-se $F(x) = P(X \leq x) = 1 - e^{-\lambda x}$, então $P(X > x) = 1-[1- e^{-\lambda x}] = e^{-\lambda x}$

Logo, alternativamente à resolução anterior tem-se $P(X>700) = 1-P(X\leq 700) = 1- [1- e^{-700/600}] = e^{-7/6}$

Distribuição Exponencial: Exemplo

Exemplo 5.20 (alterado) – Suponha-se que o tempo de vida útil de um dado componente é uma v.a. X com distribuição exponencial com média igual a 600 horas.

- Sabendo que o componente já durou 400 horas, qual a probabilidade de durar ainda mais 700 horas.

$$P(X > 1100 \mid X > 400) = \frac{P(X > 1100)}{P(X > 400)} = \frac{e^{-11/6}}{e^{-4/6}} = e^{-7/6}.$$

1100=700+400

Dada a falta de memória da exponencial,

$P(X > 700) = P(X > 1100 \mid X > 400)$. Mais simples

https://fenix.iseg.ulisboa.pt/downloadFile/281608120794416/

Distribuição Exponencial: Exemplo

- Num conjunto de 10 componentes a funcionar de forma independente qual a probabilidade de algum deles durar menos de 100 horas

$$P(\min X_i < 100) = \int_0^{100} \frac{1}{60} e^{-x/60} dx = 1 - e^{-100/60} \approx 0.8111$$

já que $\min X_i \sim \text{Ex}(1/60)$. Note-se que $P(X < 100) = 1 - e^{-1/6} \approx 0.1535$

Se E(X) = 1/λ = 600, então tem-se λ = 1/600

Como X ~ Exp(1/600), então Min Xi ~ Exp(10/600 = 1/60)

- $X_1, X_2, \ldots, X_k$ v.a. **independentes** então, $X_i \sim \text{Ex}(\lambda) \Rightarrow Y = \min_i X_i \sim \text{Ex}(k\lambda)$.

Distribuição Exponencial

Uma v.a. X tem distribuição exponencial com parâmetro λ >0 se sua função densidade de probabilidade é dada por

$$f(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}}, & x \geq 0 \\ 0, & \text{c.c.} \end{cases}$$

Notação: X ~ Exp(λ)

A função densidade de probabilidade de uma variável aleatória com **Distribuição Exponencial** pode ser escrita com esta parametrização.

Valor esperado e variância

$$E(X) = \lambda \quad \text{e} \quad V(X) = \lambda^2$$

Distribuição Exponencial: Exercícios do Murteira et al (2015)

3

Variáveis Aleatórias Contínuas

50. Um banco atende em média dois clientes em cada 3 minutos. Considere que o número de clientes atendidos é um processo de Poisson.

a) Qual a probabilidade de decorrerem 3 minutos sem qualquer cliente atendido?

b) Qual a probabilidade de o atendimento de um cliente demorar mais do que 3 minutos?

c) Compare os resultados das duas alíneas anteriores.

d) Qual a probabilidade de o atendimento de um cliente demorar entre 3 e 6 minutos?

Exercício 50 (a)

X - v.a. nº clientes atendidos por cada 3 minutos

$X \sim Po(2)$ $\lambda = E(X) =$ média de clientes atendidos por cada 3 minutos

(a)

$P(X=0) \approx 0.1353$ $\longrightarrow$ poissonpdf (2,0) ; $P(X=0) = \frac{2^0 e^{-2}}{0!} = e^{-2} \approx 0.1353$

Exercício 50 (b)

Analisa-se agora o tempo de atendimento por cliente, que corresponde à distribuição exponencial. Seja

Y - v.a. tempo atendimento por cliente, em minutos $\Rightarrow Y \sim Ex(\lambda)$

<u>Determinar λ da exponencial</u>

Na exponencial: $E(Y) = \frac{1}{\lambda}$.

Como $E(Y)$ é o tempo médio / cliente (minutos), tem-se que: $E(Y) = 3/2$.

