Estatística II

Aulas 6 & 7

Nelson Faustino [1]

[1]**Faculdade de Economia (FEUC)**
Universidade de Coimbra – Portugal

nelson@fe.uc.pt

Conteúdos

Função **qexp**

Síntaxe

Comandos R para calcular função distribuição exponencial negativa:

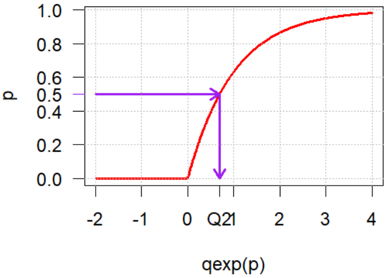
- **qexp(p)** – Calcula, por defeito, o quantil de ordem p (q) a partir de $P(T \leq q) = p$ $(T \sim Ex(1))$;
- **qexp(p,lower.tail = FALSE)** – Calcula, por defeito, o quantil de ordem p (q) a partir de $P(T > q_p) = 1 - p$ $(T \sim Ex(1))$;
- **pexp(q,rate)** – Calcula o quantil de ordem p (q)a partir de $P(T \leq q) = q$ $(\lambda =$**rate**$)$ para $T \sim Ex(\lambda)$.
- **pexp(t,rate,lower.tail = FALSE)** – Calcula o quantil de ordem p (q) a partir de $P(T > q) = 1 - p$ $(\lambda =$**rate**$)$ para $T \sim Ex(\lambda)$.

Distribuição Exponencial Negativa Estandartizada

Representação Gráfica dos Quantis de ordem α ($q_\alpha = -\ln(1-\alpha)$)

Mediana em R

```
> qexp(0.5)
[1] 0.6931472
> -log(0.5)
[1] 0.6931472
```

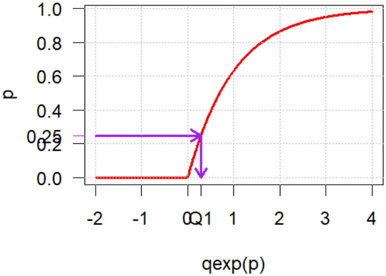


Distribuição Exponencial Negativa Estandartizada

Representação Gráfica dos Quantis de ordem α ($q_\alpha = -\ln(1-\alpha)$)

1^o quartil em R

```
> qexp(0.25)
[1] 0.2876821
> -log(0.75)
[1] 0.2876821
```

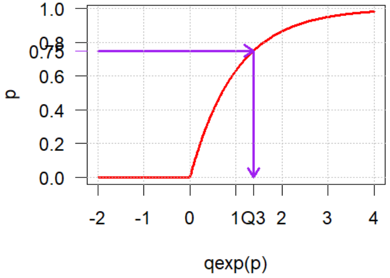


Distribuição Exponencial Negativa Estandartizada

Representação Gráfica dos Quantis de ordem α ($q_\alpha = -\ln(1-\alpha)$)

3^o quartil em R

```
> qexp(0.75)
[1] 1.386294
> -log(0.25)
[1] 1.386294
```



Distribuição Exponencial Negativa

Função de distribuição et al.

Função de Distribuição ($X \sim Ex(\lambda)$)

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

Função densidade de probabilidade ($f(x) = F'(x)$)

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (x > 0) \\ 0 & (\text{outros } x) \end{cases}$$

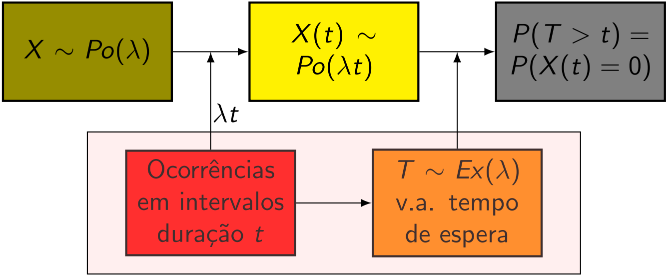
Propriedades de Interesse

$P(X > h) = e^{-\lambda h}$ $(h > 0)$ tem as seguintes interpretações:

1. **Integral Impróprio:** $\int_h^{+\infty} \lambda e^{-\lambda x} dx = e^{-\lambda h}$.
2. **Falta de Memória:** $P(X > x + h | X > x) = P(X > h)$.
3. **Processo de Poisson:** $e^{-\lambda h} = e^{-\lambda h} \frac{(\lambda h)^0}{0!}$ $(x = 0)$.

