Estatística I

Licenciatura em Gestão do Desporto (LGD)

2.º Ano/1.º Semestre

2023/2024

Aulas Teórico-Práticas N.ºs 22 e 23 (Semana 13)

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.

Distribuição Exponencial

Variáveis Aleatórias Contínuas

1

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

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

  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

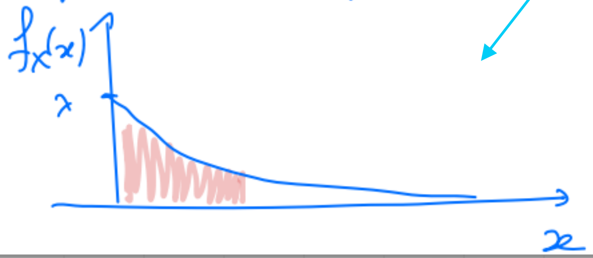
Distribuição Exponencial

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 $\lambda > 0$)

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

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



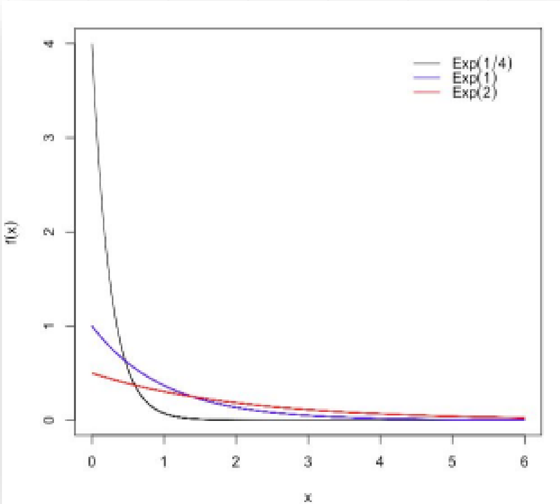
Note-se que 1/λ representa o valor médio de X ~ Exp(λ) e λ representa uma frequência média.

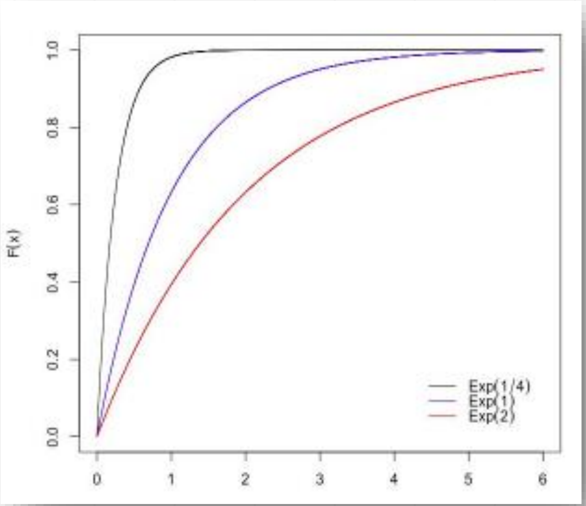
$$F_X(x) = \begin{cases} 0 & x < 0 \\ \int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

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 $E(X) = \frac{1}{\lambda}$, $\text{Var}(X) = \frac{1}{\lambda^2}$, $\text{CV} = 1$, $\gamma_1 = 2$, $\gamma_2 = 9$.
- 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": $X \sim \mathrm{Ex}(\lambda) \Rightarrow P(X > x+h \mid X > x) = P(X > h)$.
- $X_1, X_2, \ldots, X_k$ v.a. **independentes** então, $X_i \sim \mathrm{Ex}(\lambda) \Rightarrow Y = \min_i X_i \sim \mathrm{Ex}(k\lambda)$.

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 $P(X>700) = 1-P(X\leq700) = 1- [1- e^{-700/600}] = e^{-7/6}$

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 $\min X_i \sim \mathrm{Ex}(1/60)$. 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, $X_i \sim \mathrm{Ex}(\lambda) \Rightarrow Y = \min_i X_i \sim \mathrm{Ex}(k\lambda)$.

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

Variáveis Aleatórias Contínuas

2

Exemplo: Cerca de 15 clientes utilizam o caixa eletrônico por hora. Supondo que a distribuição de tempos de chegada é exponencial, qual é a probabilidade de que o tempo de chegada entre clientes consecutivos seja:

(a) Menor que três minutos?

(b) Maior que 3 minutos?

(c) Entre 2 e 4 minutos?