Logo, $\frac{1}{\lambda} = \frac{3}{2} \Leftrightarrow \lambda = 2/3$

Assim, $Y \sim Ex(2/3) \Rightarrow f(y) = \frac{2}{3} e^{-\frac{2}{3}y} \quad (y>0)$

Então: $P(Y>3) = \int_3^{+\infty} f(y)\,dy = \int_3^{+\infty} \frac{2}{3} e^{-\frac{2}{3}y}\,dy = \left[-e^{-\frac{2}{3}y}\right]_3^{+\infty} = -e^{-\infty} - \left(-e^{-2}\right) =$

$= e^{-2} \simeq 0.1353$

Exercício 50 (c)

As probabilidades calculadas em (a) e (b) são iguais.

Faz sentido, pois se o tempo de atendimento de um cliente for superior a 3 minutos (b

então nenhum cliente é atendido em 3 minutos (a).

Exercício 50 (d)

$$P(3 \leq Y \leq 6) = \int_3^6 \frac{2}{3} e^{-\frac{2}{3}y}\, dy = \left[-e^{-\frac{2}{3}y}\right]_3^6 = -e^{-4} - \left(-e^{-2}\right) = e^{-2} - e^{-4} \approx 0.117$$

4

Distribuição do Qui-Quadrado

Variáveis Aleatórias Contínuas

Distribuição do Qui-Quadrado: Função Densidade de Probabilidade

Definição: A v. a. contínua X segue uma **distribuição Qui-Quadrado com $\boldsymbol{n}$ graus de liberdade**, i. e., $\boldsymbol{X \sim \chi^2_n}$, se a sua função densidade de probabilidade é:

$$\boxed{f(x) = \frac{1}{2^{\frac{n}{2}}\,\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}},\ x > 0,\ n > 0,}$$

onde $\Gamma(\alpha)$ é a função Gama, definida por $\Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx$.

Distribuição do Qui-Quadrado: Função Densidade de Probabilidade

O parâmetro caracterizador desta distribuição é n.

Principais características:

- A v. a. só toma valores positivos;
- É uma função não simétrica.

A forma da distribuição depende dos graus de liberdade (Figura 5.12), tornando-se menos assimétrica com o aumento do número de graus de liberdade. Esta distribuição está tabelada (Anexo B).

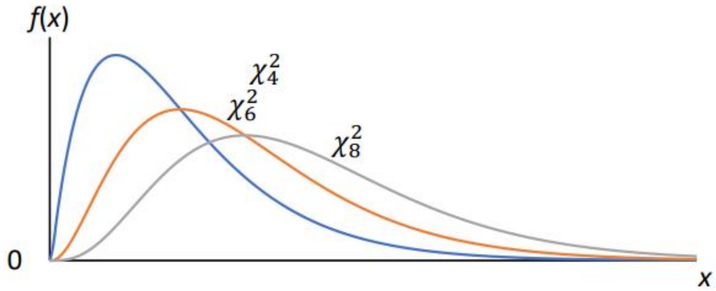


***Figura 5.12:** Função densidade de probabilidade distribuição Qui-Quadrado para diferentes graus de liberdade.*

[Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu]

Distribuição do Qui-Quadrado

- A distribuição do qui-quadrado (χ^2) é uma distribuição de probabilidade contínua.

- É uma das distribuições de probabilidade mais usadas em inferência estatística – *ex.*, testes de hipóteses, construção de intervalos de confiança.

- A distribuição do qui-quadrado com n graus de liberdade 'surge' quando tomamos o quadrado de uma distribuição normal padrão.

  (*i.e.*, a distribuição de uma soma de quadrados de n variáveis aleatórias independentes normalmente distribuídas)

- Se X tem uma distribuição do qui-quadrado com n graus de liberdade, escreve-se

  $$\boxed{X \sim \chi^2\,(n)}$$

- A distribuição do qui-quadrado tem um parâmetro: n – nº de termos independentes num somatório de quadrados (*i.e.*, o número de Xis)

Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu

Distribuição do Qui-Quadrado

- **Teorema da aditividade**

Se X_i $(i = 1, 2, 3, \dots, n)$ são variáveis aleatórias independentes e se $X_i \sim \chi^2(n_i)$

então: $\Sigma X_i \sim \chi^2_{(m)}$

(*i.e.*, ΣX_i é uma distribuição do qui-quadrado com $m = \Sigma n_i$ graus de liberdade)

Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu

Distribuição do Qui-Quadrado

- **Outros teoremas**

  - Se Z tem uma distribuição normal padrão, então Z^2 tem distribuição do qui-quadrado com 1 grau de liberdade.