Exponencial Negativa vs. Poisson

Génese da distribuição exponencial a partir da distribuição de Poisson



- $F(t) = P(T \leq t)$ **para** $T \sim Ex(\lambda)$ – dá-nos a probabilidade do **tempo de espera pela chegada do primeiro acontecimento em** $]0, t]$.
- $P(T > t) = P(X(t) = 0)$ – dá-nos probabilidade de **não haver qualquer chegada** no intervalo $]0, t]$.

Exemplo 5.19 (Poisson vs. Exponencial)

- O parâmetro média de clientes por hora ($\lambda = 20$) é determinado a partir das **propriedades de processos de Poisson** ($E(X) = V(X) = \lambda$).
- A probabilidade calculada coincide com a probabilidade de não ter atendido clientes nos primeiros 5 minutos de abertura da loja.
- A propriedade de falta de memória permite-nos p.e. concluir que a probabilidade de o comerciante ter de esperar mais de 12 minutos pelo cliente, supondo que já está aguardando à mais de 7 minutos **coincide com a probabilidade calculada neste exemplo (porquê?)**.

Distribuição Exponencial Negativa

Exemplos de [Murteira et al. (2015), pp. 285–286]

Exemplo 5.20 (Falta de Memória)

- No **Exemplo 5.20**, o parâmetro λ de $X \sim Ex(\lambda)$ é determinado a partir da equivalência $E(X) = \frac{1}{\lambda} \Leftrightarrow \lambda = \frac{1}{E(X)}$.
- No contexto do exercício, a propriedade de 'falta de memória' permite-nos assegurar que $P(X > x + 700|X > x) = P(X > 1100|X > 400)$, para todo o $x > 0$ (porquê?).
- Com recurso a **processos de Poisson**, a probabilidade calculada – $P(X > 700)$ permite-nos uma possível interpretação alternativa: componente electrónica **não registou qualquer avaria** nas primeiras 700 horas de funcionamento.

FEUC FACULDADE DE ECONOMIA UNIVERSIDADE DE COIMBRA

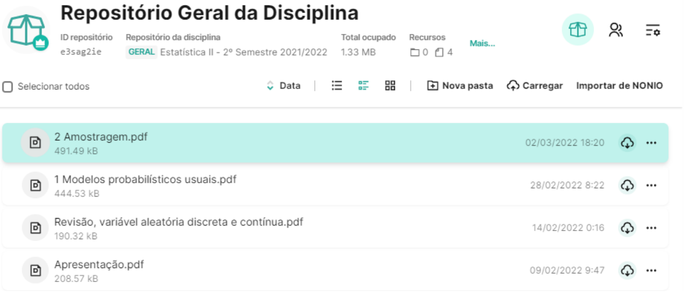
Conteúdos

Amostragem

Leituras Recomendadas

Complemento dos Slides

- **Páginas 347 a 365** de [Murteira et al. (2015), Capítulo 6].
- **Páginas 1 a 15** do ficheiro `2 Amostragem.pdf`, disponível no **Repositório Geral da Disciplina (UC Student).**



Amostra Casual – vide [Murteira et al. (2015), Definição 6.1]

O vector aleatório $(X_1, X_2, \ldots, X_n)$ define uma **amostra casual** quando as v.a.'s X_i $(i = 1, 2, \ldots, n)$ são **independentes e identicamente distribuídas (i.i.d.)**.