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

Exercício: Distribuição Exponencial

- A variável aleatória X= tempo

- Note que foi dada a frequência de chegada $\lambda = 1/\beta = 15$ /hora

-3 min = 0,05 horas

-2 min = 0,0333 horas

-4 min = 0,0666 horas

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

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

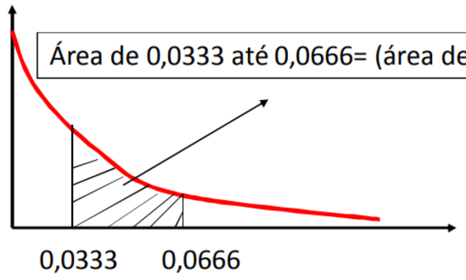
Exercício: Distribuição Exponencial

(a) $P(t < 0{,}05) = 1 - e^{-\lambda t} = 1 - e^{-(15).(0{,}05)} = 0{,}5276$

(b) $P(t \geq 0{,}05) = 1 - P(t < 0{,}05) = 1 - 0{,}5276 = 0{,}4734$

(c)

Área de 0,0333 até 0,0666= (área de 0 até 0,0666) – (área de 0 até 0,033)



$P(0{,}0333 < t < 0{,}0666) = P(t<0{,}0666) - P(t<0{,}0333) = e^{-(15).(0{,}0666))} - e^{-(15).(0{,}0333))}$

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

Distribuição Exponencial: Exercícios do Murteira et al (2015)

Variáveis Aleatórias Contínuas

3

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)

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

Teorema do Limite Central

4

Distribuição Assintótica da Soma e da Média de V.A.´s: Teorema do Limite Central (TLC)

- Na maioria das situações é difícil determinar a distribuição da soma de variáveis (mesmo que sejam independentes!). O teorema seguinte **justifica a grande utilidade** e **importância da distribuição normal** (quer em probabilidades quer em estatística).

- O teorema que vamos ver de seguida diz-nos que, para um grande número de v.a.'s com idêntico comportamento (distribucional), seja ele qual for, e mesmo que seja praticamente desconhecido, a distribuição da soma das v.a.'s aproxima-se de uma distribuição Normal; e tem uma distribuição que é tão mais próxima da Normal, quanto maior for o número de v.a.'s da soma.

TLC

Teorema do Limite Central

Seja $X_1, \ldots, X_n$ uma sequência de v.a. independentes e identicamente distribuídas com valor esperado $\mu < \infty$ e variância $\sigma^2 < \infty$. Considere-se $S_n = \sum_{i=1}^{n} X_i$ e $\bar{X}_n = \frac{S_n}{n}$, então quando $n \to +\infty$

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \overset{a}{\sim} N(0,1),$$

ou seja,

$$\lim_{n \to +\infty} P\left(\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \leq z\right) = \Phi(z),$$

onde $\Phi(.)$ é a função de distribuição da normal reduzida i.e., $N(0,1)$.

Slides da Professora Conceição Amado

TLC

Teorema do Limite Central

De modo equivalente, se tem

$$\frac{\bar{X}_n - E\left(\bar{X}_n\right)}{\sqrt{V\left(\bar{X}_n\right)}} = \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} \overset{a}{\sim} N\left(0,1\right),$$

ou seja,

$$\lim_{n \to +\infty} P\left(\frac{\bar{X}_n - E\left(\bar{X}_n\right)}{\sqrt{V\left(\bar{X}_n\right)}} \leq z\right) = \Phi(z)$$

Slides da Professora Conceição Amado

Valor Médio e Variância de Somas e Médias de V.A.´s Independentes

Observar que:

- $$E\left(S_n\right) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = nE(X_1) = n\mu$$

  e como $X_1, X_2 \cdots X_n$ são v.a. independentes tem-se

  $$V\left(S_n\right) = V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = nV(X_1) = n\sigma^2$$

- $$E\left(\bar{X}_n\right) = E\left(\sum_{i=1}^{n} X_i/n\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = E(X_1) = \mu$$

  e como $X_1, X_2 \cdots X_n$ são v.a. independentes tem-se

  $$V\left(\bar{X}_n\right) = V\left(\sum_{i=1}^{n} X_i/n\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n}V(X_1) = \frac{\sigma^2}{n}$$

