

Parte 3

Processo de Poisson.

Parâmetros de localização, escala e forma.

Momentos.

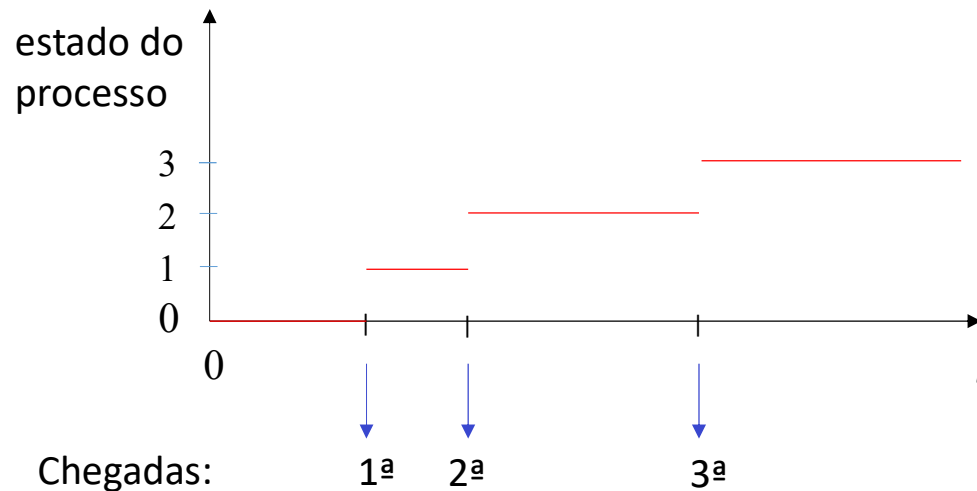
Convergências estocásticas.

Transformadas e aplicações.

Teorema Limite Central (TLC) e LGN.

3.1. Processo de Poisson (introdução)

É um processo de chegadas (aleatórias) ao longo do tempo, a partir de um instante inicial $t = 0$, sendo o **estado do processo no instante t** dado pelo **nº de chegadas até t** .



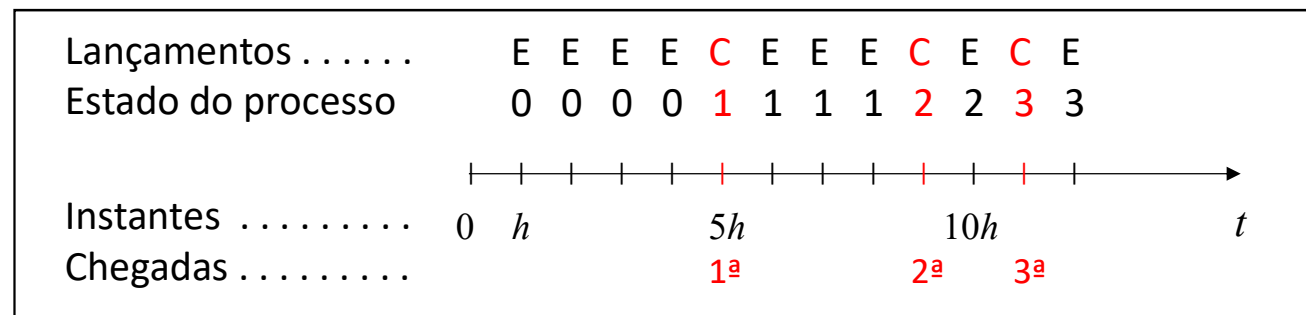
Exemplos: clientes que chegam a um serviço,
automóveis que passam numa portagem automática,
sinistros comunicados a uma companhia de seguros, etc.

O processo de Poisson pode obter-se como limite de um **processo de Bernoulli**:

Uma moeda- p é lançada em instantes sucessivos ($h, 2h, 3h, \dots$) e o valor do processo no n -ésimo instante, nh , é dado pelo **nº total de caras até esse instante**.

Dividindo o intervalo $[0, t]$ em n intervalos de amplitude h ficamos com $nh = t$ e fazendo $p = \lambda h$, com λ constante, temos que o **nº de total de caras (i.e., chegadas) até t** tem distribuição $bi(n, p)$.

Esta $bi(n, p)$ aproxima-se, quando $n \rightarrow \infty$ e $h \rightarrow 0$ ($\because p = \lambda h \rightarrow 0$), a uma lei $Poisson(np)$, com $np = n\lambda h = \lambda t$, ou seja, a uma lei **$Poisson(\lambda t)$** .



Nota: Como os lançamentos são independentes, os nºs de caras (chegadas) em intervalos de tempo disjuntos dois a dois são v.a. mutuamente independentes; e os nºs de caras em intervalos de tempo de igual amplitude têm igual lei de probabilidade (o nº de lançamentos é igual). Diz-se então que o processo tem **incrementos independentes e estacionários**.

Processo de Poisson de intensidade λ

O processo de Poisson (PP) de **taxa** ou **intensidade** λ é um processo $\{N_t\}_{t \geq 0}$ de chegadas ao longo do tempo t ($t \geq 0$). N_t representa o “nº de chegadas no intervalo $[0, t]$ ”; o instante da 1ª chegada e os intervalos de tempo entre chegadas sucessivas representam-se por T_1, T_2, T_3, \dots . O PP pode definir-se de 3 formas equivalentes:

- $N_0 = 0$
- $N_t \sim \text{Poisson}(\lambda t)$
- tem incrementos indep. e estacionários*

- $N_0 = 0$
- T_i são i.i.d. $\text{Exp}(\lambda)$
- tem incrementos indep. e estacionários*

- $N_0 = 0$
- $P(N_{t+h} - N_t = 1) = \lambda h + o(h)^\dagger$
- $P(N_{t+h} - N_t > 1) = o(h)$
- tem incrementos indep. e estac.*

* **incrementos independentes** significa que o nº de chegadas em quaisquer intervalos de tempo disjuntos dois a dois são v.a. mutuamente independentes; **incrementos estacionários** significa que em intervalos de tempo de igual amplitude, a distribuição do nº de chegadas é a mesma: $N_{t+h} - N_t$ e $N_{s+h} - N_s$ têm a mesma lei de probabilidade.

† quando $h \rightarrow 0$; uma função $f(x)$ diz-se um $o(g(x))$ quando $x \rightarrow x_0$ se $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 0$ e escreve-se $f(x) = o(g(x))$. Por exemplo, $h^2 = o(h)$ quando $h \rightarrow 0$ porque $h^2/h = h \rightarrow 0$ quando $h \rightarrow 0$

Processo de Poisson

Como $N_t \sim \text{Poisson}(\lambda t)$, então λ é o nº médio de chegadas numa unidade de tempo ($t = 1$) e ainda

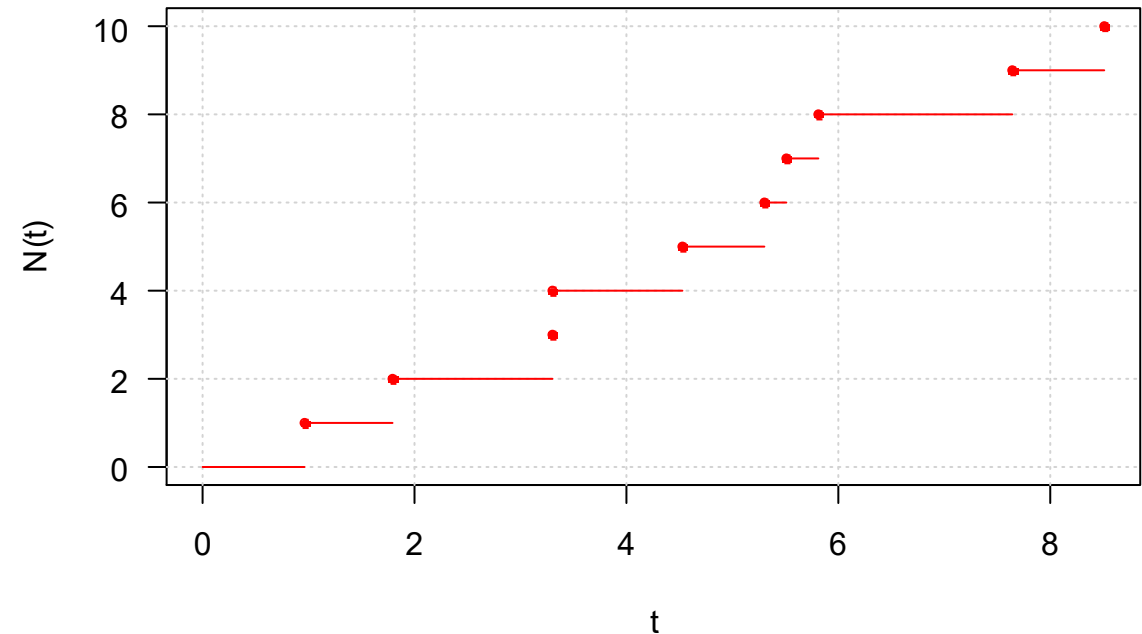
$$E(N_t) = \lambda t$$

$$\text{var}(N_t) = \lambda t$$

Prova-se* que

$$\text{cov}(N_s, N_t) = \lambda \min(s, t)$$

Trajectória (simulada) de um PP de intensidade 1



***Demonstração:** Seja $s < t$. Então $\text{var}(N_t - N_s) = \text{var}(N_t) + \text{var}(N_s) - 2\text{cov}(N_t, N_s)$ e como os incrementos são estacionários, temos $\text{var}(N_t - N_s) = \text{var}(N_{t-s})$ e portanto segue-se que $\lambda(t-s) = \lambda t + \lambda s - 2\text{cov}(N_t, N_s)$, donde $\text{cov}(N_t, N_s) = \frac{1}{2}(\lambda t + \lambda s - \lambda(t-s)) = \lambda s$. Logo $\text{cov}(N_s, N_t) = \lambda \min(s, t)$.

