IV

Transformação uniformizante (e mais simulação).
Parâmetros de localização, escala e forma.
Momentos.
Desigualdades sobre momentos.
Convergências estocásticas.
Transformadas e aplicações.
Teorema Limite Central (TLC) e LGN.

Transformação uniformizante

Dada uma v.a. X com fd contínua $F(.)$, então $Y = F(X) \frown U[0,1]$, i.e., $F_Y(y) = \begin{cases} 0 & , y < 0 \\ y & ,0 \le y < 1 \\ 1 & , y \ge 1 \end{cases}$

Reciprocamente, se $Y \frown U[0,1]$, então $X = F^{-1}(Y)$ tem fd $F(.)$.

Demonstração: Note-se que sendo F contínua, o seu contradomínio é o intervalo de 0 a 1, donde temos $F_Y(y) = 0$, para $y < 0$ e $F_Y(y) = 1$ para $y > 1$. Determina-se então $F_Y(y)$, para $0 < y < 1$:

$$F_Y(y) = P(Y \le y) = P(F(X) \le y) \underset{\substack{\downarrow \\ F \text{ é não decrescente}}}{=} P(X \le F^{-1}(y)) = F(F^{-1}(y)) = y$$

Logo $Y = F(X) \frown U[0,1]$.

Reciprocamente, $P(X \le x) = P(F^{-1}(Y) \le x) = P(Y \le F(x)) = F(x)$

Nota: Este resultado está na origem do método de "inversão da fd" para geração de NPA's com dada distribuição contínua, a partir de NPA's uniformes.

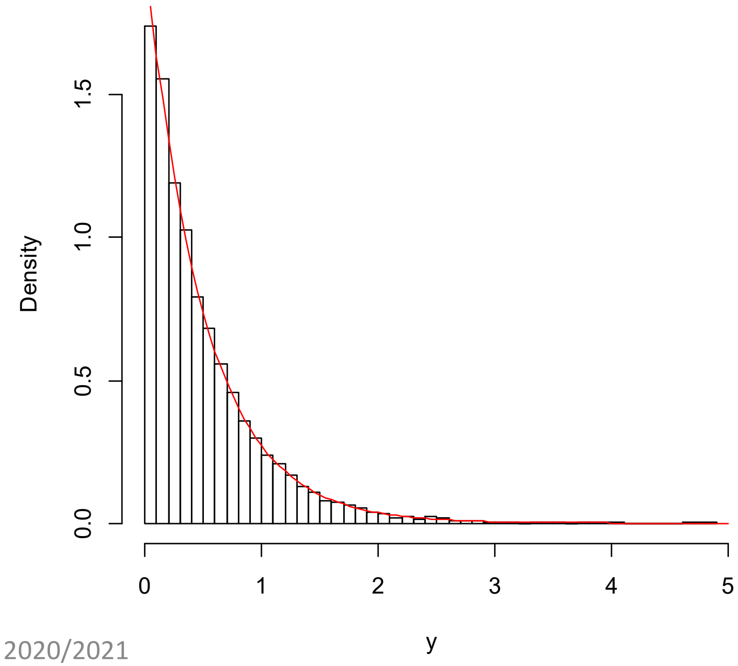
Exercício: (nº 62) Aplique a transformação uniformizante ao caso $X \frown Exp(\lambda)$, simulando uma amostra de dados $Exp(2)$, a partir de dados $U[0,1]$.

Resolução: Neste caso $F(x) = (1 - e^{-\lambda x})\, I_{[0,+\infty[}(x)$ é uma função contínua, donde $F(X) \frown U(0,1)$. Temos ainda $F^{-1}(y) = -\frac{1}{\lambda}\log(1-y),\ 0 < y < 1,$ pois

$$y = 1 - e^{-\lambda x} \Leftrightarrow e^{-\lambda x} = 1 - y \Leftrightarrow -\lambda x = \log(1-y) \Leftrightarrow x = -\tfrac{1}{\lambda}\log(1-y)$$

Logo, a partir de uma amostra de dados `y` com distribuição $U(0,1)$ obtemos uma amostra $Exp(\lambda)$ fazendo `x <- -log(1-y)/`λ

```
y <- runif(10000)
x <- -log(1-y)/2
hist(x,50,freq=F)
curve(dexp(x,2),0,col=2,add=T)
```



Parâmetros de localização, escala e forma.

Numa família (paramétrica) de v.a.'s 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.

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

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.

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

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

$Poisson(\theta)$

Famílias de localização-escala

Uma família de v.a.'s 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 (que se chamam parâmetros de localização e escala)

$N(\lambda,\delta)$

Nota: no caso de famílias de v.a.'s 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 $f(x;\theta) = \frac{1}{\theta}g(\frac{x}{\theta})$.

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

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).