Formulário

AMOSTRAGEM. DISTRIBUIÇÕES POR AMOSTRAGEM

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \;; \qquad S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n} = \frac{\sum_{i=1}^{n} X_i^2}{n} - \bar{X}^2 \;; \quad (n-1)S'^2 = n\,S^2$$

$$E(\bar{X}) = \mu \;; \qquad \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n} \;; \qquad E(S^2) = \frac{n-1}{n}\sigma^2 \;; \qquad E(S'^2) = \sigma^2$$

TLC: Aproximações

Observações:

- O TLC deve ser usado apenas quando as v.a.'s $X_1, \ldots, X_n$ não têm distribuição normal!
- A demonstração do teorema exige algumas ferramentas matemáticas avançadas.
- As v.a. $X_1, \ldots, X_n$ podem ser discretas ou contínuas.
- Geralmente considera-se n grande se $n \geq 30$
- As distribuições binomial e de Poisson podem ser aproximadas pela distribuição normal (na secção anterior vimos que podem ser escritas como somas de variáveis aleatórias). Assim:
  - $X \sim Bin(n, p)$ pode ser aproximada por $\tilde{X} \sim N(np, np(1-p))$, a aproximação tem menor erro quando $np > 5$ e $n(1-p) > 5$
  - $X \sim Poisson(\lambda)$ pode ser aproximada por $\tilde{X} \sim N(\lambda, \lambda)$, a aproximação tem menor erro quando $\lambda > 5$

Teorema de De Moivre-Laplace: Binomial - Normal

T. Limite Central: Aplicação a aproximações para distribuições discretas

Corolário 5.3 (T. de De Moivre-Laplace) – Dada a sucessão de variáveis aleatórias iid, $X_1, X_2, \ldots, X_n, \ldots$, com distribuição de Bernoulli de média $E(X_i) = \theta$ e, portanto, $\mathrm{Var}(X_i) = \theta(1-\theta)$, tem-se,

$$\frac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1).$$

Nota: $X = \sum_{i=1}^{n} X_i \sim B(n;\theta)$. Quando n é grande, utilizar o corolário. Mas...

com Correção de Continuidade.

Murteira et al (2015)

Aproximações Baseadas no TLC: Binomial - Normal

- Probabilidades associadas a uma distribuição Binomial, B(n,p), podem ser aproximadas utilizando uma distribuição Normal, N(μ,σ), com $\mu = np$ e $\sigma = \sqrt{np(1-p)}$.

  Para que a aproximação não seja muito má, devemos ter $np \geq 5$ e $n(1-p) \geq 5$.

Aproximação da Normal para a Distribuição Binomial

- Quanto mais perto p estiver de 0,5, melhor será a aproximação da normal para a distribuição binomial
- Quanto maior o tamanho da amostra, n, melhor será a aproximação da normal para a distribuição binomial
- Regra Geral:
  - A distribuição normal pode ser utilizada para aproximar a distribuição binomial se

$$np \geq 5 \text{ e } n(1-p) \geq 5$$

SlidesBioestatística6 (ua.pt)

Aproximações Baseadas no TLC: Poisson - Normal

- Probabilidades associadas a uma distribuição de Poisson, P(λ), podem ser aproximadas utilizando uma distribuição Normal, N(μ,σ), com $\mu = \lambda$ e $\sigma = \sqrt{\lambda}$.

  A aproximação será tanto melhor quanto maior for λ.

SlidesBioestatística6 (ua.pt)

TLC: Correção de Continuidade

A variante do T. L. C. para distribuições discretas introduz a correção de continuidade que permite aproximar uma distribuição discreta por uma distribuição contínua, neste caso a distribuição Normal.

Correção de continuidade: Seja X uma v. a. discreta, com variação Δ (i.e., X pode tomar os valores $0, \Delta, 2\Delta, 3\Delta, \ldots$), com $E(X) = \mu_X$ e $Var(X) = \sigma_X^2$. Então a correção de continuidade é efetuada da seguinte forma:

$$P(X = k) \approx P\left(k - \frac{\Delta}{2} \leq X \leq k + \frac{\Delta}{2}\right).$$

ProbabilidadesEstatistica2019.pdf

TLC: Correção de Continuidade

Note-se que nas distribuições usuais, por exemplo, Binomial e Poisson, $\Delta = 1$, pois tratam-se de distribuições de contagem logo