Processo de Poisson

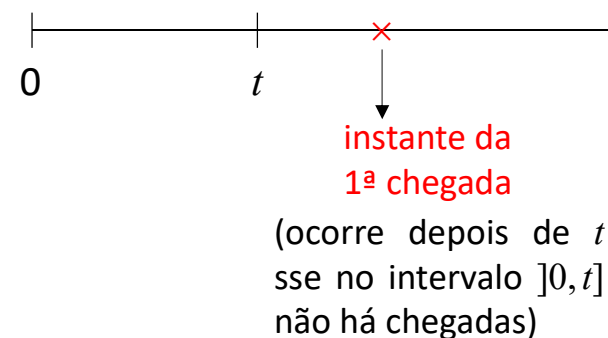
Recorde-se que T_1 é o instante da 1ª chegada e T_2, T_3, \dots são os intervalos de tempo entre as chegadas consecutivas.

É possível provar que T_1, T_2, T_3, \dots são mutuamente independentes e $Exp(\lambda)$, a partir da formulação $N_t \sim Poisson(\lambda t)$. A prova é simples para o caso T_1 :

Sendo T o instante (aleatório) da primeira chegada, temos $T > t$ sse não houver chegadas no intervalo $]0, t]$, donde

$$P(T > t) = P(N_t = 0) = e^{-\lambda t}, \text{ para } t > 0$$

Logo $T \sim Exp(\lambda)$



Processo de Poisson: generalizações

O processo de Poisson (PP) permite inúmeras generalizações, entre as quais:

- PP **composto** (as chegadas podem ser múltiplas – em vez de unitárias – com dada distribuição de probabilidade, discreta ou outra)

Exemplos: total de passageiros, num PP de chegadas de aviões a um terminal; indemnizações pagas por uma seguradora num PP de sinistros (chegadas); total de artigos comprados por clientes num PP de chegadas a uma loja; etc.

- PP **não homogéneo** (a intensidade do processo é função do tempo, i.e., $\lambda = \lambda(t)$)

Exemplo: passagem de automóveis numa portagem automática, com intensidade variável no tempo (horas de ponta, períodos de fraca intensidade, etc.)

- Processo **pontual** de Poisson (no plano ou espaço), em que o nº de pontos (chegadas) num subconjunto de volume V é $\text{Poisson}(\lambda V)$

Exemplos: colónias de bactérias numa superfície; defeitos num tecido produzido num tear; partículas de poeira no espaço; etc.

Processo de Poisson: generalizações

- Processos **de nascimento** (chegadas) em que as taxas de chegada dependem do tamanho da população e processos **de nascimento e morte** (com taxas de saídas)
Exemplos: processo de nascimento com taxas de chegada proporcionais ao tamanho da população; processos de imigração – emigração; filas de espera; etc.
- Processos de **renovamento** (os intervalos de tempo entre chegadas consecutivas são i.i.d. com distribuição arbitrária)
Exemplos: sistemas de falha e substituição instantânea de aparelhos; gestão de stocks; etc.

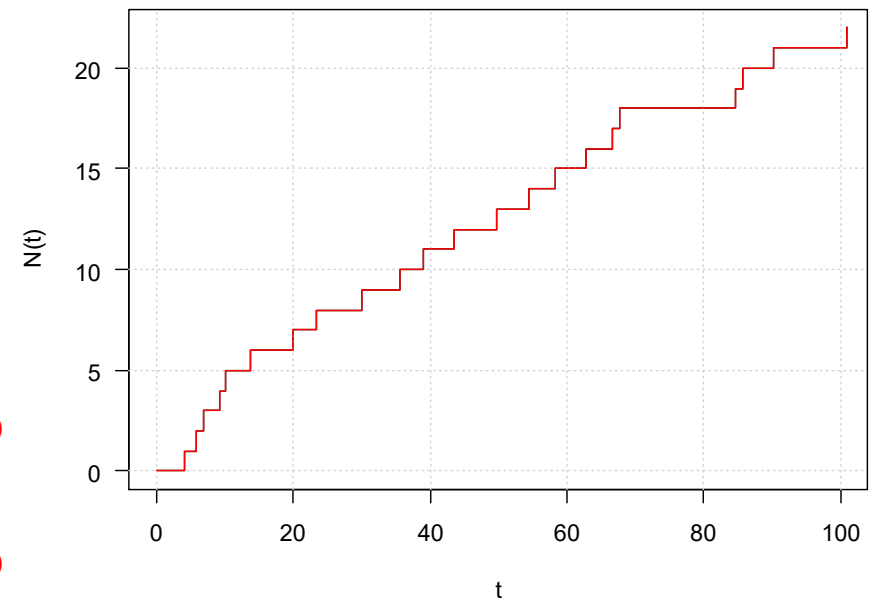
Nota: Uma variante do Processo de Bernoulli é o passeio aleatório, estudado anteriormente, que por sua vez leva ao movimento browniano (ou processo de Wiener-Lévy) por passagem ao limite, considerando saltos de amplitude $\pm \delta$, em instantes espaçados por t/n unidades de tempo, e fazendo $\delta \rightarrow 0$ e $n \rightarrow \infty$ de forma conveniente (de forma a que o processo limite tenha valor médio μt e variância $\sigma^2 t$; a prova faz-se usando transformadas de Laplace...).

Exercício: (nº 69) (i) Simule um processo de chegadas com intervalos de tempo (entre chegadas consecutivas) $Exp(\lambda)$, independentes, com $\lambda = 1/5$, no intervalo de tempo $[0, 100]$. Represente graficamente a trajectória do processo.

(ii) Para $t = 25$, estime a distribuição de $N(t) = \text{“nº de chegadas até ao instante } t \text{”}$, por meio de simulação (com 100 mil réplicas).

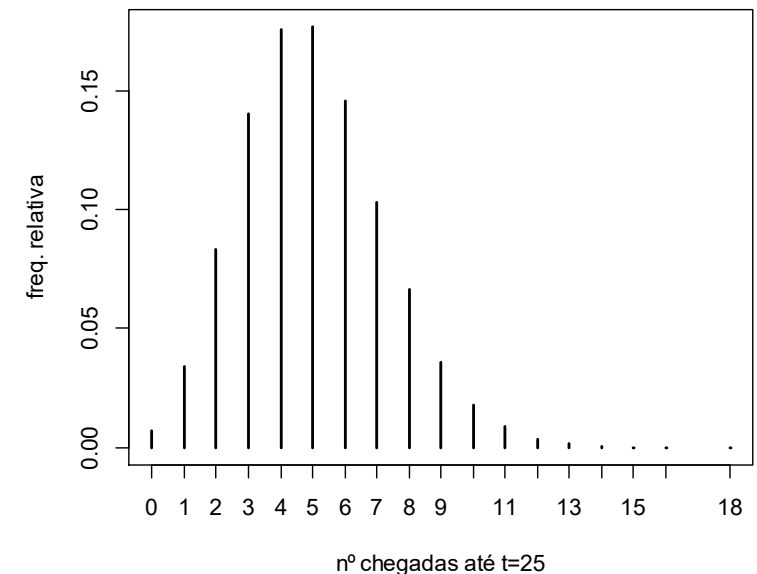
Resolução: (i) Temos $E(X) = 1/\lambda$, se $X \sim Exp(\lambda)$

```
# em média há 100/5 = 20 chegadas em [0,100];  
# simulam-se 40 intervalos de tempo Exp(1/5):  
tps <- rexp(40,1/5)  
# 1 + nº de chegadas até t=100:  
fim <- min(which(cumsum(tps)>100))  
# gráfico da trajectória do processo:  
plot(c(0,cumsum(tps[1:fim])),0:fim,type="s", las=1,  
      xlim=c(0,100),ylim=c(0,fim), xlab="t", ylab="N(t)")  
# acrescentando uma grelha:  
grid()  
points(c(0,cumsum(tps[1:fim])),0:fim,type="s",col=2)
```



Resolução: (ii)

Para $t = 25$, simulam-se 10^5 valores de $N(t)$ = “nº de chegadas até ao instante t ” e elaboram-se a correspondente tabela de frequências relativas dos valores obtidos.



```
# 1) recorrendo a matriz 50x10^5 de interv.tp
simul <- matrix(rexp(50*10^5,1/5),nr=50)
n.25 <- apply(simul,2,function(x) min(which(cumsum(x)>25))-1 )
plot(table(n.25)/100000,xlab="nº chegadas até t=25",ylab="freq. relativa")
table(n.25)/100000
      0      1      2      3      4      5      6      7      8      9     10 ...
0.00701 0.03388 0.08331 0.14005 0.17556 0.17673 0.14573 0.10299 0.06648 0.03573 0.01801 ...
round(dpois(0:20,1/5*25),5)
[1] 0.00674 0.03369 0.08422 0.14037 0.17547 0.17547 0.14622 0.10444 0.06528 0.03627 0.01813 ...

# 2) recorrendo a um ciclo for ...
```

Resolução: (ii)

```
# 2) recorrendo a um ciclo for
```

```
n.25 <- 0
```

```
for (i in 1:100000)
```

```
{ t <- 0; n <- 0;
```

```
  while(t<25)
```

```
    {t <- t+rexp(1,1/5); n <- n+1} ;
```

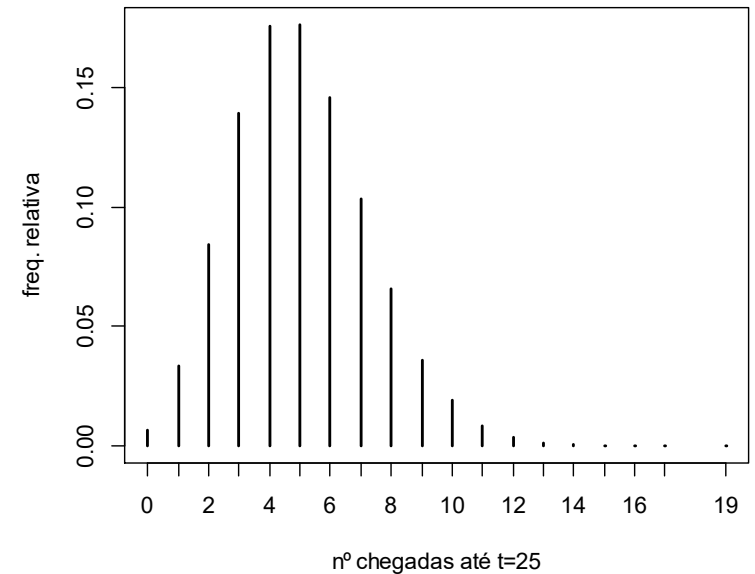
```
    n.25[i] <- n-1 }
```

```
table(n.25)/100000
```

0	1	2	3	4	5	6	7	8	9
0.00660	0.03345	0.08455	0.13954	0.17601	0.17642	0.14590	0.10344	0.06600	0.03584
10	11	12	13	14	15	16	17	19	
0.01895	0.00819	0.00331	0.00117	0.00041	0.00013	0.00006	0.00002	0.00001	

```
round(dpois(0:19,1/5*25),5)
```

0.00674	0.03369	0.08422	0.14037	0.17547	0.17547	0.14622	0.10444	0.06528	0.03627
0.01813	0.00824	0.00343	0.00132	0.00047	0.00016	0.00005	0.00001	0.00000	0.00000



3.2. Parâmetros de localização, escala e forma.

Numa família (paramétrica) de v.a. X com fd $F(x; \theta)$,

θ ($\theta \in \mathbb{R}$), diz-se um **parâmetro de localização** se a distribuição de $X - \theta$ não depender de θ .
A família diz-se então **família de localização**.