$X_1, X_2, \ldots, X_n$ são:

- **Independentes** quando $P(X_i \leq x_i \wedge X_j \leq x_j) = P(X_i \leq x_i).P(X_j \leq x_j)$, para todo o $i \neq j$;
- **Identicamente distribuídas** quando correspondem a n 'cópias' [idênticas] da v.a. $X \sim F$, i.e.

$$P(X_i \leq x_i) = F(x_i), \text{ para todo o } i = 1, 2, \ldots, n.$$

Distribuição conjunta da amostra casual

$$F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)=F(x_1)F(x_2)\ldots F(x_n),$$

onde $F(x_i)=P(X\leq x_i)$ $(i=1,2,\ldots n)$ denota a **função distribuição de probabilidade** da população $X\sim F$.

Função [densidade] de probabilidade da amostra casual

$$f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)=f(x_1)f(x_2)\ldots f(x_n),$$

onde f denota a **função [densidade] de probabilidade** da população $X\sim F$.

Alguns exemplos de estatísticas que iremos considerar a partir de uma amostra casual $(X_1, X_2, \ldots, X_n)$ de $X \sim F$.

- **Mínimo da Amostra:** $X_{(1)} = \min\limits_{1 \leq i \leq n} (X_i)$
- **Máximo da Amostra:** $X_{(n)} = \max\limits_{1 \leq i \leq n} (X_i)$
- **Média da amostra:** $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$
- **Variância da amostra:** $S^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2$
- **Variância corrigida da amostra:** $(S')^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$

Variáveis Aleatórias Independentes

vide [Murteira et al. (2015), pp. 362-363]

Mínimo da Amostra

$$X_{(1)} = \min_{1\leq i\leq n} (X_i)$$

Máximo da Amostra

$$X_{(n)} = \max_{1\leq i\leq n} (X_i)$$

Probabilidade Envolvendo Mínimo da Amostra

$$X_{(1)} > x \overset{\text{ordenação}}{\iff} X_i > x, \quad \forall_{1\leq i\leq n} \overset{i.i.d.'s}{\Longrightarrow} P(X_{(1)} > x) = (P(X > x))^n$$

Probabilidade Envolvendo Máximo da Amostra

$$X_{(n)} \leq x \overset{\text{ordenação}}{\iff} X_i \leq x, \quad \forall_{1\leq i\leq n} \overset{i.i.d.'s}{\Longrightarrow} P(X_{(n)} \leq x) = (P(X \leq x))^n$$

Variáveis Aleatórias Independentes

Média e Variância

Média da População

$E(b) = b$

$E(aX+b) = aE(X)+b \; (a, b \in \mathbb{R})$

Variância da população

$V(b) = 0$

$V(aX+b) = a^2 V(X) \; (a, b \in \mathbb{R})$

Média v.a.'s independentes

$$Y = \sum_{i=1}^{n} a_i X_i \overset{\text{linearidade}}{\Longrightarrow} E(Y) = \sum_{i=1}^{n} a_i E(X_i)$$

Variância v.a.'s independentes

$$Y = \sum_{i=1}^{n} a_i X_i \overset{\text{independência das v.a.'s}}{\Longrightarrow} V(Y) = \sum_{i=1}^{n} a_i^2 V(X_i)$$

Independência de variáveis aleatórias

Reformulação de resultados de [Murteira et al. (2015), Capítulo 5]

Soma de v.a.'s i.i.d.'s

$$\sum_{i=1}^{n} X_i = n\overline{X}$$

Reformulação dos resultados

- **Binomial:** Teorema 5.1
- **Binomial Negativa:** Teorema 5.2
- **Poisson:** Teorema 5.3
- **Normal:** Corolário 5.1

Propriedades de distribuições envolvendo amostras casuais

- **Binomial:** $X_i \sim B(k, \theta) \Longrightarrow n\overline{X} \sim B(nk, \theta)$
- **Binomial Negativa:** $X_i \sim BN(r, \theta) \Longrightarrow n\overline{X} \sim BN(nr, \theta)$
- **Poisson:** $X_i \sim Po(\lambda) \Longrightarrow n\overline{X} \sim Po(n\lambda)$
- **Normal:** $X_i \sim N(\mu, \sigma^2) \Longrightarrow n\overline{X} \sim N(n\mu, n\sigma^2)$