$$\boxed{Z^2 = \left(\frac{X-\mu}{\sigma}\right)^2 \cap \chi^2_{(1)}}$$

  - Se X_i (i = 1, 2, 3, …, n) são variáveis aleatórias independentes e normalmente distribuídas com média μ_i e desvio-padrão σ_i, então:

$$\boxed{\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2 \cap \chi^2_{(n)}}$$

Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu

Distribuição do Qui-Quadrado

- É uma distribuição positiva e não simétrica.
- E[X] = n
- Var[X] = 2n
- Está tabelada para algumas probabilidade e n ≤ 30.
- Quando n > 30, pode usar-se a aproximação à distribuição normal:

$$\boxed{\sqrt{2X} - \sqrt{2n} \,\dot{\frown}\, N(0,1)}$$

Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu

Se $X \sim \chi^2_n$, então $\mu_X = E(X) = n$ e $\sigma^2_X = Var(X) = 2n$.

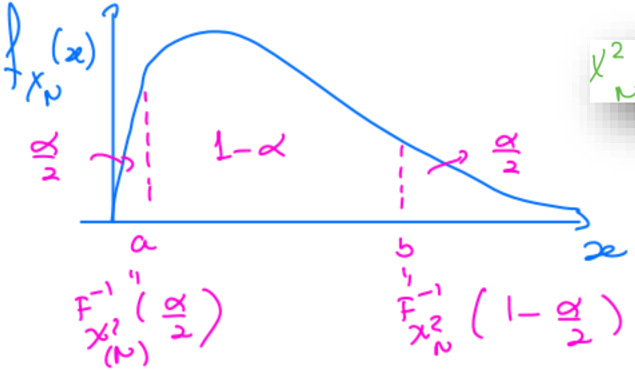
Usualmente utiliza-se a notação $\chi^2_{n;\,\alpha}$ para representar o quantil de probabilidade α de uma v. a. $X \sim \chi^2_n$. Portanto, $\chi^2_{n;\,\alpha}$ corresponde ao menor valor k tal que $P(X \leq k) = \alpha$.

Distribuição do Qui-Quadrado: Resumo

Propriedades:

i) $f_{\chi^2_N}(x) > 0$ para $x > 0$

ii) Não é uma distribuição simétrica



χ^2_N lê-se "distribuição do qui-quadrado com N graus de liberdade"

Distribuição do Qui-Quadrado

Formulário

- **QUI-QUADRADO** $X \sim \chi^2(n)$, (n inteiro positivo).

$$X \sim \chi^2(n) \Leftrightarrow X \sim G\left(n/2; 1/2\right) \; ; E(X) = n \; ; \; \mathrm{Var}(X) = 2n \; ; \; M_X(s) = (1-2s)^{-\frac{n}{2}}, s < \frac{1}{2} \quad ; \; \gamma_1 = \sqrt{\frac{8}{n}} \; ; \; \gamma_2 = 3 + \frac{12}{n}$$

Propriedades:

- $X_i \sim \chi^2_{(n_i)}$ ($i = 1, 2, \ldots, k$) independentes $\Rightarrow \sum_{i=1}^{k} X_i \sim \chi^2_{(n)}$ com $n = \sum_{i=1}^{k} n_i$
- $X \sim G(n; \lambda) \Leftrightarrow 2\lambda X \sim \chi^2(2n)$
- $X_i \sim N(0,1)$, ($i = 1,2,\ldots,n$) independentes $\Rightarrow \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$
- $X \sim \chi^2(n) \Rightarrow \sqrt{2X} - \sqrt{2n-1} \overset{a}{\sim} N(0,1)$

Distribuição Gama
(ver slides a seguir)

Distribuição do Qui-Quadrado: Exercícios

5

Variáveis Aleatórias Contínuas

1. Suponha que $X \sim \chi^2_{(10)}$. Qual o valor de <u>a</u> que garante $P[X \leq a] = 0{,}9$.

2. Suponha que $X \sim \chi^2_{(10)}$.

   a) Quais os valores de <u>a</u> e <u>b</u> que garantem:

   $P[X < a] = 0{,}95$ e $P[X > b] = 0{,}975$

   b) Calcule <u>c</u> e <u>d</u> tais que $P[c < X < d] = 0{,}9$.

   c) Calcule $P[X > 30{,}6]$.