θ ($\theta > 0$), diz-se um **parâmetro de escala** se a distribuição de X/θ não depender de θ .
A família diz-se então **família de escala**.

Os parâmetros que não são de localização nem escala são chamados **parâmetros de forma**.

$$X \sim N(\theta, 1) \Rightarrow X - \theta \sim N(0, 1)$$

$$X \sim U[0, \theta] \Rightarrow X/\theta \sim U[0, 1]$$

$$X \sim \text{Exp}(1/\theta) \Rightarrow X/\theta \sim \text{Exp}(1)$$

$$\text{Poisson}(\theta)$$

Famílias de localização-escala

Uma família de v.a. X com fd $F(x; \lambda, \delta)$, $\lambda \in \mathbb{R}$, $\delta > 0$, diz-se de **localização-escala** se a distribuição de $(X - \lambda)/\delta$ não depender de λ nem de δ (estes são os parâmetros de localização e escala)

$$N(\lambda, \delta)$$

Nota: no caso de famílias de v.a. absolutamente contínuas, θ é um parâmetro

(i) de localização ($\theta \in \mathbb{R}$), se a fdp se escrever na forma $f(x; \theta) = g(x - \theta)$;

(ii) de escala ($\theta > 0$), se a fdp se puder escrever na forma $f(x; \theta) = \frac{1}{\theta} g\left(\frac{x}{\theta}\right)$.

e (λ, δ) é um parâmetro (bivariado) de localização-escala se $f(x; \lambda, \delta) = \frac{1}{\delta} g\left(\frac{x - \lambda}{\delta}\right)$

Correspondem às famílias $\theta + X$, θX e $\lambda + \delta X$, construídas a partir de uma v.a. X com fdp g (cf. Slides 150-151).

3.3. Momentos

Dada uma v.a. X qualquer, definem-se (caso existam...) os **momentos simples** e os **momentos centrais** por

$$\mu'_n = E(X^n)$$

momento de ordem n

$$\mu_n = E((X - \mu)^n)$$

momento central de ordem n

Casos particulares:

$$\mu'_1 = \mu = E(X)$$

valor médio

$$\mu_2 = \sigma^2 = E((X - \mu)^2)$$

variância

Os momentos μ_3 e μ_4 entram na definição do coeficiente de assimetria e curtose,

$$\beta_1 = E\left(\left(\frac{X - \mu}{\sigma}\right)^3\right)$$

coeficiente de assimetria

$$\beta_2 = E\left(\left(\frac{X - \mu}{\sigma}\right)^4\right)$$

coeficiente de curtose

que medem, respectivamente, a assimetria e o peso das caudas da distribuição. Numa distribuição simétrica, temos $\beta_1 = 0$, mas a recíproca não é verdadeira (vd. exercício nº 37); e também todos os momentos de ordem ímpar (caso existam) são nulos, mas a recíproca não é verdadeira: há distribuições com todos os momentos de ordem ímpar nulos, mas que não são simétricas!

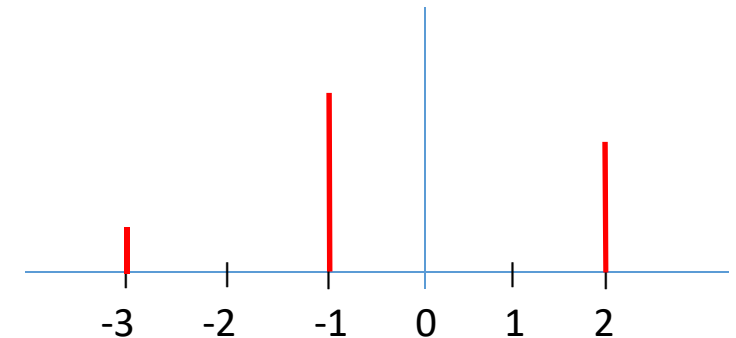
Exercício: (nº 37, slide 103) Calcule o coeficiente de assimetria da v.a. discreta X com fmp (não simétrica)

$$X: \begin{cases} -3 & -1 & 2 \\ \frac{1}{10} & \frac{5}{10} & \frac{4}{10} \end{cases}$$

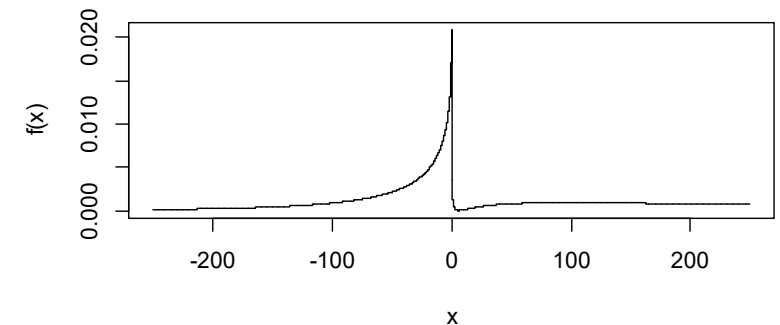
Resolução: Calculam-se os momentos

$$\mu = -\frac{3}{10} - \frac{5}{10} + \frac{8}{10} = 0 \quad \left(\sigma^2 = E(X^2) = \frac{9}{10} + \frac{5}{10} + \frac{16}{10} = 3 \right)$$

$$\mu_3 = E(X^3) = -\frac{27}{10} - \frac{5}{10} + \frac{32}{10} = 0 \quad \text{donde} \quad \beta_1 = \frac{1}{\sigma^3} E((X - \mu)^3) = 0$$



Nota: A fdp $f(x) = \frac{1}{48}(1 - \text{sgn}(x)) \sin(|x|^{1/4}) e^{-|x|^{1/4}}$,
devida a Stieltjes, não é simétrica, no entanto todos
os momentos de ordem ímpar são nulos!



Num par aleatório (X, Y) temos os **momentos conjuntos** de ordem (r, s) , simples e centrais, definidos, para $r, s \in \{0, 1, 2, \dots\}$ por (caso existam os valores médios considerados)

$$\mu'_{r,s} = E(X^r Y^s)$$

momento conjunto de ordem (r, s)

$$\mu_{r,s} = E\left((X - \mu_X)^r (Y - \mu_Y)^s\right)$$

momento conjunto central de ordem (r, s)
covariância, para $r = s = 1$

Casos particulares: $\mu_{2,0} = \text{var}(X)$, $\mu_{0,2} = \text{var}(Y)$ e $\mu_{1,1} = \text{cov}(X, Y) = \mu'_{1,1} - \mu'_{1,0} \mu'_{0,1}$

Nota: Os momentos desempenham um papel importante em probabilidades e estatística. Em certos casos, a sequência dos momentos identifica a distribuição.

3.4. Convergências estocásticas

- convergência fraca / em lei / em distribuição
- convergência em probabilidade

Outros tipos de convergência:
convergência em média quadrática
convergência em média de ordem r
convergência forte / quase certa.

Seja X_n uma sucessão de v.a. com fd F_n ($n = 1, 2, \dots$) e X uma v.a. com fd F .

Diz-se que X_n converge em lei ou converge em distribuição para X se

$$F_n(x) \xrightarrow{n \rightarrow \infty} F(x), \quad \forall x : x \text{ é ponto de continuidade de } F,$$

e escreve-se $X_n \xrightarrow{L} X$ ou $X_n \xrightarrow{d} X$

(diz-se também que F_n converge em lei ou em distribuição, ou converge fracamente para F)

Convergências estocásticas

- convergência fraca / em lei / em distribuição
- convergência **em probabilidade**

Seja X uma v.a. e X_n uma sucessão de v.a. ($n = 1, 2, \dots$).