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

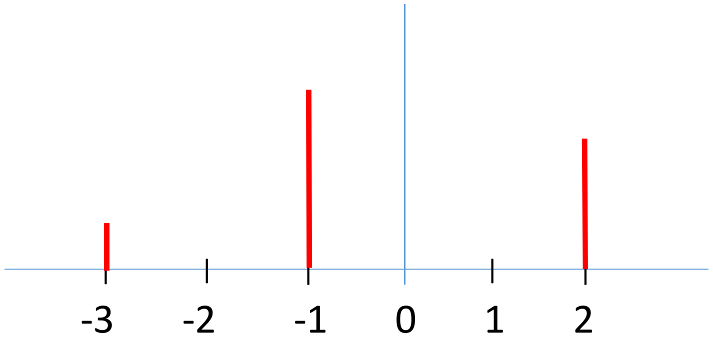
$$\beta_2 = E\left(\left(\frac{X-\mu}{\sigma}\right)^4\right)$$

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 mais, 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) 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

$$\mu = -\frac{3}{10} - \frac{5}{10} + \frac{8}{10} = 0$$

$$\sigma^2 = E(X^2) = \frac{9}{10} + \frac{5}{10} + \frac{16}{10} = 3$$

$$\mu_3 = E(X^3) = -\frac{27}{10} - \frac{5}{10} + \frac{32}{10} = 0$$

donde $\beta_1 = \frac{1}{\sigma^3} E((X-\mu)^3) = 0$

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

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

momento conjunto de ordem $(r\ ,\ s)$

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

momento conjunto central de ordem $(r\ ,\ s)$

covariância, para $r = s = 1$

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

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

Desigualdades sobre momentos

Teorema: Dada uma v.a. X e uma função não negativa, $g:\mathbb{R}\to\mathbb{R}$, tal que $E(g(X))$ exista, então, para qualquer $a>0$ temos

$$P\big(g(X)\geq a\big)\leq \frac{1}{a}E\big(g(X)\big)$$

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

Considerando $A=\{x: g(x)\geq a\}$, temos

$$E(g(X))=\int_{-\infty}^{+\infty} g(x)f(x)\,dx=\int_A g(x)f(x)\,dx+\int_{\bar{A}} \overbrace{g(x)f(x)}^{\geq 0 \text{ porque } g \text{ é não negativa}}\,dx\geq$$

$$\geq \int_A g(x)f(x)\,dx\geq \int_A a\,f(x)\,dx=a\,P(g(X)\geq a)$$

Desigualdades de Markov e de Chebyshev

Do teorema anterior decorrem as desigualdades de Markov e de Chebychev, tomando respectivamente $g(x) = |x|^r$, $a = a^r$ e $g(x) = (x - \mu)^2$, $a = (r\sigma)^2$ no teorema anterior,

$$P\left(|X| \geq a\right) \leq \frac{1}{a^r} E\left(|X|^r\right)$$ desigualdade de Markov

para $a > 0, r > 0$

$$P\left(|X - \mu| \geq r\sigma\right) \leq \frac{1}{r^2}$$ desigualdade de Chebyshev

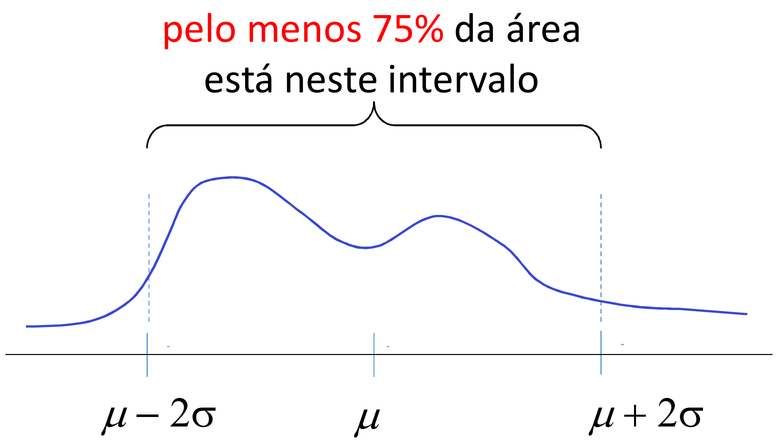
para $r > 0$

Da desigualdade de Chebychev decorre a LGN no caso de variância finita

A desigualdade de Chebychev pode também escrever-se na forma

$$P\left(|X-\mu|<r\sigma\right)\geq 1-\frac{1}{r^2}$$

que permite limitar a probabilidade de uma v.a. qualquer (seja qual for a sua distribuição de probabilidade) estar compreendida entre $\mu - r\sigma$ e $\mu + r\sigma$ (i.e., estar afastada de μ menos que r vezes o seu desvio padrão).
Por exemplo, no caso $r=2$ significa que pelo menos 75% da probabilidade está distribuída no intervalo $]\,\mu-2\sigma\ ,\ \mu+2\sigma\,[$.