População Normal $N(\mu, \sigma^2)$

Vide [Murteira et al. (2015), Corolário 5.2]

Distribuição da média da amostra

$$(X_1, \ldots X_n) \longrightarrow n\overline{X} \sim N(n\mu, n\sigma^2) \xrightarrow[\frac{\sigma^2}{n}]{\mu} \overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Obs:

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{n\overline{X} - n\mu}{\sqrt{n\sigma^2}} \sim N(0,1)$$

Média e Variância da População – vide [Murteira et al. (2015), Teorema 6.1 & Teorema 6.2]

Se $(X_1, X_2, \ldots, X_n)$ é uma **amostra casual** de uma população para a qual existem média $\mu = E(X_i)$ e variância $\sigma^2 = V(X_i)$ $(i = 1, 2, \ldots, n)$, tem-se:

- **Valor esperado da média da amostra:** $E(\overline{X}) = \mu$;
- **Variância da média da amostra:** $V(\overline{X}) = \dfrac{\sigma^2}{n}$;
- **Valor esperado da variância da amostra:** $E(S^2) = \dfrac{n-1}{n}\sigma^2$.

Variância Corrigida da Amostra

$$
\begin{aligned}
(S')^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2 \\
&= \frac{n}{n-1}\,\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2 \\
&= \frac{n}{n-1}S^2
\end{aligned}
$$

$$
\begin{aligned}
\Longrightarrow E((S')^2) &= \frac{n}{n-1}\overbrace{E(S^2)}^{=\frac{n-1}{n}\sigma^2} \\
&= \frac{\cancel{n}}{\cancel{n-1}}\,\frac{\cancel{n-1}}{\cancel{n}}\sigma^2 = \sigma^2
\end{aligned}
$$

Conteúdos

Exercícios Resolvidos

Exercício 50., página 337

Dados do Exercício 50.

Distribuição de Probabilidade: $X(t) \sim Po(2)$, sendo $\lambda t = 2$ a média, $t = 3$ minutos a unidade de tempo ('*em média dois clientes por cada 3 minutos*').

Resolução 50.a)

Usando a fórmula da função de probabilidade de $X(3) \sim Po(2)$, segue que $P(X(3) = 0) = e^{-2}\frac{2^0}{0!} = e^{-2}$ dá-nos o valor exacto da probabilidade pedida ('*3 minutos sem qualquer cliente atendido*').

50.a) via `> dpois(x,lambda)`

```
> dpois(0,2)
[1] 0.1353353
> exp(-2)*2^0/factorial(0)
[1] 0.1353353
```

Formulação 50.b)

Para resolvermos o item **50.b)**, precisamos de considerar variável aleatória $T \sim Ex(2/3)$ para descrever o **número de atendimentos consecutivos por minuto**, onde $\lambda = 2/3$ – calculada a partir dos dados do enunciado – dá-nos a média de atendimentos por minuto.

Resolução 50.b) usando função de distribuição de $T \sim Ex(2/3)$

Sendo $F(t) = 1 - e^{-\frac{2t}{3}}$ $(t \geq 0)$ a função de distribuição de $T \sim Ex(2/3)$, segue que

$$P(T > 3) = 1 - \overbrace{P(T \leq 3)}^{=F(3)} = \cancel{1} - (\cancel{1} - e^{-\frac{2.\cancel{3}}{\cancel{3}}}) = e^{-2}.$$

50.b) via `> pexp(x,rate)` **&** `> pexp(x,rate,lower.tail=FALSE)`

```
> 1-pexp(3,2/3)
[1] 0.1353353
> pexp(3,2/3,lower.tail = FALSE)
[1] 0.1353353
```

Resolução 50.c)

Verificámos que a probabibilidade do tempo de espera demorar mais que 3 minutos **coincidir** com a probabilidade de ocorrer 3 minutos sem qualquer atendimento, i.e.