3. Suponha que $Y \sim \chi^2_{(60)}$. Calcule $P[Y < 12]$.

1. Suponha que $X \sim \chi^2_{(10)}$. Qual o valor de a que garante $P[X \leq a] = 0,9$.

2. Suponha que $X \sim \chi^2_{(10)}$.

   a) Quais os valores de a e b que garantem:

      $P[X < a] = 0,95$ e $P[X > b] = 0,975$

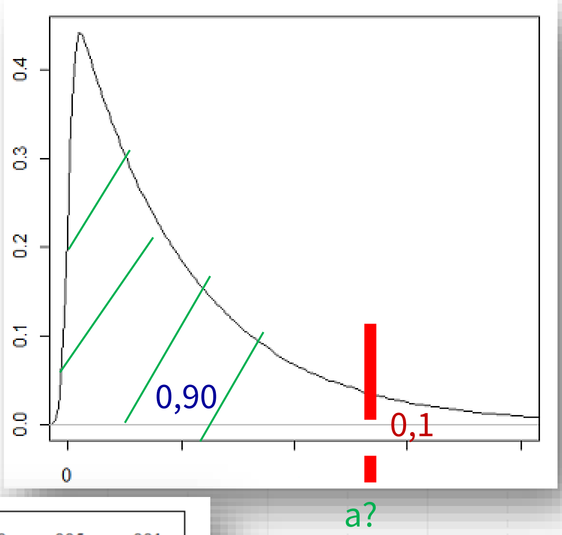
   b) Calcule c e d tais que $P[c < X < d] = 0,9$.

   c) Calcule $P[X > 30,6]$.

3. Suponha que $Y \sim \chi^2_{(60)}$. Calcule $P[Y < 12]$.

Exercício 1

Área total é igual a 1



$P(X \leq a) = 0,90 \Leftrightarrow a = F(0,90)^{-1} = 15,987$

$\chi^2_{n.\varepsilon} : P(X > \chi^2_{n.\varepsilon}) = \varepsilon$

n \ ε	.995	.990	.975	.950	.900	.750	.500	.250	.100	.050	.025	.010	.005	.001
1	.000	.000	.001	.004	.016	.102	.455	1.323	2.706	3.841	5.024	6.635	7.879	10.827
2	.010	.020	.051	.103	.211	.575	1.386	2.773	4.605	5.991	7.378	9.210	10.597	13.815
3	.072	.115	.216	.352	.584	1.213	2.366	4.108	6.251	7.815	9.348	11.345	12.838	16.266
4	.207	.297	.484	.711	1.064	1.923	3.357	5.385	7.779	9.488	11.143	13.277	14.860	18.466
5	.412	.554	.831	1.145	1.610	2.675	4.351	6.626	9.236	11.070	12.832	15.086	16.750	20.515
6	.676	.872	1.237	1.635	2.204	3.455	5.348	7.841	10.645	12.592	14.449	16.812	18.548	22.457
7	.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	24.321
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	26.124
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	27.877
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	29.588

1. Suponha que $X \sim \chi^2_{(10)}$. Qual o valor de a que garante $P[X \leq a] = 0{,}9$.
2. Suponha que $X \sim \chi^2_{(10)}$.

   a) Quais os valores de a e b que garantem:

      $P[X < a] = 0{,}95$ e $P[X > b] = 0{,}975$

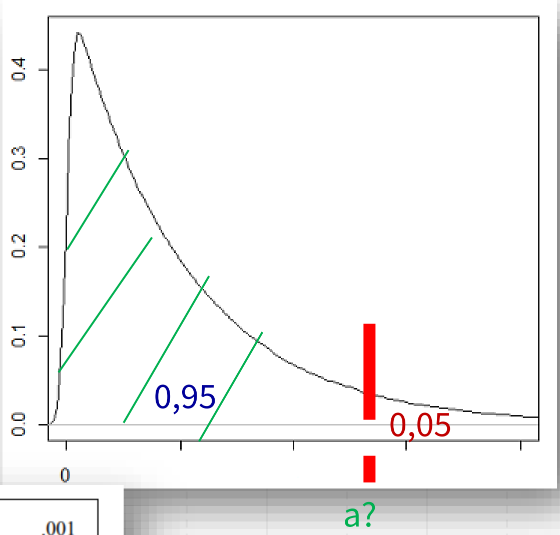
   b) Calcule c e d tais que $P[c < X < d] = 0{,}9$.

   c) Calcule $P[X > 30{,}6]$.