Diz-se que X_n **converge em probabilidade** para X se para qualquer $\varepsilon > 0$,

$$P(|X_n - X| > \varepsilon) \xrightarrow{n \rightarrow \infty} 0$$

e escreve-se $X_n \xrightarrow{P} X$

Nota: Trata-se da convergência da probabilidade do acontecimento $\{|X_n - X| > \varepsilon\}$ e não da convergência das v.a. X_n para a v.a. X ; a condição na definição equivale a $P(|X_n - X| \leq \varepsilon) \xrightarrow{n \rightarrow \infty} 1$

Exemplo 26: Dadas v.a. $X_n \sim \text{bi}(1, p)$, com fmp $X_n : \begin{cases} 0 & 1 \\ 1 - \frac{1}{n} & \frac{1}{n} \end{cases}$

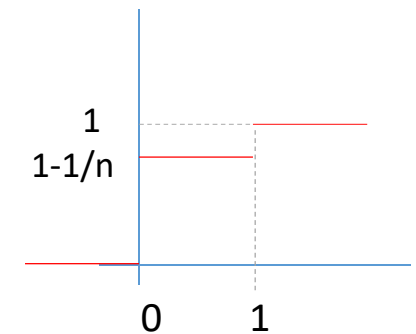
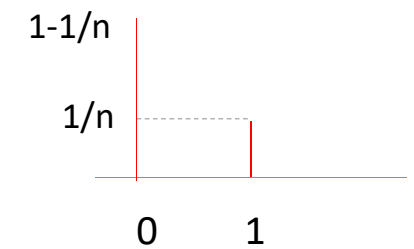
temos $P(|X_n| > \varepsilon) = \begin{cases} P(X_n = 1) = \frac{1}{n} & \text{se } 0 < \varepsilon < 1 \\ 0 & \text{se } \varepsilon \geq 1 \end{cases}$

donde $P(|X_n| > \varepsilon) \rightarrow 0$ quando $n \rightarrow \infty$

Logo $X_n \xrightarrow{P} 0$

Por outro lado, temos $F_n(x) = P(X_n \leq x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 - \frac{1}{n} & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases}$

donde $F_n(x) \xrightarrow{n \rightarrow \infty} \begin{cases} 0 & \text{se } x < 0 \\ 1 & \text{se } x \geq 0 \end{cases}$ e portanto $X_n \xrightarrow{L} 0$



Propriedades (convergência em lei e em probabilidade)

$$X_n \xrightarrow{L} X \Rightarrow \begin{cases} X_n + c \xrightarrow{L} X + c \\ cX_n \xrightarrow{L} cX \quad (c \neq 0) \end{cases}$$

$$X_n \xrightarrow{P} X \Leftrightarrow X_n - X \xrightarrow{P} 0$$

$$X_n \xrightarrow{P} X, \quad Y_n \xrightarrow{P} Y \Rightarrow X_n \pm Y_n \xrightarrow{P} X \pm Y$$

$$X_n \xrightarrow{P} X \Rightarrow cX_n \xrightarrow{P} cX$$

⋮

Teoremas (convergência em lei e em probabilidade)

$$1) \quad X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{L} X \quad (\text{a recíproca não é verdadeira})$$

$$2) \quad X_n \xrightarrow{L} c \Rightarrow X_n \xrightarrow{P} c$$

Nota:

A convergência em probabilidade não implica a convergência do momento de ordem r ($r = 1, 2, \dots$).

Por exemplo, se $X_n : \begin{cases} 0 & n \\ 1 - \frac{1}{n} & \frac{1}{n} \end{cases}$ temos $X_n \xrightarrow{P} 0$ mas $E(X_n^r) = \frac{1}{n} n^r \not\rightarrow 0$ quando $n \rightarrow \infty$

3.5. Transformadas e aplicações – ver separata “Transformadas”

Transformada de Laplace, função geradora de momentos (f.g.m.), função geradora de probabilidades (f.g.p.), função característica (f.c.).

$$L(t) = E(e^{-tX})$$

se este valor médio existir numa vizinhança de $t = 0$

$$M(t) = E(e^{tX})$$

se este valor médio existir numa vizinhança de $t = 0$

$$P(z) = E(z^X)$$

para v.a. discretas com suporte $\{0, 1, 2, \dots\}$

$$\varphi(t) = E(e^{itX})$$

existe para todo $t \in \mathbb{R}$

3.6. Teorema Limite Central (TLC)

Seja X_1, X_2, \dots uma sucessão de v.a. i.i.d. com valor médio μ e variância σ^2 , e seja $S_n = X_1 + X_2 + \dots + X_n$. Então $\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$

(analogamente, $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$, sendo $\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$)

Logo, para n suficientemente grande*, temos que a f.d. de S_n é aproximadamente $N(n\mu, \sigma\sqrt{n})$. Analogamente, a f.d. de \bar{X} é aproximadamente $N(\mu, \frac{\sigma}{\sqrt{n}})$

* se a distribuição de X não for muito assimétrica, geralmente $n = 30$ (ou até menos) é suficiente

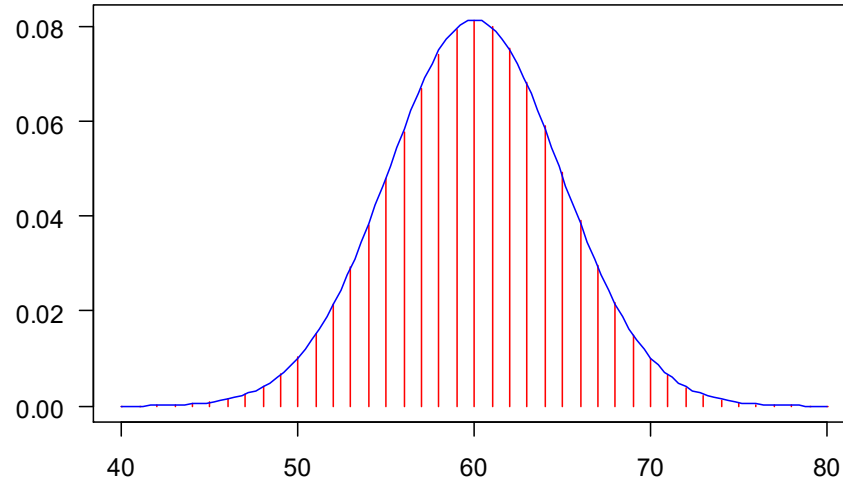
TLC – caso particular Bernoulli, i.e., $X \sim bi(1, p)$

O caso particular $X \sim bi(1, p)$ deve-se a A. de Moivre (c. 1733) e a Laplace.

Neste caso temos $S_n \sim bi(n, p)$ e então

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} Z$$

Ilustração com gráfico da fmp $bi(n, p)$, $n=100$, $p=0.6$ e fdp normal sobreposta:



fmp $bi(100, 0.6)$

fdp $N(60, \sqrt{24})$

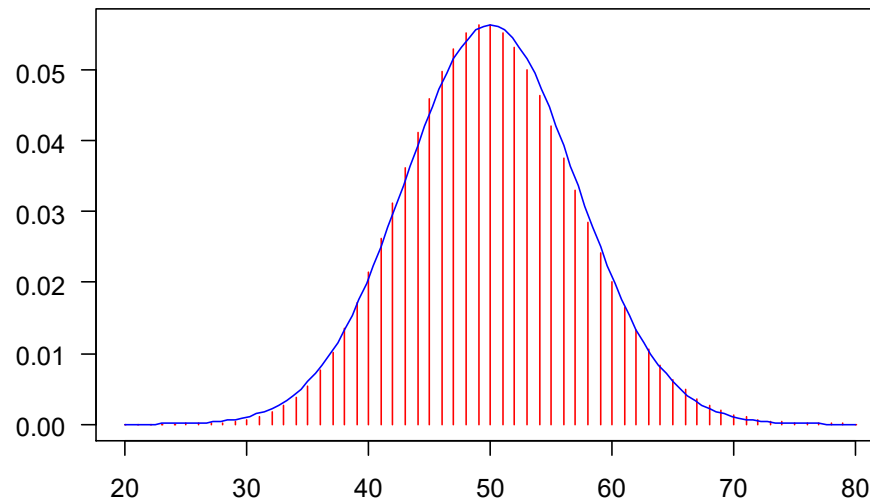
$$S_n \approx N(np, \sqrt{np(1-p)})$$

TLC – caso particular *Poisson*

No caso $X \sim \text{Poisson}(\lambda)$ temos $S_n \sim \text{Poisson}(n\lambda)$ e então

$$\frac{S_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow{d} Z$$

Ilustração com gráfico de fmp $\text{Poisson}(n\lambda)$, $\lambda = 1$, $n = 50$, e fdp normal sobreposta:



fmp $\text{Poisson}(50)$

fdp $N(50, \sqrt{50})$

$X \approx N(\lambda, \sqrt{\lambda})$, para λ grande

TLC – Correcção de continuidade

(para melhorar a aproximação de uma v.a. discreta à normal, no TLC)

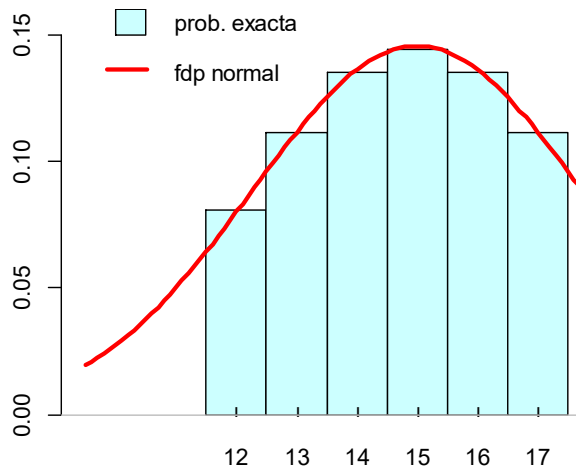
Por exemplo, para calcular uma aproximação de $P(12 \leq X \leq 17)$, com $X \sim bi(\underbrace{30}_n, \underbrace{0.5}_p)$,
toma-se $P(11.5 < Y < 17.5)$, com $Y \sim N(\underbrace{15}_{np}, \underbrace{\sqrt{7.5}}_{np(1-p)})$

```
# probabilidade exacta:
pbinom(17, 30, 0.5) - pbinom(11, 30, 0.5)
[1] 0.7189585
# prob aproximada pelo TLC, sem correcção:
pnorm(17, 15, sqrt(7.5)) - pnorm(11, 15, sqrt(7.5))
[1] 0.6953321
# prob aproximada pelo TLC, com correcção:
pnorm(17.5, 15, sqrt(7.5)) - pnorm(11.5, 15, sqrt(7.5))
[1] 0.7187235
```

TLC – Correção de continuidade (ilustração)