$$P(T > 3) = e^{-2} = P(X(3) = 0).$$

Resolução 50.d)

Teremos mais uma vez de considerar a função distribuição de probabilidade, $F(t) = 1 - e^{-\frac{2t}{3}}$ $(t \geq 0)$ da variável aleatória $T \sim Ex(2/3)$ concluir que a probabilidade pedida – atendimento de um cliente demorar entre 3 a 6 minutos – é calculada por

$$\begin{aligned} P(3 \leq T \leq 6) &= \overbrace{P(T \leq 6)}^{=F(6)} - \overbrace{P(T < 3)}^{=F(3)} \\ &= (\cancel{1} - \overbrace{e^{-\frac{2.6}{3}}}^{=e^{-4}}) - (\cancel{1} - e^{-\frac{2.\cancel{3}}{\cancel{3}}}) \\ &= e^{-2} - e^{-4}. \end{aligned}$$

50.d) via `> pexp(x,rate)`

```
> pexp(6,2/3)-pexp(3,2/3)
[1] 0.1170196
> exp(-2)-exp(-4)
[1] 0.1170196
```

Observação 50.

Se, ao invés tivéssemos considerado a v.a. $T^* \sim Ex(2)$ para descrever o **número de atendimentos consecutivos em intervalos de 3 minutos**, os comandos *R* abaixo produzem soluções equivalentes para **50.b)** & **50.d)**, respectivamente:

	Probabilidade a calcular	Comando R
50.b)	$P(T^* > 1)$	$> 1 - \mathtt{pexp}(1,2)$
50.d)	$P(1 < T^* < 2)$	$> \mathtt{pexp}(2,2) - \mathtt{pexp}(1,2)$

Exercícios Resolvidos

Exercício 51., página 337

Dados do Exercício 51.

Distribuição de Probabilidade: $X(t) \sim Po(2t)$, sendo $\lambda = 2$ a média de aviões por hora.

Formulação 51.a)

Considerando a v.a. $X(1/4) \sim Po(1/2)$ para contabilizar o **número de chegadas em intervalos de 15 minutos (1/4 hora)**, segue que a v.a $T \sim Ex(1/2)$ dá-nos o **tempo de espera entre duas aterragens consecutivas em intervalos de 15 minutos.**

Resolução 51.a)

Da propriedade $P(T > 1/4) = P(X(1/4) = 0)$ – probabilidade do tempo de chegada ser superior a 1/4 hora coincidir com a probabilidade de não haver qualquer chegada em intervalo de tempo $]0, 1/4]$ – segue que $P(T < 1/4) = 1 - e^{-1/2}$ **[Porquê?]**.

Exercícios Resolvidos

Exercício 51., página 337

Formulação 51.b)

Considerando a v.a. $X(1/2) \sim Po(1)$ para contabilizar o **número de chegadas em intervalos de 30 minutos (1/2 hora)**, segue que a v.a $T^* \sim Ex(1/2)$ dá-nos o **tempo de espera entre duas aterragens consecutivas em intervalos de 30 minutos.**

Resolução 51.b) via propriedade $P(T^* > 1/2) = P(X(1/2) = 0)$

Da propriedade $P(T^* > 1/2) = P(X(1/2) = 0)$ – probabilidade do tempo de chegada ser superior a 1/2 hora coincidir com a probabilidade de não haver qualquer chegada em intervalo de tempo $]0, 1/2]$ – segue que $P(T^* > 1/2) = e^{-1}$ **[Porquê?]**.

Exercícios Resolvidos

Exercício 51, página 337

Resolução 51. em R

	Probabilidade a calcular	Comando R
51.a)	$P(T < 1/4)$	$>$ pexp(1/4, 2)
	$1 - P(X(4) = 0)$	$>$ 1 − dpois(0, 1/2)
51.b)	$P(T^* > 1/2)$	$>$ 1 − pexp(1/2, 2)
	$P(X(1/2) = 0)$	$>$ dpois(0, 1)

Resolução 51. – comandos alternativos

```
> ppois(0,1/2,lower.tail=FALSE)
[1] 0.3934693
> pexp(1/2,2,lower.tail = FALSE)
[1] 0.3678794
```

Exercícios Resolvidos

Exercício 52., página 337

Variável Aleatória 52.