Devido à sua grande generalidade, esta desigualdade é por vezes grosseira (tal como no caso $X \frown N(\mu, \sigma)$ em que $P(\mu-2\sigma<X<\mu+2\sigma)=0.9545$); noutros casos pode até ser atingida a igualdade na fórmula.

Exercício: (nº 66) Mostre que em n lançamentos de um dado equilibrado, a probabilidade de o nº de ases estar entre $\frac{n}{6}-\sqrt{n}$ e $\frac{n}{6}+\sqrt{n}$ é $\geq 31/36$

Resolução: O nº de ases é uma v.a. $bi(n,1/6)$, com valor médio $\mu = n/6$ e variância $\sigma^2 = 5n/36$. Pela desigualdade de Chebyshev, temos

$$P\left(\left| X-\tfrac{n}{6}\right| < r\frac{\sqrt{5n}}{6}\right)\geq 1-\tfrac{1}{r^2} \quad \text{donde} \quad P\left(\tfrac{n}{6}-r\frac{\sqrt{5n}}{6} < X < \tfrac{n}{6}+r\frac{\sqrt{5n}}{6}\right)\geq 1-\tfrac{1}{r^2}$$

Escolhendo $r=\frac{6}{\sqrt{5}}$ temos finalmente

$$P\left(\tfrac{n}{6}-\sqrt{n} < X < \tfrac{n}{6}+\sqrt{n}\right)\geq 1-\tfrac{5}{36}=\tfrac{31}{36}=0.86(1)$$

A probabilidade é
0.9845 para n=10
0.9925 para n=10^3
0.9927 para n=10^7

Convergências

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

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

Seja X_n uma sucessão de v.a.'s 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

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

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

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

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

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

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

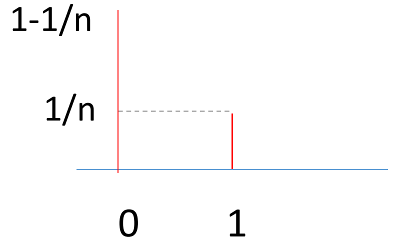
$$P(|X_n - X| > \varepsilon) \xrightarrow[n\to\infty]{} 0$$

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

(note-se que se trata da convergência da probabilidade do acontecimento $\{|X_n - X| > \varepsilon\}$ e não da convergência das v.a.'s X_n para a v.a. X)

Exemplo 23: Dadas as v.a.'s X_n discretas com fmp $X_n : \begin{cases} 0 & 1 \\ 1-\frac{1}{n} & \frac{1}{n} \end{cases}$

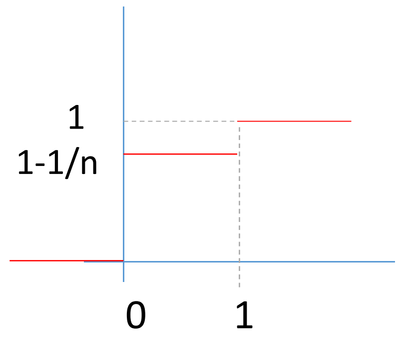
temos $P(|X_n| > \varepsilon) = \begin{cases} P(X_n = 1) = \frac{1}{n} & \text{se } 0 < \varepsilon < 1 \\ 0 & \text{se } \varepsilon \geq 1 \end{cases}$



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

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

Propriedades (convergência em lei e em probabilidade)

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

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

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

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

$\vdots$

Teoremas (convergência em lei e em probabilidade)

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

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

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

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 $E(X_n^r) = \frac{1}{n}\, n^r \not\to 0$

Transformadas e aplicações – ver “Slides T+TP – Parte 4.5”

transformada de Laplace, função geradora de momentos, função geradora de probabilidades, função característica

$$L(t) = E(e^{-tX})$$ 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, …}

$$\varphi(t) = E(e^{itX})$$ existe para $t \in \mathbb{R}$

Teorema Limite Central (TLC)

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

$\Big($analogamente, $\dfrac{\overline{X} - \mu}{\sigma}\sqrt{n} \xrightarrow{d} Z \frown N(0,1)$, sendo $\overline{X} = \frac{1}{n}(X_1 + X_2 \ldots + X_n)$ $\Big)$

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 $\overline{X}$ é aproximadamente $N(\mu, \frac{\sigma}{\sqrt{n}})$