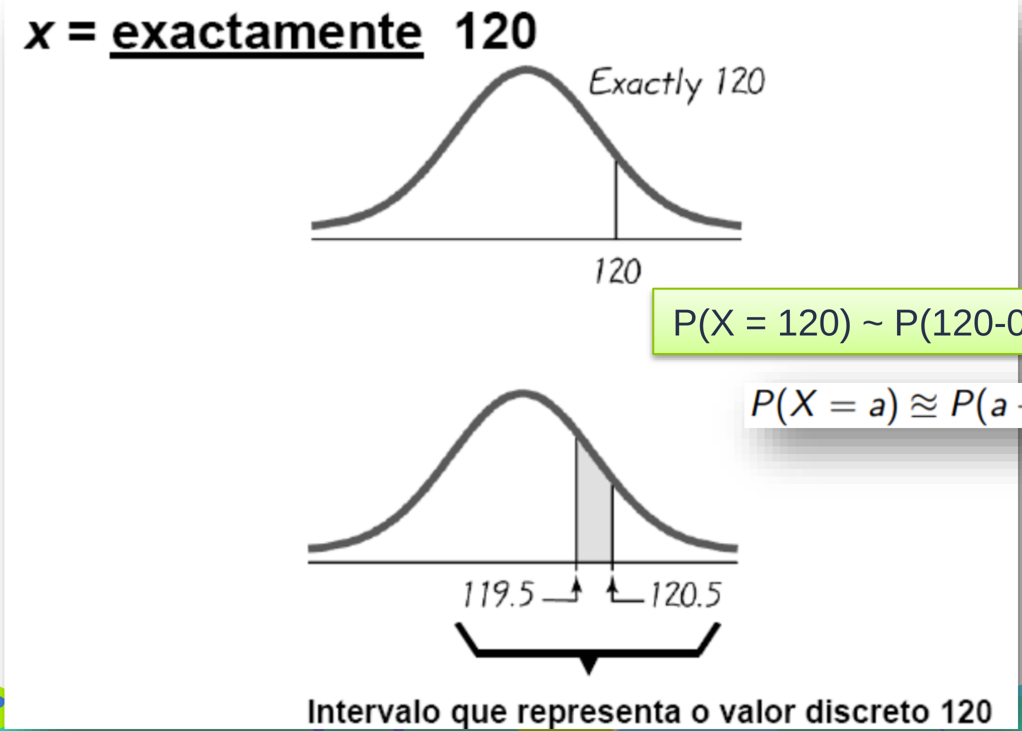
$$P(X = k) \approx P(k - 0{,}5 \leq X < k + 0{,}5).$$

Com a aplicação das propriedades da distribuição Normal, podem generalizar-se as seguintes regras:

- $P(X \leq k) \approx P\left(Z \leq \frac{k + 0{,}5 - \mu_X}{\sigma_X}\right) = \Phi\left(\frac{k + 0{,}5 - \mu_X}{\sigma_X}\right);$
- $P(X < k) \approx P\left(Z < \frac{k - 0{,}5 - \mu_X}{\sigma_X}\right) = \Phi\left(\frac{k - 0{,}5 - \mu_X}{\sigma_X}\right);$
- $P(X \geq k) \approx P\left(Z \geq \frac{k - 0{,}5 - \mu_X}{\sigma_X}\right) = 1 - \Phi\left(\frac{k - 0{,}5 - \mu_X}{\sigma_X}\right);$
- $P(X > k) \approx P\left(Z > \frac{k + 0{,}5 - \mu_X}{\sigma_X}\right) = 1 - \Phi\left(\frac{k + 0{,}5 - \mu_X}{\sigma_X}\right).$

ProbabilidadesEstatistica2019.pdf

Aproximações Baseadas no TLC: Correção de Continuidade



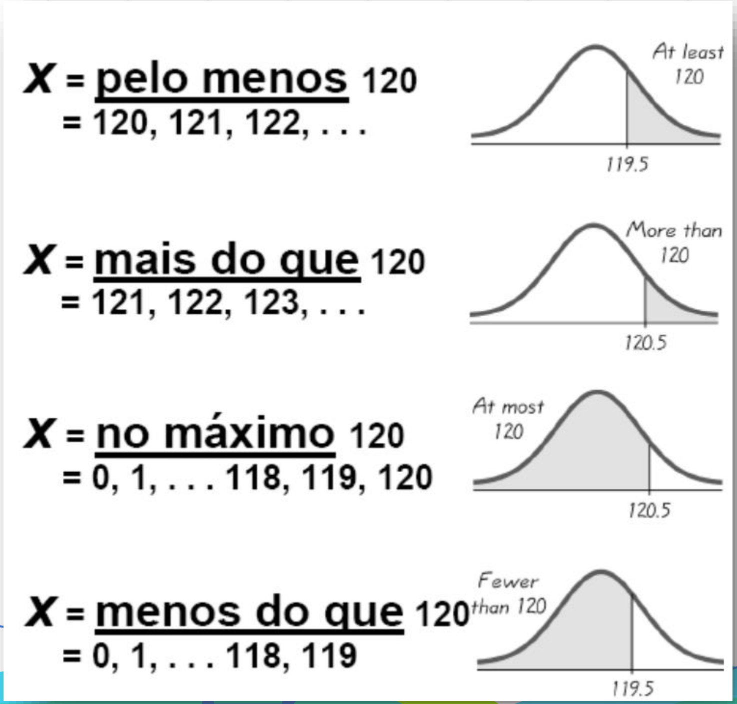
Note-se que $\tilde{X}$ é uma v.a. Normal que "aproxima" X (sendo esta uma v.a. discreta).