Sendo $E(X) = 1000$ horas a duração média de vida de uma lâmpada, segue que $X \sim Ex(1/1000)$ descreve **tempo de vida de uma lâmpada** (usei a propriedade $E(X) = \frac{1}{\lambda}$ de $X \sim Ex(\lambda)$).

O que pretendemos calcular em 52.?

Probabilidade $P(Y < 50)$, onde a v.a. Y é construída da seguinte forma:

1. Começo por considerar 12 **v.a.'s independentes e identicamente distribuídas (i.i.d.'s)** $X_i \sim Ex(1/1000)$ **(**$i = 1, 2, \ldots, 12$**)** para descrever o tempo de vida de cada lâmpada da caixa com 12 lâmpadas;
2. Defino a v.a. Y como $Y = \min\{X_1, X_2, \ldots, X_{12}\}$ para descrever a **lâmpada com menor tempo duração na caixa de 12**.

Exercícios Resolvidos

Exercício 52., página 337

Resolução 52.

Do facto de $X_1, X_2, \ldots, X_{12}$ serem i.i.d.'s, e da equivalência

$$\min\{X_1, X_2, \ldots, X_{12}\} \geq y \iff X_1 \geq y \ \wedge \ X_2 \geq y \ \wedge \ \ldots \ \wedge \ X_{12} \geq y$$

segue que

$$\begin{aligned} P(Y \geq y) &= P(X_1 \geq y).P(X_2 \geq y) \ldots P(X_{12} \geq y) \\ &= (P(X \geq y))^{12} \quad \textbf{[Porquê?]} \end{aligned}$$

Logo $P(Y < 50) = 1 - (P(X \geq 50))^{12} = 1 - e^{-\frac{12 \times 50}{1000}}$ dá-nos a probabilidade pedida.

Resolução 52. em R

```
> 1-(1-pexp(50,1/1000))^12
[1] 0.4511884
```

Estatísticas de Interesse

Mínimo ($X_{(1)}$) e Máximo ($X_{(n)}$) da Amostra

$$X \to (X_1, X_2, \ldots, X_n)$$

$$X_{(1)} = \min_{1 \leq i \leq n} (X_i)$$

$$P(X_{(1)} > x) \;\text{—}\; (1 - F(x))^n$$

$$P(X_{(1)} \leq x) \;\text{–}\; 1 - (1 - F(x))^n$$

$$X_{(n)} = \max_{1 \leq i \leq n} (X_i)$$

$$P(X_{(n)} > x) \;\text{—}\; 1 - (F(x))^n$$

$$P(X_{(n)} \leq x) \;\text{—}\; (F(x))^n$$

Estatísticas de Interesse

Média ($\overline{X}$) e Variância (S^2) da amostra

$$X \rightarrow (X_1, X_2, \ldots, X_n)$$

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

$$\mu = E(X) \quad \text{——} \quad E(\overline{X}) = \mu$$

$$\sigma^2 = V(X) \quad \text{——} \quad V(\overline{X}) = \frac{\sigma^2}{n}$$

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

$$\sigma^2 = V(X) \quad \text{—} \quad E(S^2) = \frac{n-1}{n}\sigma^2$$

$$E(X^k) \ (k = 2, 4) \quad V(S^2) = \ldots$$

Exercícios Resolvidos

Exercício 2., página 388

Dados do Exercício 2.

Distribuição de Probabilidade da População: $X \sim Po(0.3)$ corresponde ao número de gralhas de uma página.