---

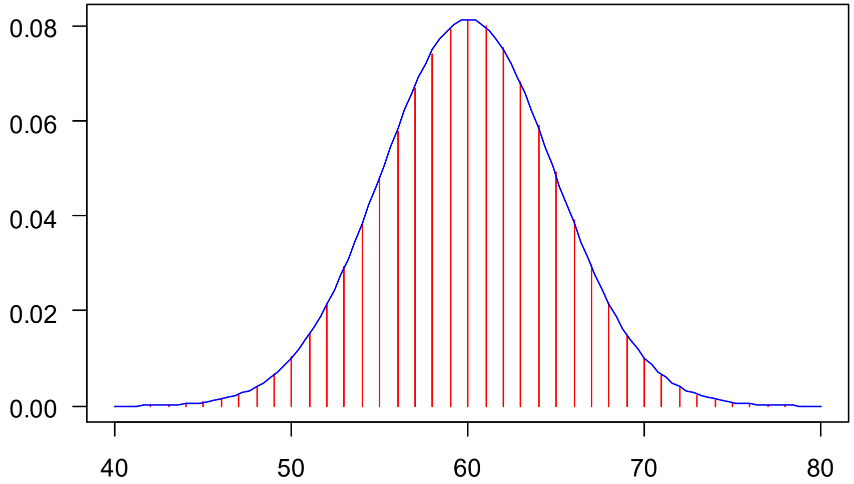
* se a distribuição de X não for muito assimétrica, geralmente $n = 30$ (ou até menos) é suficiente

TLC – caso particular $X \frown bi(1,p)$

O caso particular $X \frown bi(1,p)$ deve-se a A. de Moivre (c. 1773) e a Laplace.

Neste caso temos $S_n \frown bi(n,p)$ e então $\boxed{\dfrac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} Z}$

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



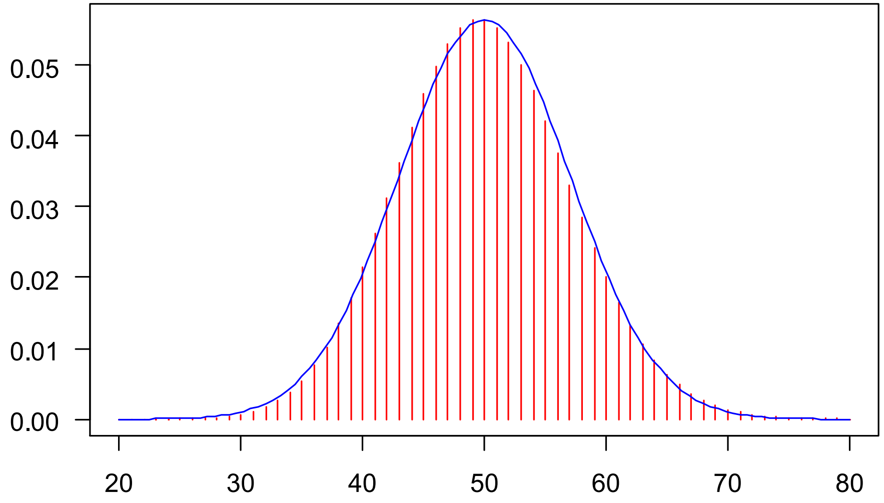
fmp $bi(100,0.6)$
fdp $N(60,\sqrt{24})$

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

TLC – caso particular *Poisson*

No caso $X \frown Poisson(\lambda)$ temos $S_n \frown Poisson(n\lambda)$ e então $\boxed{\dfrac{S_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow{d} Z}$

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



fmp $Poisson(50)$

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

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

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 $X \frown bi(\underbrace{30}_{n},\underbrace{0.5}_{p})$,

toma-se $P(11.5 < Y < 17.5)$, com $Y \frown N(\underbrace{15}_{np},\sqrt{\underbrace{7.5}_{np(1-p)}})$

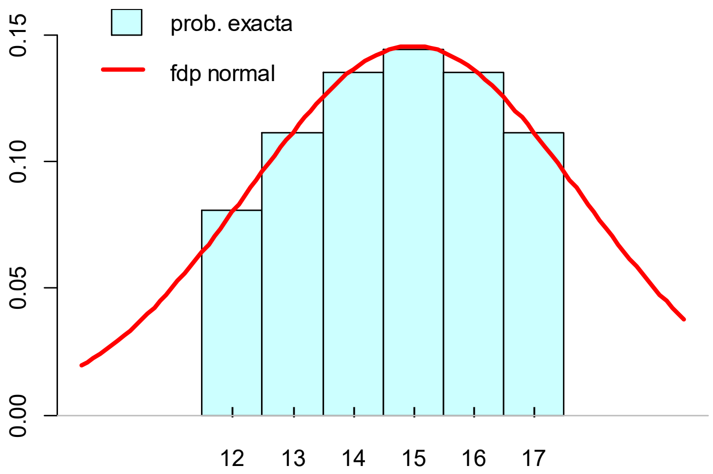
```
# probabilidade exacta:
pbinom(17,30,0.5)-pbinom(11,30,0.5)
[1] 0.7189585
# prob aproximada pelo TLC, sem correcção:
pnorm(17,15,sqrt(7.5))-pnorm(11,15,sqrt(7.5))
[1] 0.6953321
# prob aproximada pelo TLC, com correcção:
pnorm(17.5,15,sqrt(7.5))-pnorm(11.5,15,sqrt(7.5))
[1] 0.7187235
```