3. Suponha que $Y \sim \chi^2_{(60)}$. Calcule $P[Y < 12]$.

Exercício 2 a)

Área total é igual a 1



$P(X < a) = 0{,}95 \Leftrightarrow a = F(0{,}95)^{-1} = 18{,}307$
$P(X > b) = 0{,}975 \Leftrightarrow b = F(0{,}025)^{-1} = 3{,}247$

$\chi^2_{n.\varepsilon} : P(X > \chi^2_{n.\varepsilon}) = \varepsilon$

n \ ε	.995	.990	.975	.950	.900	.750	.500	.250	.100	.050	.025	.010	.005	.001
1	.000	.000	.001	.004	.016	.102	.455	1.323	2.706	3.841	5.024	6.635	7.879	10.827
2	.010	.020	.051	.103	.211	.575	1.386	2.773	4.605	5.991	7.378	9.210	10.597	13.815
3	.072	.115	.216	.352	.584	1.213	2.366	4.108	6.251	7.815	9.348	11.345	12.838	16.266
4	.207	.297	.484	.711	1.064	1.923	3.357	5.385	7.779	9.488	11.143	13.277	14.860	18.466
5	.412	.554	.831	1.145	1.610	2.675	4.351	6.626	9.236	11.070	12.832	15.086	16.750	20.515
6	.676	.872	1.237	1.635	2.204	3.455	5.348	7.841	10.645	12.592	14.449	16.812	18.548	22.457
7	.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	24.321
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	26.124
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	27.877
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	29.588

1. Suponha que $X \sim \chi^2_{(10)}$. Qual o valor de a que garante $P[X \leq a] = 0{,}9$.

2. Suponha que $X \sim \chi^2_{(10)}$.

   a) Quais os valores de a e b que garantem:

      $P[X < a] = 0{,}95$ e $P[X > b] = 0{,}975$

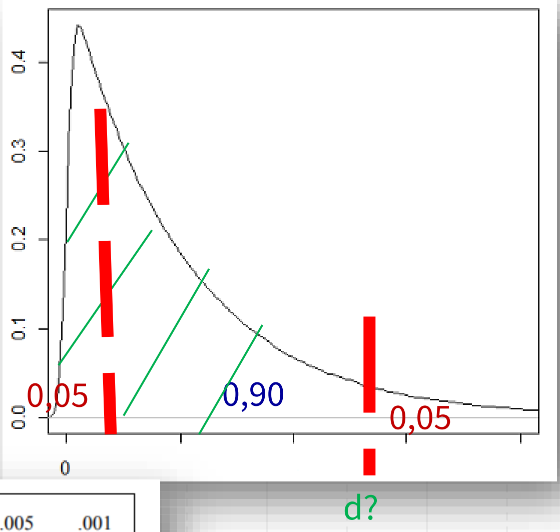
   b) Calcule c e d tais que $P[c < X < d] = 0{,}9$.

   c) Calcule $P[X > 30{,}6]$.

3. Suponha que $Y \sim \chi^2_{(60)}$. Calcule $P[Y < 12]$.

Exercício 2 b)

Área total é igual a 1



$P(c < X < d) = 0{,}90 \Leftrightarrow P(X < d) - P(X < c) = F(d) - F(c) = 0{,}90$

Logo, tem-se $d = F(0{,}95)^{-1} = 18{,}307$ e $c = F(0{,}05)^{-1} = 3{,}940$

$\chi^2_{n.\varepsilon} : P(X > \chi^2_{n.\varepsilon}) = \varepsilon$

ε / n	.995	.990	.975	.950	.900	.750	.500	.250	.100	.050	.025	.010	.005	.001
1	.000	.000	.001	.004	.016	.102	.455	1.323	2.706	3.841	5.024	6.635	7.879	10.827
2	.010	.020	.051	.103	.211	.575	1.386	2.773	4.605	5.991	7.378	9.210	10.597	13.815
3	.072	.115	.216	.352	.584	1.213	2.366	4.108	6.251	7.815	9.348	11.345	12.838	16.266
4	.207	.297	.484	.711	1.064	1.923	3.357	5.385	7.779	9.488	11.143	13.277	14.860	18.466
5	.412	.554	.831	1.145	1.610	2.675	4.351	6.626	9.236	11.070	12.832	15.086	16.750	20.515
6	.676	.872	1.237	1.635	2.204	3.455	5.348	7.841	10.645	12.592	14.449	16.812	18.548	22.457
7	.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	24.321
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	26.124
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	27.877
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	29.588