Probabilidade a calcular em 2.a)

A probabilidade a calcular a partir da amostra casual $(X_1, X_2, X_3, X_4, X_5)$ corresponde a

$$P(X_1 = 1 \wedge X_2 = 1, X_3 = 0 \wedge X_4 = 0 \wedge X_5 = 0) \overset{i.i.d.'s}{=} (f(1))^2 (f(0))^3,$$

sendo $f(x) = P(X = x)$ $(X \sim Po(0.3))$.

2.a) via `> dpois(x,lambda)`

```
> (dpois(1,0.3))^2*(dpois(0,0.3))^3
[1] 0.02008171
```

Exercícios Resolvidos

Exercício 2., página 388

Formulação de 2.b)

Amostra casual: Temos agora uma amostra casual de dimensão 20 de $X \sim Po(0.3)$. Neste caso, usando a **estatística média da amostra casual** ($\overline{X}$), podemos dizer que o número total de gralhas em 20 páginas pode ser descrito pela v.a. $Y = 20\overline{X}$.

Probabilidade a calcular em 2.b)

A probabilidade pedida corresponde a

$$P(20\overline{X} \geq 8) = 1 - P(20\overline{X} \leq 7),$$

onde $20\overline{X} \sim Po(20 \times 0.3) = Po(6)$.

2.a) via > `ppois(x,lambda)`

```
> ppois(1,0.3)^5
[1] 0.8284667
```

Formulação de 2.c)

Amostra casual: A amostra casual passou a ser agora de dimensão 50 de $X \sim Po(0.3)$.

Resolução de 2.c)

Com base nas fórmulas

$$E(\overline{X}) = E(X) \ \& \ V(\overline{X}) = \frac{V(X)}{50} \ (n = 50),$$

e nas igualdades $E(X) = V(X) = 0.3$, retiramos que

$$E(\overline{X}) = 0.3 \quad \wedge \quad V(\overline{X}) = 0.006$$

Exercícios Resolvidos

Exercício 2., página 388

Formulação de 2.c)

Amostra casual: A amostra casual passou a ser agora de dimensão 50 de $X \sim Po(0.3)$.

Resolução de 2.c)

Com base nas fórmulas

$$E(\overline{X}) = E(X) \text{ \& } V(\overline{X}) = \frac{V(X)}{50} \ (n = 50),$$

e nas igualdades $E(X) = V(X) = 0.3$, retiramos que

$$E(\overline{X}) = 0.3 \quad \wedge \quad V(\overline{X}) = 0.006$$

Formulação de 2.d)

Estatística máximo da amostra: Note que:

$$\max(X_i) \leq 1 \iff X_1 \leq 1 \ \wedge \ X_2 \leq 1 \ \wedge \ X_3 \leq 1 \ \wedge \ X_4 \leq 1 \ \wedge \ X_5 \leq 1$$

Probabilidade a calcular em 2.d)

Da equivalência acima segue que

$$P(\max(X_i) \leq 1) \overset{i.i.d's}{=} (P(X \leq 1))^5$$ [**porquê?**]

2.d) via `> ppois(x,lambda)`

```
> 1-ppois(7,6)
[1] 0.2560202
```

Exercícios Resolvidos

Exercício 2., página 388

Formulação de 2.e)

Temos que considerar a v.a. $Y \sim B(n, \theta)$, sendo $n = 100$ e $\theta = P(X = 0)$ $(X \sim Po(0.3))$

Probabilidade a calcular em 2.e)

A probabilidade a calcular – $P(Y \geq 80) = 1 - P(Y \leq 79)$ – pode ser realizada segundo o seguinte procedimento

	Probabilidade a calcular	Comando R
Passo 1	$\theta = P(X = 0)$	`> theta <-dpois(0, 0.3)`
Passo 2	$P(Y \geq 80)$	`> 1 - pbinom(79, 100, theta)`
Output		[1] 0.1060093

Exercícios Resolvidos

Exercício 6., página 389

Formulação de 6.

População: $X \sim Ex(1/15)$ de média $E(X) = 15$.
Estatísticas de interesse: $X_{(1)}$ resp. $X_{(5)}$ para descrever o menor resp. maior dos atrasos.

Probabilidades a calcular em 6.