TLC – Correcção de continuidade

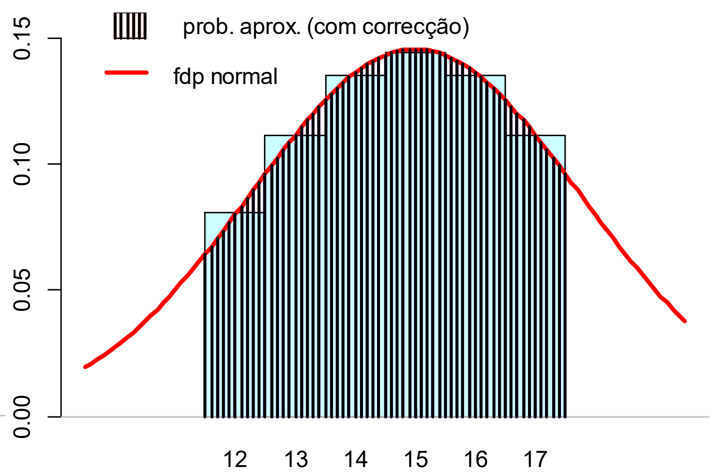
(ilustração)

$$X \frown bi(\underbrace{30}_{n},\ \underbrace{0.5}_{p})$$

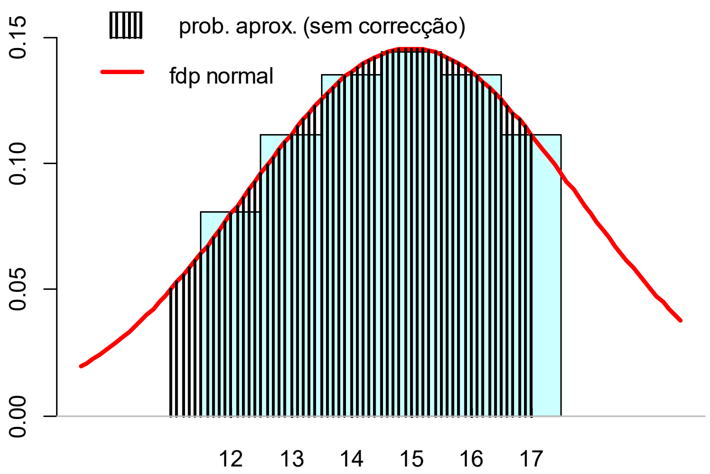
$$Y \frown N(\underbrace{15}_{np}, \sqrt{\underbrace{7.5}_{np(1-p)}})$$