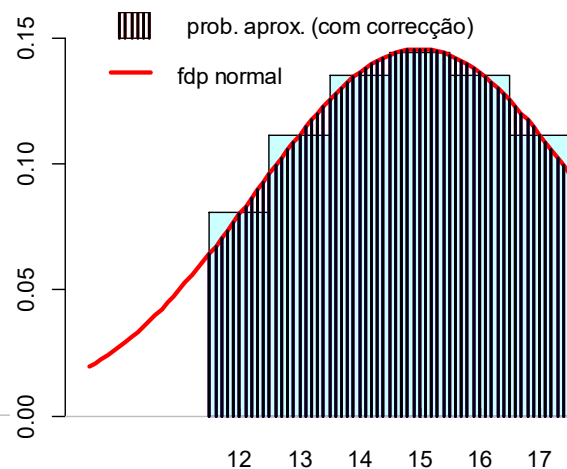
$$X \sim bi(\underbrace{30}_n, \underbrace{0.5}_p)$$

$$Y \sim N(\underbrace{15}_{np}, \underbrace{\sqrt{7.5}}_{np(1-p)})$$



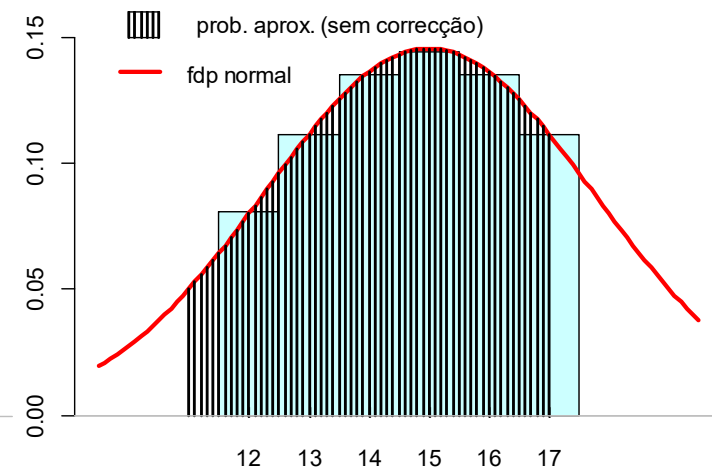
$$P(12 \leq X \leq 17)$$

0.7190



$$P(11.5 < Y < 17.5)$$

0.7187

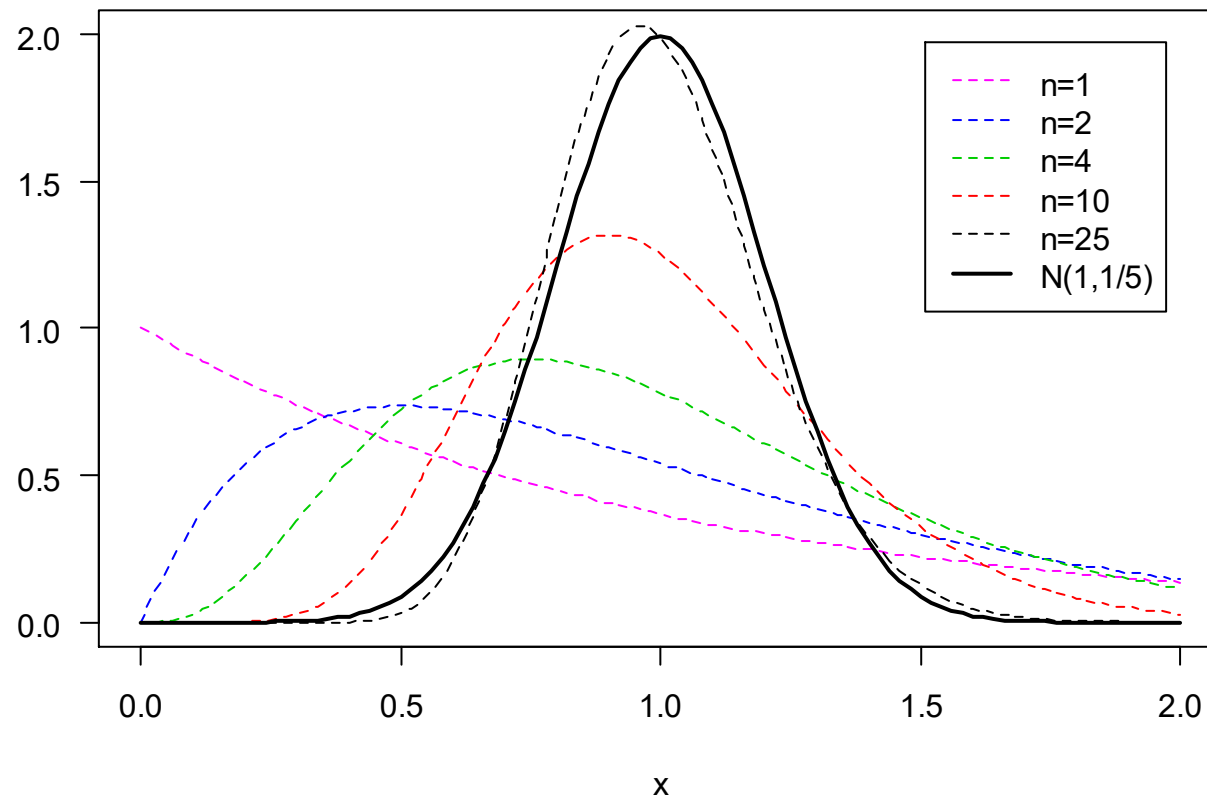


$$P(11 < Y \leq 17)$$

0.6953

TLC – ilustração no caso $X_i \sim \text{Exp}(\lambda)$, com $\lambda = 1$

f.d.p. da média de n v.a. i.i.d. $\text{Exp}(1)$



Note-se que neste caso se sabe que

$$S_n \sim \text{Gama}(n, \lambda)$$

$$\bar{X} \sim \text{Gama}(n, n\lambda)$$

Pelo TLC,

$$\bar{X} \approx N\left(\frac{1}{\lambda}, \frac{1}{\lambda\sqrt{n}}\right)$$

para n grande

Lei dos grandes números (LGN)

Se X_1, X_2, \dots são i.i.d., com $\mu = E(X)$, então $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \mu$

$$\forall \varepsilon > 0, \quad P(|\bar{X} - \mu| \leq \varepsilon) \xrightarrow{n \rightarrow \infty} 1$$

Caso particular: $f_n(A)$ = “freq. relativa do acontecimento A , em n repetições independentes de uma experiência aleatória”, com $p = P(A)$. Então as v.a.

$$X_i = \begin{cases} 1 & , \text{se } A \text{ ocorreu na } i\text{-ésima repetição} \\ 0 & , \text{caso contrário} \end{cases}$$

são i.i.d. $bi(1, p)$, com $\mu = E(X) = p$ e

$$f_n(A) = \frac{\text{n}^\circ \text{ de vezes que } A \text{ ocorreu}}{n} = \frac{X_1 + \dots + X_n}{n} = \bar{X}$$

Logo, se n for grande, $f_n(A)$ será uma boa estimativa de $P(A)$

LGN – demonstração no caso de variância finita (pela desigualdade de Markov)

Sejam X_1, X_2, \dots i.i.d., com $\mu = E(X)$ e $\text{Var}(X) = \sigma^2 < +\infty$. Seja $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.
Então temos $E(\bar{X}) = \mu$ e $\text{Var}(\bar{X}) = E((\bar{X} - \mu)^2) = \frac{\sigma^2}{n}$

Pela desigualdade de Markov*, com $r = 2$, temos

$$\forall \varepsilon > 0, \quad P(|\bar{X} - \mu| \leq \varepsilon) \geq 1 - \frac{E((\bar{X} - \mu)^2)}{\varepsilon^2} = 1 - \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \rightarrow \infty} 1$$

donde $\bar{X} \xrightarrow{P} \mu$

$$^*P(|X| > a) \leq P(|X| \geq a) \leq \frac{1}{a^r} E(|X|^r) \quad \text{ou então} \quad P(|X| \leq a) \geq 1 - \frac{1}{a^r} E(|X|^r)$$

LGN – demonstração no caso de variância finita (pelo TLC)

$$\begin{aligned}\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P(|\bar{X} - \mu| < \varepsilon) &= \lim_{n \rightarrow \infty} P(-\varepsilon < \bar{X} - \mu < \varepsilon) = \\ &= \lim_{n \rightarrow \infty} P\left(-\frac{\varepsilon\sqrt{n}}{\sigma} < \frac{\bar{X} - \mu}{\sigma}\sqrt{n} < \frac{\varepsilon\sqrt{n}}{\sigma}\right) \stackrel{TLC}{=} \\ &= \lim_{n \rightarrow \infty} P\left(-\frac{\varepsilon\sqrt{n}}{\sigma} < Z < \frac{\varepsilon\sqrt{n}}{\sigma}\right) = 1\end{aligned}$$

A LGN, no caso particular $X \sim bi(1, p)$, foi demonstrada por *J. Bernoulli*, tendo sido publicada (postumamente) em 1713 na obra *Ars Conjectandi*. No caso geral (sem a exigência de variância finita) a LGN deve-se a *A. Khinchine* (1929).

O TLC, no caso particular $X \sim bi(1, p)$, é também conhecido por teorema de *de Moivre – Laplace*. A primeira versão do TLC como “lei universal” é da autoria de *Laplace* (1810). No caso mais geral (v.a. i.i.d. quaisquer, com variância finita), o TLC deve-se a *J. Lindeberg* (1922).

Há várias generalizações do TLC (relaxando as hipóteses de independência e/ou de igual distribuição e/ou de variância finita). E também há versões multivariadas do TLC e suas generalizações. No caso univariado com v.a. i.i.d. quaisquer (sem a exigência de variância finita) obtêm-se as distribuições chamadas “estáveis” (a normal, a Cauchy, a Lévy e outras; as estáveis formam uma família de 4 parâmetros; a normal é a única que tem variância finita).

TLC – demonstração (apenas no caso de existir t. Laplace dos X_i)

Seja $U_i = X_i - \mu$. Então os U_i são i.i.d. com valor médio 0 e variância σ^2 , e t. Laplace

$$L_U(t) = e^{-\mu t} L_X(t), |t| < t_0. \text{ Seja } Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n U_i = \sum_{i=1}^n \frac{U_i}{\sigma\sqrt{n}}$$

Então $L_{Z_n}(t) = \left(L_U\left(\frac{t}{\sigma\sqrt{n}}\right) \right)^n$, pelas propriedades dadas sobre t. Laplace.