Com base na ordenação abaixo

$$X_{(1)} \geq 18 \Longleftrightarrow X_1 \geq 18 \wedge X_2 \geq 18 \wedge X_3 \geq 18 \wedge X_4 \geq 18 \wedge X_5 \geq 18$$
$$X_{(5)} \leq 30 \Longleftrightarrow X_1 \leq 30 \wedge X_2 \leq 30 \wedge X_3 \leq 30 \wedge X_4 \leq 30 \wedge X_5 \leq 30$$

e após algumas simplificações:

Probabilidade	F. distribuição	Comando R
$P(X_{(1)} < 18)$	$1-(1-F(18))^5$	`> 1 - (1 - pexp(18, 1/15))^5`
$P(X_{(5)} > 30)$	$1-(F(30))^5$	`> 1 - pexp(30, 1/15)^5`

Exercícios Resolvidos

Exercício 12., página 389

Formulação de 6.

População: $X \sim N(\mu, \sigma^2)$ com $\mu = 800$ e $\sigma = 100$
Estatísticas de interesse:

- **a)** $X_{(1)}$ para descrever a primeira lâmpada que deixa de funcionar.
- **b)** $X_{(4)}$ para descrever a última lâmpada que deixa de funcionar.
- **c)** $\overline{X} \sim N(800, (100)^2/4)$ para descrever a duração média da amostra.

Probabilidades a calcular em 12.

	Probabilidade a calcular	após simplificações
a)	$P(X_{(1)} > 900)$	$(P(X > 900))^4$
b)	$P(X_{(4)} > 1000)$	$1 - (P(X \leq 1000))^4$
c)	$P(\lvert\overline{X} - 800\rvert > 100)$	$P\left(\overline{X} > 900\right) + P\left(\overline{X} < 700\right)$

Exercícios Resolvidos

Exercício 12., página 390

12.a) via > pnorm(x,mean,sd)

```
> (1-pnorm(900,800,100))^4
[1] 0.0006336039
```

12.b) via > pnorm(x,mean,sd)

```
> 1-(pnorm(1000,800,100))^4
[1] 0.08794195
```

12.c) via > pnorm(x,mean,sd)

```
> 1-pnorm(900,800,50)+pnorm(700,800,50)
[1] 0.04550026
```

FEUC FACULDADE DE ECONOMIA UNIVERSIDADE DE COIMBRA

Conteúdos

Exercícios Propostos

Complemento das Aulas 5 & 6.

Distribuição Normal

- **págs. 341-342:** 69., 70., 72.

Distribuição Exponencial

- **págs. 341– 342:** 77., 80.

Murteira, B., C. Silva Ribeiro, J. Andrade e Silva, C. Pimenta, F. Pimenta (2015), **Introdução à Estatística, 3ª Edição**', Escolar Editora.

Exercícios Propostos

Complemento da Aula 7

Amostragem vs. Binomial

- **pág. 388:** 1.

Amostragem vs. Normal

- **págs. 389–390:** 7., 11.

Amostragem vs. Exponencial Negativa

- **pág. 390:** 13.

Exercício 16., p. 391

Alguns comentários:

- **a)** Vide [Murteira et al. (2015), p. 355, Definição 6.2].
- **b)** **Esta afirmação é sempre verdadeira**, para qualquer população $X \sim F$ (propriedade $E(\overline{X}) = E(X)$) Só seria falsa se substituíssemos a palavra média por variância (propriedade $V(\overline{X}) = \frac{V(X)}{n}$).
- **c)** De acordo com o enunciado, terá de verificar se a desigualdade $V(\overline{X}) \leq V(X)$ é sempre satisfeita.
- **d)** As estatísticas $\overline{X}_1$ e $\overline{X}_2$ induzem cópias da mesma função distribuição (vide [Murteira et al. (2015), p. 352]).
- **e)** **Obviamente falso**, pois a igualdade de v.a.'s i.i.d.'s refere-se à média/variância/função distribuição e não à igualdade de v.a.'s.