$$\underbrace{P(12 \le X \le 17)}_{\textbf{0.7190}}$$

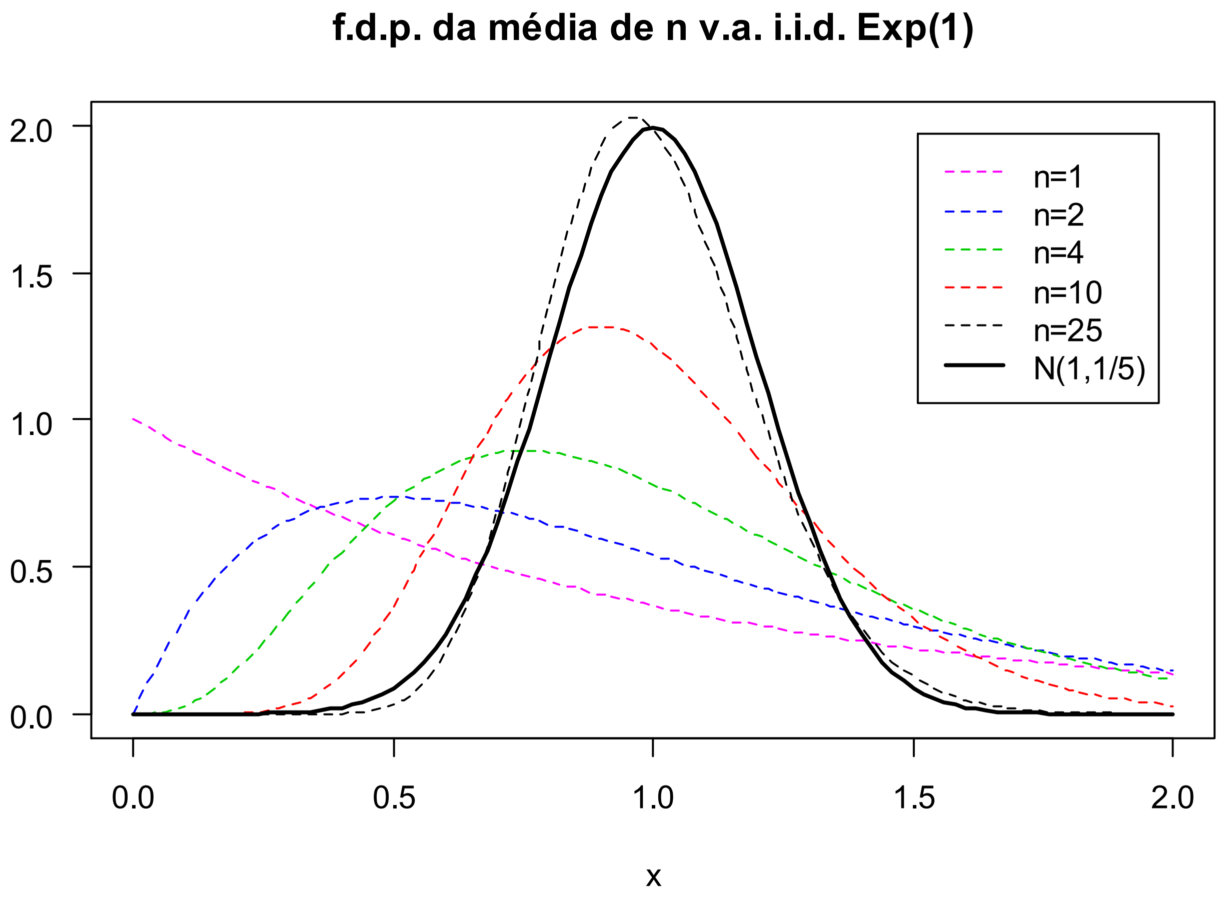


$$\underbrace{P(11.5 < Y < 17.5)}_{\textbf{0.7187}}$$



$$\underbrace{P(11 < Y \le 17)}_{\textbf{0.6953}}$$

TLC – ilustração no caso $X_i \frown Exp(\lambda)$, com $\lambda = 1$



Note-se que neste caso temos

$$S_n \frown Gama(n, \lambda)$$

$$\overline{X} \frown Gama(n, n\lambda)$$

Pelo TLC,

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

para n grande

Lei dos grandes números (LGN)

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

$$\forall \varepsilon > 0, \quad P\left(|\overline{X} - \mu| \leq \varepsilon\right) \underset{n\to\infty}{\to} 1$$

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

$$f_n(A) = \frac{\text{nº de vezes que } A \text{ ocorreu}}{n} = \frac{X_1 + \ldots + X_n}{n} = \overline{X}$$

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 $X_1, X_2, \ldots$ i.i.d., com $\mu = E(X)$ e $\mathrm{Var}(X) = \sigma^2 < +\infty$. Seja $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Então temos $E(\overline{X}) = \mu$ e $\mathrm{Var}(\overline{X}) = E\left((\overline{X}-\mu)^2\right) = \frac{\sigma^2}{n}$

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

$$\forall \varepsilon > 0,\ \ P\left(|\overline{X}-\mu| \le \varepsilon\right) \ \ge\ 1 - \frac{E((\overline{X}-\mu)^2)}{\varepsilon^2} = 1 - \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to\infty]{} 1$$

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

$^*P\left(|X| > a\right) \le P\left(|X| \ge a\right) \le \frac{1}{a^r} E\left(|X|^r\right)$ ou então $P\left(|X| \le a\right) \ge 1 - \frac{1}{a^r} E\left(|X|^r\right)$

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

$$\forall \varepsilon > 0,\ \lim_{n\to\infty} P\left(|\overline{X}-\mu|<\varepsilon\right) = \lim_{n\to\infty} P\left(-\varepsilon < \overline{X}-\mu < \varepsilon\right) =$$

$$= \lim_{n\to\infty} P\left(-\frac{\varepsilon\sqrt{n}}{\sigma} < \frac{\overline{X}-\mu}{\sigma}\sqrt{n} < \frac{\varepsilon\sqrt{n}}{\sigma}\right) \underset{TLC}{=}$$

$$= \lim_{n\to\infty} P\left(-\frac{\varepsilon\sqrt{n}}{\sigma} < Z < \frac{\varepsilon\sqrt{n}}{\sigma}\right) = 1$$

A LGN, no caso particular $X \frown bi(1, p)$, 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 *Khinchine* (1929).

O TLC, no caso particular $X \frown bi(1, p)$, é também conhecido por teorema de *de Moivre – Laplace*. No caso geral, 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.

TLC – demonstração parcial (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

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

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

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

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

donde $L_{Z_n}(t) = \left(L_U\left(\dfrac{t}{\sigma\sqrt{n}} \right) \right)^n = \left(1 + \dfrac{t^2}{2n} + o(\tfrac{1}{n}) \right)^n \xrightarrow[n\to\infty]{} \boxed{e^{t^2/2}}$ t. Laplace de $Z \frown N(0,1)$ (função contínua na origem)

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) concluímos que $Z_n \xrightarrow{d} Z$

Exercício: (nº 71) 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 $\sigma^2 = \mathrm{Var}(X) = 0.39 + 0.41 - 0.02^2 = 0.7996$.