1. Suponha que $X \sim \chi^2_{(10)}$. Qual o valor de a que garante $P[X \leq a] = 0,9$.

2. Suponha que : $X \sim \chi^2_{(10)}$.

   a) Quais os valores de a e b que garantem:

      $P[X < a] = 0,95$ e $P[X > b] = 0,975$

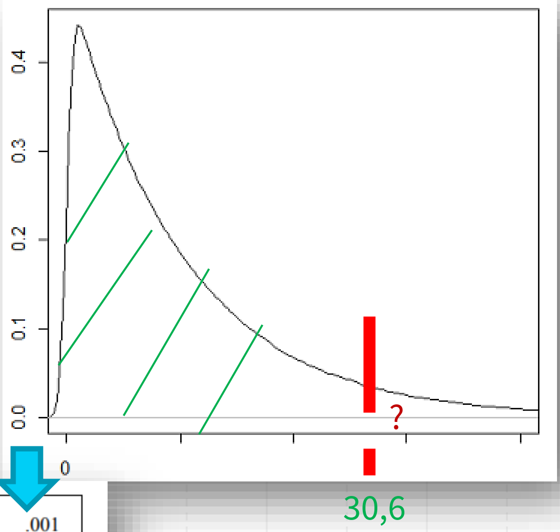
   b) Calcule c e d tais que $P[c < X < d] = 0,9$.

   c) Calcule $P[X > 30,6]$.

3. Suponha que $Y \sim \chi^2_{(60)}$. Calcule $P[Y < 12]$.

Exercício 2 c)

Área total é igual a 1



$P(X > 30,6) \sim P(X > 29,588) = 0,001$

$\chi^2_{n.\varepsilon} : P(X > \chi^2_{n.\varepsilon}) = \varepsilon$

n \ ε	.995	.990	.975	.950	.900	.750	.500	.250	.100	.050	.025	.010	.005	.001
1	.000	.000	.001	.004	.016	.102	.455	1.323	2.706	3.841	5.024	6.635	7.879	10.827
2	.010	.020	.051	.103	.211	.575	1.386	2.773	4.605	5.991	7.378	9.210	10.597	13.815
3	.072	.115	.216	.352	.584	1.213	2.366	4.108	6.251	7.815	9.348	11.345	12.838	16.266
4	.207	.297	.484	.711	1.064	1.923	3.357	5.385	7.779	9.488	11.143	13.277	14.860	18.466
5	.412	.554	.831	1.145	1.610	2.675	4.351	6.626	9.236	11.070	12.832	15.086	16.750	20.515
6	.676	.872	1.237	1.635	2.204	3.455	5.348	7.841	10.645	12.592	14.449	16.812	18.548	22.457
7	.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	24.321
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	26.124
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	27.877
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	29.588

1. Suponha que $X \sim \chi^2_{(10)}$. Qual o valor de a que garante $P[X \leq a] = 0{,}9$.

2. Suponha que $X \sim \chi^2_{(10)}$.

   a) Quais os valores de a e b que garantem:

      $P[X < a] = 0{,}95$ e $P[X > b] = 0{,}975$

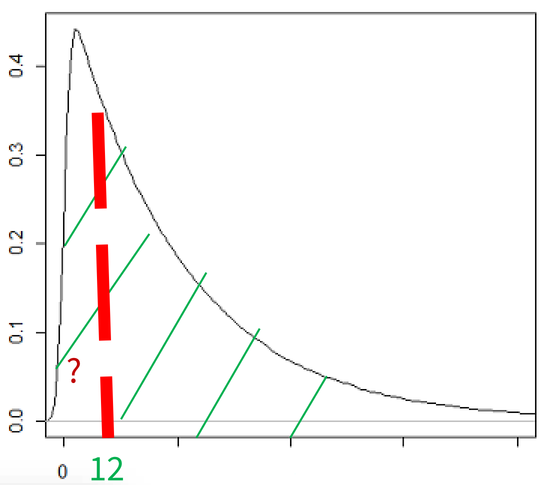
   b) Calcule c e d tais que $P[c < X < d] = 0{,}9$.

   c) Calcule $P[X > 30{,}6]$.