Expandindo $L_U(x)$ em série de Maclaurin, obtemos (lembrar que $L_U^{(n)}(0) = (-1)^n E(U^n)$)

$$L_U(x) = 1 - x E(U) + \frac{1}{2} x^2 E(U^2) + o(x^2) = 1 + \frac{1}{2} x^2 \sigma^2 + o(x^2)$$

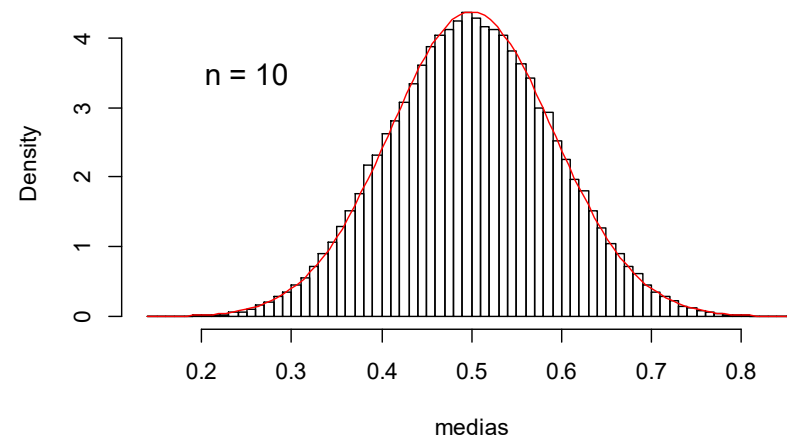
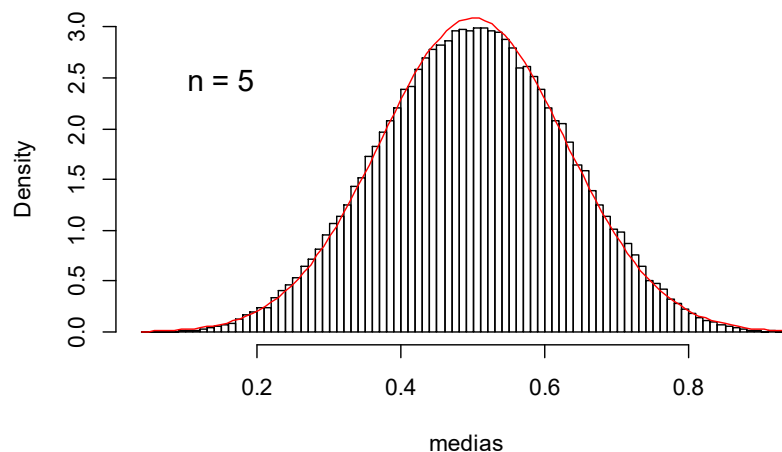
donde
$$L_{Z_n}(t) = \left(L_U\left(\frac{t}{\sigma\sqrt{n}}\right) \right)^n = \left(1 + \frac{t^2}{2n} + o\left(\frac{1}{n}\right) \right)^n \xrightarrow{n \rightarrow \infty} \boxed{e^{t^2/2}} \quad \begin{array}{l} \text{t. Laplace de } Z \sim N(0,1) \\ \text{(função contínua na origem)} \end{array}$$

Como a t. Laplace identifica a f.d. e a convergência em distribuição é equivalente à convergência das t.Laplace (para uma função contínua na origem) concluimos que $Z_n \xrightarrow{d} Z$

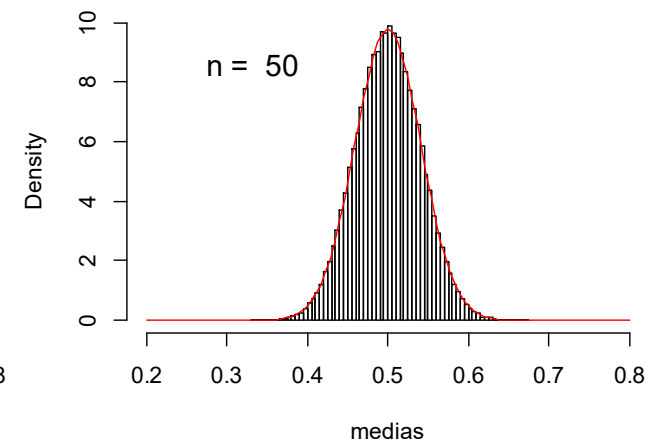
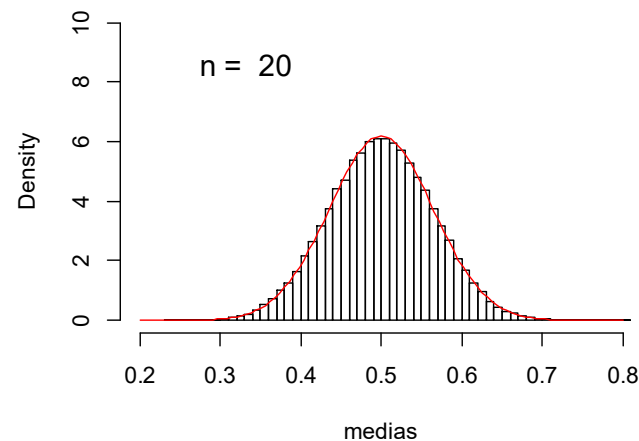
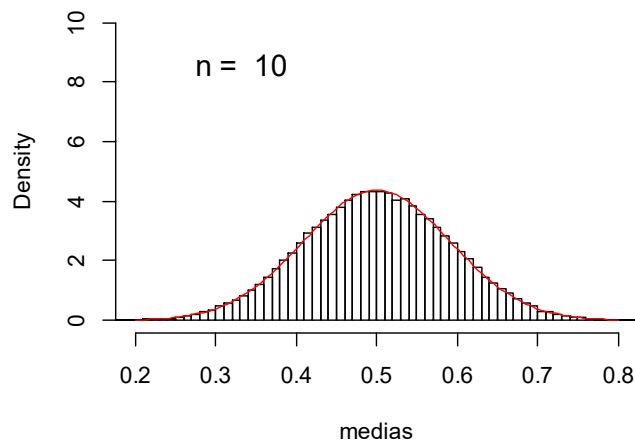
Exercícios resolvidos

Exercício: Ilustre o TLC para a média de n v.a. i.i.d. $X_i \sim U[0,1]$, simulando e sobrepondo a curva normal (fdp) adequada, para $n = 5, 10, 30$.

```
# função para simular médias de n v.a. i.i.d. U[0,1] com n° réplicas nrep:  
fmed <- function(n, nrep) colMeans( matrix(runif(n*nrep),nr=n) )  
medias <- fmed(5, 10^5)  
hist(medias, 75, freq=F, main="")  
# gráfico para n=5:  
n0 <- 5; curve(dnorm(x, 0.5, sqrt(1/12/n0)), add=T, col=2)  
legend("topleft", inset=0.05, legend=paste("n = ", n0), bty="n")
```



Exercício (cont.): Comente os resultados obtidos para as distribuições das médias, considerando $n = 10, 50, 100$.



A distribuição da média de n v.a. i.i.d. $U[0,1]$ fica cada vez mais concentrada em torno do seu valor médio, $\mu = 0.5$, à medida que n aumenta (o desvio padrão de \bar{X} é $\frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{12n}} \xrightarrow{n \rightarrow \infty} 0$).

Mais geralmente, se desconhecermos a distribuição dos dados de uma amostra de n v.a. i.i.d. com X (com $\sigma^2 < +\infty$), estimamos o valor médio de X (i.e., μ , desconhecido) por \bar{X} , com uma precisão que aumenta quando n aumenta.

Exercício: (nº 65) Em cada dia uma certa acção A pode descer \$1, manter-se, ou subir \$1 com probs 0.39, 0.2, 0.41 resp. (as alterações diárias são independentes). Calcule um valor aproximado da prob de ao fim de 700 dias a cotação ter subido mais de \$10 acima do seu valor inicial, usando (i) o TLC (ii) simulação

Resolução: A fmp da v.a. X que representa a alteração diária é $X : \begin{cases} -1 & 0 & 1 \\ 0.39 & 0.2 & 0.41 \end{cases}$

Temos $\mu = E(X) = -0.39 + 0.41 = 0.02$ e $\sigma^2 = \text{Var}(X) = 0.39 + 0.41 - 0.02^2 = 0.7996$.

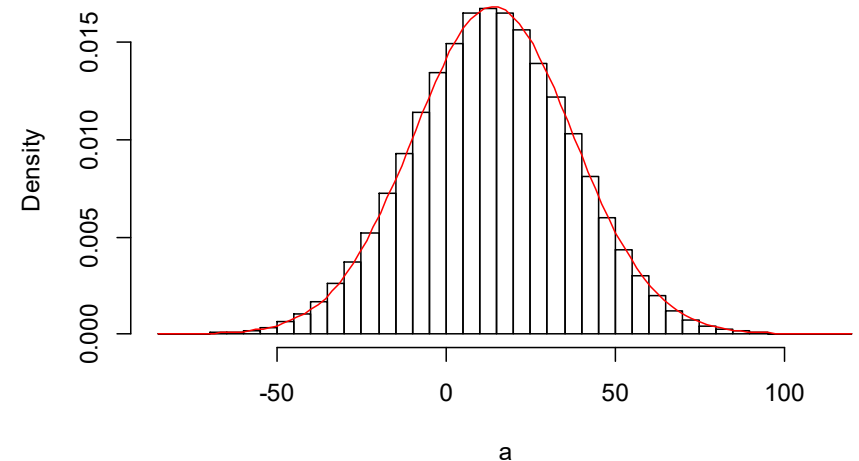
Então $S_{700} = X_1 + \dots + X_{700}$ (alteração após 700 dias), X_i i.i.d. com X , tem distribuição aproximada normal com v. médio $700 \times 0.02 = 14$ e variância $700 \times 0.7996 = 559.72$

Aplica-se então o TLC (pois $n = 700$ é suficientemente grande) para calcular um valor aproximado de $P(S_{700} > 10)$, com correcção de continuidade, ou seja, calcula-se o valor de $P(Z_{700} > 10.5)$, onde $Z_{700} \sim N(14, 559.72^{1/2})$, conforme segue adiante. Calculou-se também um valor aproximado de $P(S_{700} = 0)$, pelo TLC e por simulação.