Então $S_{700} = X_1 + \ldots + X_{700}$ (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 $Z_{700} \frown N(14, 559.72^{1/2})$, conforme segue adiante. Calculou-se também um valor aproximado de $P(S_{700} = 0)$, pelo TLC e por simulação.

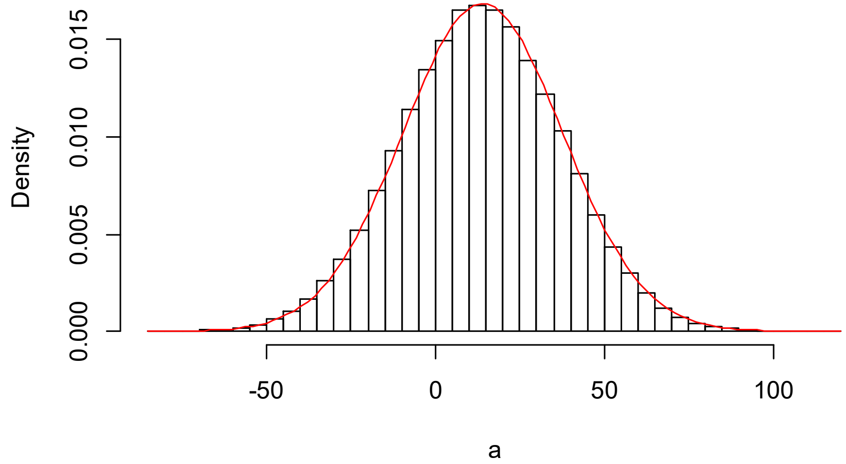
Resolução (cont.):

```
## 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)
```

$$E(S_{700}) = 700\mu = 700 \times 0.02 = 14$$

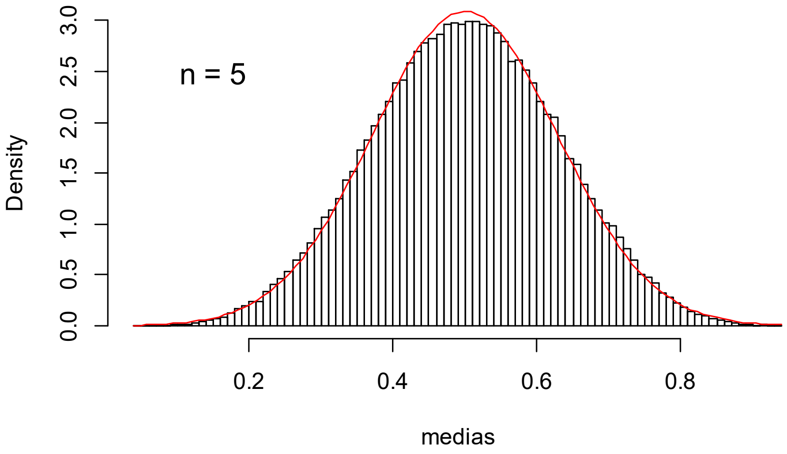
$$Var(S_{700}) = 700\sigma^2 = 700 \times 0.7996$$