P(X = 120) ~ P(120-0,5 < $\tilde{X}$ < 120+0,5)

$$P(X = a) \approxeq P(a - \delta < \tilde{X} < a + \delta)$$

SlidesBioestatística6 (ua.pt)

Aproximações Baseadas no TLC: Correção de Continuidade



P(X ≥ 120) ~ P($\tilde{X}$ ≥ 120-0,5)

P(X > 120) ~ P($\tilde{X}$ > 120+0,5)

P(X ≤ 120) ~ P($\tilde{X}$ ≤ 120+0,5)

P(X < 120) ~ P($\tilde{X}$ < 120-0,5)

Note-se que $\tilde{X}$ é uma v.a. Normal que “aproxima” X (sendo esta uma v.a. discreta).

SlidesBioestatística6 (ua.pt)

TLC: Resumo da Correção de Continuidade

Observação - aplicação do TLC para distribuições discretas

- Importante: A utilização do teorema do limite central para **distribuições discretas** apresenta um problema, uma vez que se vai aproximar um fenómeno discreto por uma distribuição contínua. Embora não exista uma solução ótima para todas as situações, na prática é comum adotar a chamada correção de continuidade. Considere-se $0 < \delta < 1$ e $\tilde{X}$ a v.a. normal que "aproxima" X, tem-se:
  - $P(X = a) \approxeq P(a - \delta < \tilde{X} < a + \delta)$
  - $P(a < X < b) \approxeq P(a + \delta \leq \tilde{X} \leq b - \delta)$
  - $P(X \leq a) \approxeq P(\tilde{X} \leq a + \delta)$
  - $P(X < a) \approxeq P(\tilde{X} < a - \delta) = P(\tilde{X} \leq a - \delta)$ (recordar que $\tilde{X}$ é contínua...)
- No caso da aplicação do TLC à binomial e à Poisson o valor típico é $\delta = 0.5$, ou seja metade da variação entre dois valores consecutivos.

Slides da Professora Conceição Amado

TLC

Formulário

TEOREMA DO LIMITE CENTRAL E COROLÁRIOS

TLC: Sendo X_i iid com $E(X_i) = \mu$ e $\mathrm{Var}(X_i) = \sigma^2 \Rightarrow \dfrac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$

Corolário: Sendo $X_i \sim B(1;\theta)$, iid $\Rightarrow \dfrac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1)$

Correcção de continuidade: $P(a \le X \le b) \approx \Phi\left(\dfrac{b + \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\dfrac{a - \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right)$, com a e b inteiros

Corolário: Sendo $X \sim \mathrm{Po}(\lambda)$, quando $\lambda \to +\infty \Rightarrow \dfrac{X - \lambda}{\sqrt{\lambda}} \overset{a}{\sim} N(0,1)$

Correcção de continuidade: $P(a \le X \le b) \approx \Phi\left(\dfrac{b + \frac{1}{2} - \lambda}{\sqrt{\lambda}}\right) - \Phi\left(\dfrac{a - \frac{1}{2} - \lambda}{\sqrt{\lambda}}\right)$, com a e b inteiros

Exemplo 1: TLC

Exemplo 1- aplicação TLC

Uma empresa de chocolates embala caixas com 100 pacotes de bombons. Os pesos por pacote são v.a.'s X_i, onde $X_i \perp\!\!\!\perp X_j$, com valor médio 0.5kg e variância 0.1kg^2. São colocadas 10 caixas numa palete. Qual é a probabilidade aproximada do peso dos bombons da palete ser superior a 510kg?