3. Suponha que $Y \sim \chi^2_{(60)}$. Calcule $P[Y < 12]$.

Exercício 3

Área total é igual a 1



$P(Y < 12) \sim 1-P(Y > 35{,}534) = 1-0{,}995 = 0{,}005$

Alternativamente, pelo TLC (aproximação à distribuição Normal)

$P(Y < 12) \sim P([Y-n]/(2n)^{-0{,}5} < (12 - 60) / 10{,}95) = P(Z < -4{,}38)$

$= \Phi(-4{,}38) = 1 - \Phi(4{,}48) \sim 1 - 0{,}999978 = 0{,}000022$ (ver mais à frente)

$\chi^2_{n.\varepsilon} : P(X > \chi^2_{n.\varepsilon}) = \varepsilon$

n \ ε	.995	.990	.975	.950	.900	.750	.500	.250	.100	.050	.025	.010	.005	.001
1	.000	.000	.001	.004	.016	.102	.455	1.323	2.706	3.841	5.024	6.635	7.879	10.827
2	.010	.020	.051	.103	.211	.575	1.386	2.773	4.605	5.991	7.378	9.210	10.597	13.815
3	.072	.115	.216	.352	.584	1.213	2.366	4.108	6.251	7.815	9.348	11.345	12.838	16.266
4	.207	.297	.484	.711	1.064	1.923	3.357	5.385	7.779	9.488	11.143	13.277	14.860	18.466
5	.412	.554	.831	1.145	1.610	2.675	4.351	6.626	9.236	11.070	12.832	15.086	16.750	20.515
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	73.403
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	86.660
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	99.608
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	112.317
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	124.839

58. Se $X \sim \chi^2(23)$, obtenha:

a) $P(14.85 < X < 32.01)$.

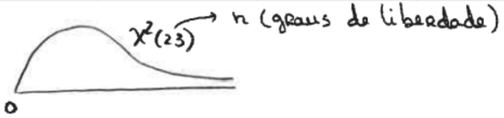
b) Os valores de a e b tais que $P(a < X < b) = 0.95$ e $P(X < a) = 0.025$.

c) A média e a variância de X.

d) $\chi^2_{23,0.05}$ e $\chi^2_{23,0.95}$.

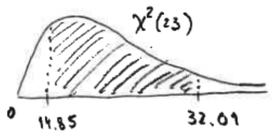
Exercício 58 (a)

$X \sim \chi^2(23)$



(a)

$P(14.85 < X < 32.01) \approx 0.8 \quad \rightarrow \quad \chi^2 cdf\,(14.85, 32.01, 23)$



$P(14.85 < X < 32.01) = P(X > 14.85) - P(X > 32.01)$

Pela tabela 6 :

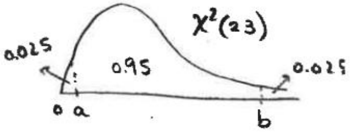
- $P(X > 14.85) = 0.9$
- $P(X > 32.01) = 0.1$

$P(14.85 < X < 32.01) = 0.9 - 0.1 = 0.8$

Exercício 58 (b), (c) e (d)

(b)

a, b : $P(a < X < b) = 0.95$ e $P(X < a) = 0.025$



Pela tabela 6 :

$P(X < a) = 0.025 \Leftrightarrow P(X \geq a) = 0.975 \Rightarrow \boxed{a = 11.689}$

$P(X > b) = 0.025 \Rightarrow \boxed{b = 38.076}$

(c)

$E(X) = n = 23$ $\quad$ $Var(X) = 2n = 2 \times 23 = 46$

(d)

$\chi^2_{23,0.05}$ é o valor numa $\chi^2(23)$ com probabilidade à direita de 0.05 $\Rightarrow \chi^2_{23,0.05} = 35.172$

$\chi^2_{23,0.95}$ é o valor numa $\chi^2(23)$ com probabilidade à direita de 0.95 $\Rightarrow \chi^2_{23,0.95} = 13.091$

Obrigada!

Questões?