Resolução (cont.):

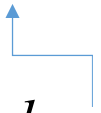
```
## por simulação:
a <- 0 ; p = c(0.39,0.2,0.41)
for (i in 1:10^5) a[i] <- sum(sample(c(-1,0,1),700, rep=T, prob=p))
sum(a>10)/10^5
[1] 0.55843
mean(a)
[1] 14.01794
var(a)
[1] 556.97229
sum(a==0)/10^5
[1] 0.01435
## pelo TLC:
pnorm(10.5,14,sqrt(700*0.7996),lower=F)
[1] 0.5588045
pnorm(0.5,14,sqrt(700*0.7996))-pnorm(-0.5,14,sqrt(700*0.7996))
[1] 0.01415351
hist(a, 50, freq=F, main="")
curve(dnorm(x,14,sqrt(700*0.7996)), add=T, col=2)
```

$$E(S_{700}) = 700\mu = 700 \times 0.02 = 14$$
$$Var(S_{700}) = 700\sigma^2 = 700 \times 0.7996 = 559.72$$



Exercício: Seja X uma v.a. com fdp f e com transformada de Laplace L . Se f for uma função par, o que se pode concluir sobre L ?

Resolução: Temos $f(x) = f(-x)$, $x \in \mathbb{R}$. Então, usando este facto e fazendo a mudança de variável $x = -y$ no 2º integral, temos

$$L(t) = \int_{-\infty}^{+\infty} e^{-tx} f(x) dx = \int_{-\infty}^{+\infty} e^{-tx} f(-x) dx = \int_{-\infty}^{+\infty} e^{-(-t)y} f(y) dy = L(-t)$$


donde se conclui que a transformada de Laplace também é uma função par

Nota: Como exemplos de v.a. com densidade par temos a $N(0, \sigma)$, $U[-c, c]$ e $Laplace(0,1)$, esta última com fdp $f(x) = \frac{1}{2} e^{-|x|}$, $x \in \mathbb{R}$. E com fmp par temos a $U\{-n, -n+1, \dots, n-1, n\}$.

Exercício: Prove que se existir o momento de ordem r ($r > 0$) de uma v.a. X então também existe o momento de ordem s , para $0 < s < r$. Ou seja, prove que

$$E(|X|^r) < +\infty \Rightarrow E(|X|^s) < +\infty, \forall s : 0 < s < r$$

Demonstração: (para o caso de v.a. X contínua com fdp f ; o caso discreto é análogo)

Temos como hipótese que $E(|X|^r) < +\infty$. Seja s tal que $0 < s < r$. Então

$$\begin{aligned} E(|X|^s) &= \int_{-\infty}^{+\infty} |x|^s f(x) dx = \int_{|x|^s \leq 1} |x|^s f(x) dx + \int_{|x|^s > 1} |x|^s f(x) dx \leq \\ &\leq \int_{|x|^s \leq 1} f(x) dx + \int_{|x|^s > 1} |x|^r f(x) dx \leq P(|X|^s \leq 1) + E(|X|^r) < +\infty \end{aligned}$$

donde se conclui que existe o momento de ordem s .

Nota: Como consequência deste resultado, se $\text{Var}(X)$ existir, também existe o valor médio de X . E se não existir o momento de ordem r , também não existe momento de ordem t , para $t > r$

Exercício: (nº 64) Num passeio aleatório simétrico, partindo de fortuna inicial 0, calcule a probabilidade de que a fortuna do jogador ao fim de 100 passos esteja compreendida entre 0 e 50 (inclusive). Compare o resultado exacto (recorde que a fortuna num jogo é dada por $Y=2X-1$, com $X \sim bi(1,0.5)$) com o resultado aproximado de $P(0 \leq Y_1 + \dots + Y_{100} \leq 50)$, obtido pelo TLC com correcção de continuidade.

Resolução:

(i) Cálculo da probabilidade exacta, usando $T = X_1 + \dots + X_{100} \sim bi(100,0.5)$

$$P(0 \leq Y_1 + \dots + Y_{100} \leq 50) = P(0 \leq 2(X_1 + \dots + X_{100}) - 100 \leq 50) = P(50 \leq T \leq 75) = 0.5398$$

```
pbinom(75,100,0.5)-pbinom(49,100,0.5)
[1] 0.5397945
sum(dbinom(50:75,100,0.5))
[1] 0.5397945
```

Resolução (cont.):

- (ii) Cálculo da probabilidade aproximada, com correcção de 0.5 no TLC, usando $X_i \sim bi(1, 0.5)$; aqui F é a fd $N(50, 5)$, pois $n\mu = np = 50$, $n\sigma^2 = np(1-p) = 25$;

$$P(0 \leq Y_1 + \dots + Y_{100} \leq 50) = P(50 \leq T \leq 75) = F(75.5) - F(49.5) = 0.5398$$

```
pnorm(75.5, 50, 5) - pnorm(49.5, 50, 5)
[1] 0.5398277
```

- (iii) Cálculo da probabilidade aproximada, com correcção de continuidade, usando $S = Y_1 + \dots + Y_{100}$, com $Y_i \sim U\{-1, 1\}$; note-se que o suporte de S é conjunto $\{-100, -98, -96, \dots, 96, 98, 100\}$, donde a correcção no TLC é de 1 e não 0.5;

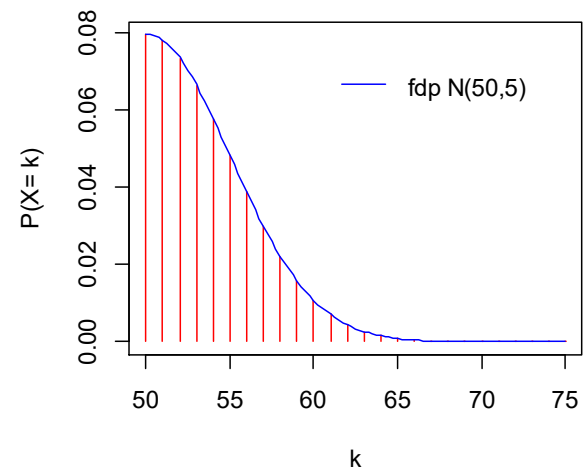
$$P(0 \leq S \leq 50) = G(51) - G(-1) = 0.5398 \quad \rightarrow \text{aqui } G \text{ representa a fd } N(0, 10), \text{ pois}$$

$\mu = E(Y) = 0$, $\sigma^2 = \text{Var}(Y) = E(Y^2) = 1$
donde $n\mu = 0$, $\sigma\sqrt{n} = 10$,

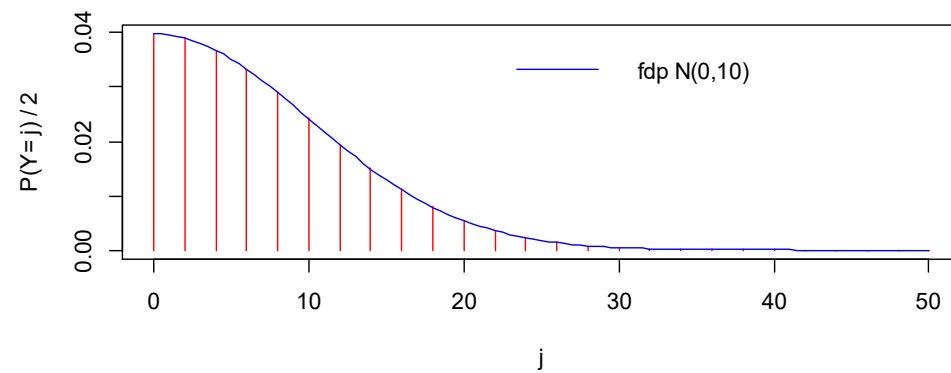
```
pnorm(51, 0, 10) - pnorm(-1, 0, 10)
[1] 0.5398277
```

Resolução (cont.): gráficos

(ii)



(iii)



Exercício: (nº 66) O tempo de atendimento de cada cliente num certo balcão tem valor médio 15 min e desvio padrão 4.5 min. Numa amostra de 50 clientes (com tempos de atendimento mutuamente independentes), calcule a probabilidade (aproximada) de que a média dos 50 tempos de atendimento

(i) exceda 16 min (ii) esteja compreendida entre 14.5 e 15.5 min.

Sejam X_1, \dots, X_n i.i.d. com valor médio $\mu = 15$ e desvio padrão $\sigma = 4.5$, com $n = 50$.

Pelo TLC, com a média $\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$, temos $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$

Logo, a f.d. de \bar{X} é aproximadamente igual à f.d. de $Z^* \sim N(\mu, \frac{\sigma}{\sqrt{n}})$, donde

(i) $P(\bar{X} > 16) = P(Z^* > 16) = 0.0581$

(ii) $P(14.5 < \bar{X} < 15.5) =$
 $= P(14.5 < Z^* < 15.5) = 0.5679$

```
dp <- 4.5/sqrt(50)
pnorm(16,15,dp,lower.tail=F)
[1] 0.05805087
pnorm(15.5,15,dp)-pnorm(14.5,15,dp)
[1] 0.5679416
```

Exercício: (nº 68) Considere a lei $Laplace(0, 1)$ com fdp $f(x) = \frac{1}{2} e^{-|x|}$, $x \in \mathbb{R}$. Considere $T \sim Laplace(0, 1)$ e ainda X e Y v.a. i.i.d. $Exp(1)$.

- (a) Prove que a transformada de Laplace de T é $L_T(t) = \frac{1}{1-t^2}$, $|t| < 1$.
- (b) Determine a t. Laplace de $X - Y$ e conclua que $X - Y$ e T têm igual distribuição.
- (c) Prove que $|T| \sim Exp(1)$.
- (d) Simule amostras da v.a. T utilizando
 - (i) a alínea (b)
 - (ii) a alínea (c)
- (e) Calcule o valor médio e variância da $Laplace(0, 1)$ por 3 processos diferentes.