Slides da Professora Conceição Amado

Exemplo 1: TLC

Resolução:

$X_i -$ 'v.a. peso do $i-$ésimo pacote (kg)'

$E(X_i) = 0.5,\ V(X_i) = 0.1\ , \forall i$ e $X_i \perp\!\!\!\perp X_j$ para $i \neq j$

$S = \sum_{i=1}^{1000} X_i$ 'v.a. peso dos bombons colocados na palete (kg)'

$$E(S) = E\left(\sum_{i=1}^{1000} X_i\right) = \sum_{i=1}^{1000} E(X_i) = 1000E(X_1) = 1000 \times 0.5 = 500$$

(porque são identicamente distribuídas (i.d.) a X_1)

$$V(S) = V\left(\sum_{i=1}^{1000} X_i\right) = \sum_{i=1}^{1000} V(X_i) = 1000V(X_1) = 1000 \times 0.1 = 100$$

porque são independentes e identicamente distribuídas (i.d.) a X_1.

Slides da Professora Conceição Amado

Exemplo 1: TLC

Resolução (cont.):

Pelo T.L.C. tem-se que:

$$\frac{S - E(S)}{\sqrt{V(S)}} = \frac{S - 500}{\sqrt{100}} \overset{a}{\sim} N(0,1),$$

$$P(S > 510) = 1 - P(S \leq 510) = 1 - P\left(\frac{S-500}{\sqrt{100}} \leq \frac{510-500}{\sqrt{100}}\right) \underset{T.L.C.}{\approx}$$

$$\underset{T.L.C.}{\approx} 1 - \Phi(1) \approx 0.1587.$$

Slides da Professora Conceição Amado

Exemplo 2: TLC

Exemplo 2 - aplicação TLC e correção de continuidade

O número de chamadas de telemóvel registadas a partir de certa "zona" numa hora tem, em condições estacionárias, distribuição de Poisson de parâmetro 1500. Calcule a probabilidade (aproximada) de ocorrerem mais de 1600 chamadas na próxima hora.

Slides da Professora Conceição Amado

Exemplo 2: TLC

Resolução:
$X \sim Poisson(1500)$, como o valor de λ é elevado não consta nas tabelas disponibilizadas, e sem a ajuda de um computador não é possível calcular a probabilidade pedida. Mas, podemos usar a distribuição normal para obter uma aproximação.
Como $X \sim Poisson(1500)$ pode ser aproximada por $\tilde{X} \sim N(1500, 1500)$

$$P(X > 1600) = 1 - P(X \leq 1600) \simeq 1 - P(\tilde{X} \leq 1600 + 0.5) =$$
$$= 1 - P(\tilde{X} \leq 1600.5) = 1 - P\left(\frac{\tilde{X} - 1500}{\sqrt{1500}} \leq \frac{1600.5 - 1500}{\sqrt{1500}}\right) =$$
$$= 1 - \Phi(2.59) \simeq 0.0048$$

Nota:
Lambda = 1500 > 5
Logo pode-se aproximar a distribuição Poisson à distribuição Normal

Slides da Professora Conceição Amado

Teorema do Limite Central

Formulário

TEOREMA DO LIMITE CENTRAL E COROLÁRIOS

TLC: Sendo X_i iid com $E(X_i)=\mu$ e $\mathrm{Var}(X_i)=\sigma^2 \Rightarrow \dfrac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\,\sigma} = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$

Corolário: Sendo $X_i \sim B(1;\theta)$, iid $\Rightarrow \dfrac{\sum_{i=1}^n X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1)$

Correcção de continuidade: $P(a \le X \le b) \approx \Phi\left(\dfrac{b+\frac{1}{2}-n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\dfrac{a-\frac{1}{2}-n\theta}{\sqrt{n\theta(1-\theta)}}\right)$, com a e b inteiros

Corolário: Sendo $X \sim \mathrm{Po}(\lambda)$, quando $\lambda \to +\infty \Rightarrow \dfrac{X-\lambda}{\sqrt{\lambda}} \overset{a}{\sim} N(0,1)$

Correcção de continuidade: $P(a \le X \le b) \approx \Phi\left(\dfrac{b+\frac{1}{2}-\lambda}{\sqrt{\lambda}}\right) - \Phi\left(\dfrac{a-\frac{1}{2}-\lambda}{\sqrt{\lambda}}\right)$, com a e b inteiros

Obrigada!

Questões?