Nota: Pela nota do slide 223, note-se que a f.d.p. da $Laplace(\lambda, \delta)$, dada por $f(x; \lambda, \delta) = \frac{1}{2\delta} e^{-|x-\lambda|/\delta}$, $x \in \mathbb{R}$, se escreve na forma $f(x) = \frac{1}{\delta} g\left(\frac{x-\lambda}{\delta}\right)$, com $g(x) = \frac{1}{2} e^{-|x|}$, $x \in \mathbb{R}$. Aqui g é a fdp de $T \sim Laplace(0,1)$ donde $W = \lambda + \delta T \sim Laplace(\lambda, \delta)$.

Resolução:

(a) A t.L de T é dada por (recorde-se que a t.L de $X \sim \text{Exp}(1)$ é $L_X(t) = \frac{1}{1+t}$, $t > -1$)

$$\begin{aligned} L_T(t) &= \int_{-\infty}^{+\infty} \frac{1}{2} e^{-tx} e^{-|x|} dx = \int_{-\infty}^0 \frac{1}{2} e^{-tx} e^x dx + \int_0^{+\infty} \frac{1}{2} e^{-tx} e^{-x} dx = \\ &= \int_0^{+\infty} \frac{1}{2} e^{-(-t)y} e^{-y} dy + \frac{1}{2} L_X(t) = \frac{1}{2} L_X(-t) + \frac{1}{2} L_X(t) = \frac{1}{2} \frac{1}{1-t} + \frac{1}{2} \frac{1}{1+t}, \\ &\quad \downarrow \qquad \qquad \downarrow \\ &\quad \text{se } t > -1 \qquad \text{se } -t > -1 \end{aligned}$$

donde $L_T(t) = \frac{1}{1-t^2}$, $|t| < 1$.

Resolução (cont.): (b) Determine a t. Laplace de $X - Y$ e conclua que $X - Y$ e T têm igual distribuição

(b) Como X e Y são independentes, então X e $-Y$ também são independentes, donde a t.Laplace de $X - Y$ é o produto das t.Laplace de X e de $-Y$. Então,

$$L_{X-Y}(t) = L_X(t)L_{-Y}(t) = L_X(t)L_Y(-t) = \frac{1}{1+t} \frac{1}{1-t}, \text{ se } t > -1 \text{ e } -t > -1, \text{ i.e., se } |t| < 1$$

↓
 X e $-Y$ também
são independentes

↓
 $L_{a+bY}(t) = e^{-at} L_Y(bt)$
(com $a = 0$ e $b = -1$)

Logo $L_{X-Y}(t) = \frac{1}{1-t^2}$, $|t| < 1$, ou seja, $X - Y$ e T têm a mesma t.L, donde se conclui finalmente que (visto que a t.Laplace caracteriza a distribuição) $X - Y \stackrel{d}{=} T$.

Resolução (cont.): (c) Prove que $|T| \sim \text{Exp}(1)$.

(c) Calcula-se a fd de $|T|$ pela definição de fd (note-se que $|T|$ é v.a. não negativa):

$$F_{|T|}(t) = P(|T| \leq t) = P(-t \leq T \leq t) = \begin{cases} 0 & , t < 0 \\ F_T(t) - F_T(-t) = 1 - \frac{1}{2}e^{-t} - \frac{1}{2}e^{-t} = 1 - e^{-t} & , t \geq 0 \end{cases}$$

Logo $F_{|T|}(t) = \begin{cases} 0 & , t < 0 \\ 1 - e^{-t} & , t \geq 0 \end{cases}$, que é a fd de uma v.a. $\text{Exp}(1)$.

Conclui-se portanto que $|T| \sim \text{Exp}(1)$.

Resolução (cont.):

(d) Simule amostras da v.a. T utilizando

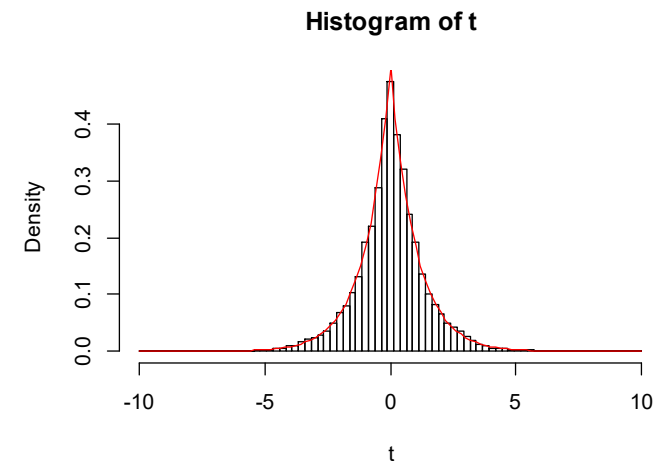
(ii) a alínea (b)

(iii) a alínea (c)

(d) Simulam-se amostras da v.a. T utilizando

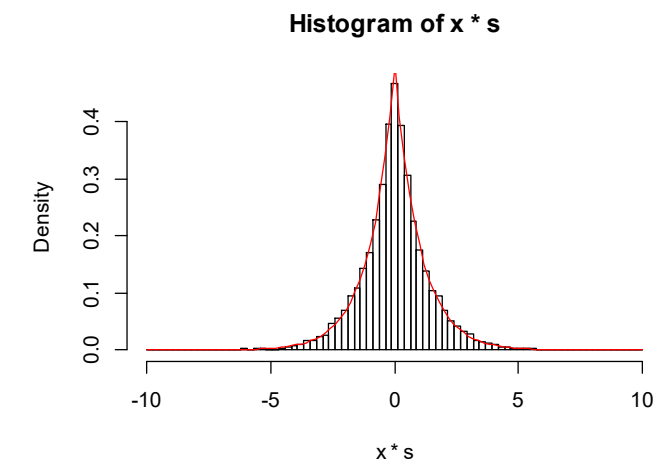
(ii) a alínea (b), i.e., $X - Y \stackrel{d}{=} T$

```
x <- rexp(10000,1); y <- rexp(10000,1)
t <- x-y
hist(t,breaks=seq(-10,10,len=80),freq=F)
curve(0.5*exp(-abs(x)),col=2,add=T)
```



(iii) a alínea (c), i.e., $|T| \sim \text{Exp}(1)$.

```
x <- rexp(10000,1);
s <- 2*rbinom(10000,1,0.5)-1
hist(x*s,breaks=seq(-10,10,len=80),freq=F)
curve(0.5*exp(-abs(x)),col=2,add=T)
```



Resolução (cont.):

(e) Calcule o valor médio e variância de T .

(e) Calculam-se os momentos $E(T)$ e $Var(T)$, com $T \sim Laplace(0,1)$.

Qual a maneira mais simples de resolver?

Por definição (integrando)? $E(T) = \int_{-\infty}^{+\infty} t f_T(t) dt = \dots$; $E(T^2) = \int_{-\infty}^{+\infty} t^2 f_T(t) dt = \dots$

usar o facto de f ser uma função par...

Pela t. Laplace (derivando)? $L_T(t) = \frac{1}{1-t^2}, |t| < 1 \Rightarrow \begin{cases} L_T'(t) = \dots & \Rightarrow E(T) = -L_T'(0) = \dots \\ L_T''(t) = \dots & \Rightarrow E(T^2) = L_T''(0) = \dots \end{cases}$

usar a fórmula dos momentos à custa das derivadas (no ponto 0) da t. Laplace

Ou de outra forma, ainda mais simples?

Usar a alínea (b) é ainda mais fácil... Solução: $E(T) = 0$ e $Var(T) = 2$

Exercício: Num processo de Poisson de intensidade λ , seja S_n o tempo de espera até à n -ésima chegada, i.e., $S_n = T_1 + \dots + T_n$, com T_1, \dots, T_n i.i.d. $Exp(\lambda)$, sendo T_i ($i = 1, 2, \dots$) os intervalos de tempo entre chegadas consecutivas.

Note-se que a n -ésima chegada ocorre depois do instante t (i.e., $S_n > t$) se e só se houver menos que n chegadas até ao instante t (i.e., $N_t < n$), ou seja, temos

$$S_n > t \quad \text{se e só se} \quad N_t < n$$

donde se conclui, uma vez que $N_t \sim \text{Poisson}(\lambda t)$,

$$P(S_n > t) = P(N_t < n) = \sum_{j=0}^{n-1} P(N_t = j) = \sum_{j=0}^{n-1} \frac{1}{j!} e^{-\lambda t} (\lambda t)^j$$

Deste resultado obtém-se a densidade de S_n por derivação:

$$f_{S_n}(t) = \frac{d}{dt} P(S_n \leq t) = -\frac{d}{dt} P(S_n > t) = -\frac{d}{dt} \sum_{j=0}^{n-1} \frac{1}{j!} e^{-\lambda t} (\lambda t)^j = \dots = \frac{1}{(n-1)!} \lambda^n e^{-\lambda t} t^{n-1}, t > 0$$

Estas densidades, $f_{S_n}(t) = \frac{1}{(n-1)!} \lambda^n e^{-\lambda t} t^{n-1}$, $t > 0$ são conhecidas pelo nome $Gama(n, \lambda)$. Note-se que n ($n > 0$) é um parâmetro de forma e $\delta = 1/\lambda$ ($\delta > 0$) é um parâmetro de escala. A distribuição $Gama(n, \lambda)$ é portanto a lei da soma de n v.a. i.i.d. $Exp(\lambda)$. Consequentemente, a sua transformada de Laplace é dada por

$$L_{S_n}(t) = \left(\frac{\lambda}{\lambda + t} \right)^n, |t| < \lambda$$

e os momentos de ordem $r = 0, 1, 2, \dots$ por

$$E(S_n^r) = \frac{(n+r-1)!}{(n-1)! \lambda^r}$$

No caso particular $n = 1$ temos o resultado $Gama(1, \lambda) \equiv Exp(\lambda)$, como era de esperar, pois $S_1 = T_1$.