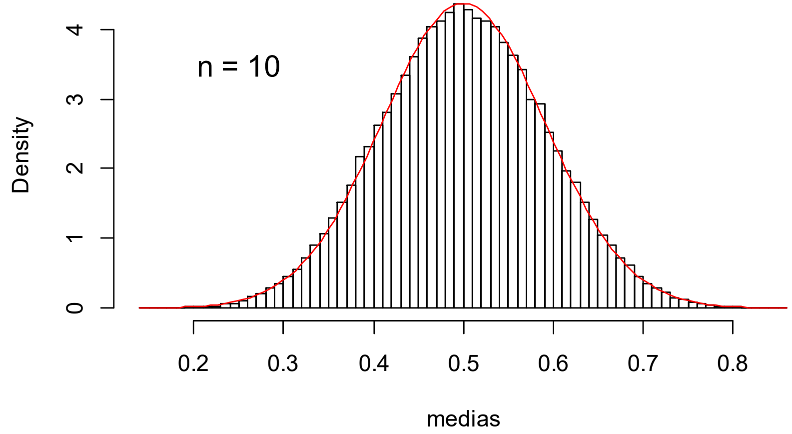
$$= 559.72$$



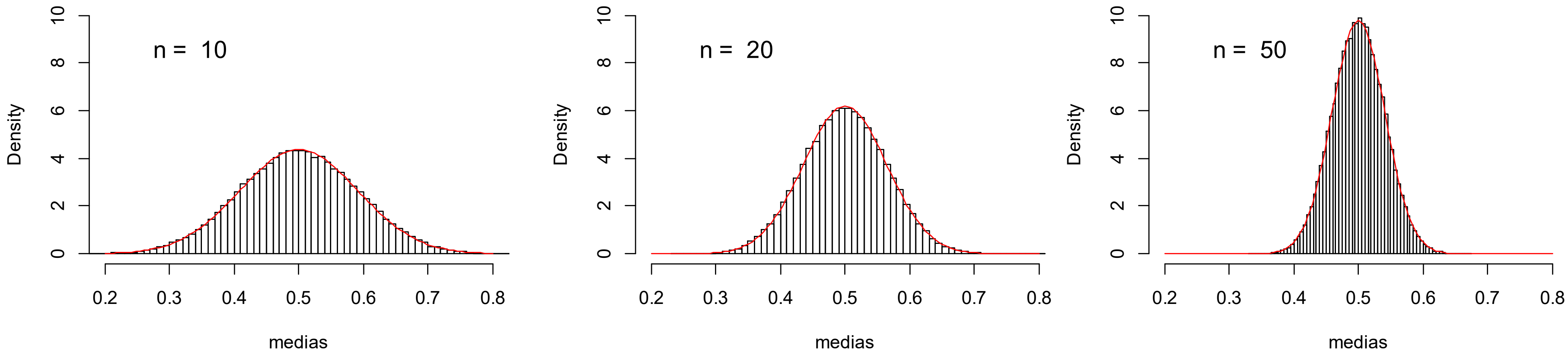
Exercício: Ilustre o TLC para a média de n v.a. i.i.d. $X_i \frown U[0,1]$, simulando e sobrepondo a curva normal (fdp) adequada, para $n = 5, 10, 30$.

```
# função pª simular médias de n valores U[0,1] com nº réplicas nrep
fmed <- function(n,nrep) colMeans(matrix(runif(n*nrep),nr=n))
medias <- fmed(5,10^5)
hist(medias,75,freq=F, main="")
curve(dnorm(x,0.5,sqrt(1/12/5)),add=T,col=2)
text(0.15,2.5,"n = 5",cex=1.3)
```





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 $\overline{X}$ é $\dfrac{\sigma}{\sqrt{n}} = \dfrac{1}{\sqrt{12\,n}} \xrightarrow[n\to\infty]{} 0$).

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 $\overline{X}$, com uma precisão que aumenta quando n aumenta.

Exercício: (nº 70) 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 a $Laplace(0,1)$, esta última com fdp $f(x) = \frac{1}{2} e^{-|x|}$, $x \in \mathbb{R}$. E com fmp par temos a $U\{-n, -n+1, \ldots, n-1, n\}$, *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

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

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

$$E(|X|^s) = \int_{-\infty}^{+\infty} |x|^s f(x)\,dx = \int_{|x|^s \le 1} |x|^s f(x)\,dx + \int_{|x|^s > 1} |x|^s f(x)\,dx \le$$

$$\le \int_{|x|^s \le 1} f(x)\,dx + \int_{|x|^s > 1} |x|^r f(x)\,dx \le P(|X|^s \le 1) + E(|X|^r) < +\infty$$

donde se conclui que existe o momento de ordem s.

Como consequência deste resultado, se 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 69) 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 $X \frown bi(1{,}0.5)$) com o resultado aproximado de $P(0 \leq Y_1 + \ldots + Y_{100} \leq 50)$, obtido pelo TLC com correcção de continuidade.

Resolução:

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

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

$$= 0.5398$$

```
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 $X_i \frown bi(1,0.5)$; aqui F é a fd $N(50, 5)$, pois $n\mu = np = 50$, $n\sigma^2 = np(1-p) = 25$;

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

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

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

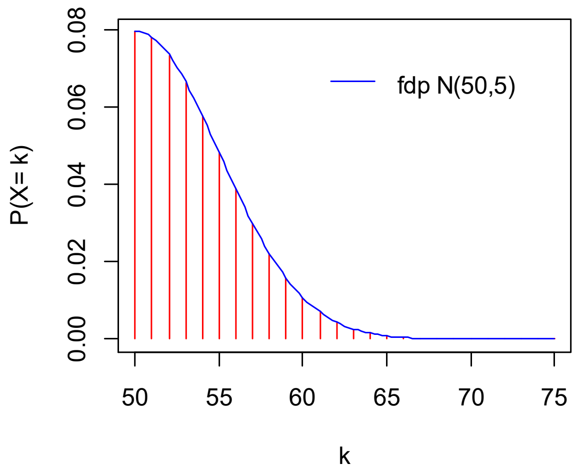
$$P(0 \le S \le 50) = G(51) - G(-1) = 0.5398$$

aqui G representa a fd $N(0, 10)$, pois $\mu = E(Y) = 0$, $\sigma^2 = \mathrm{Var}(Y) = E(Y^2) = 1$ donde $n\mu = 0$, $\sigma\sqrt{n} = 10$,

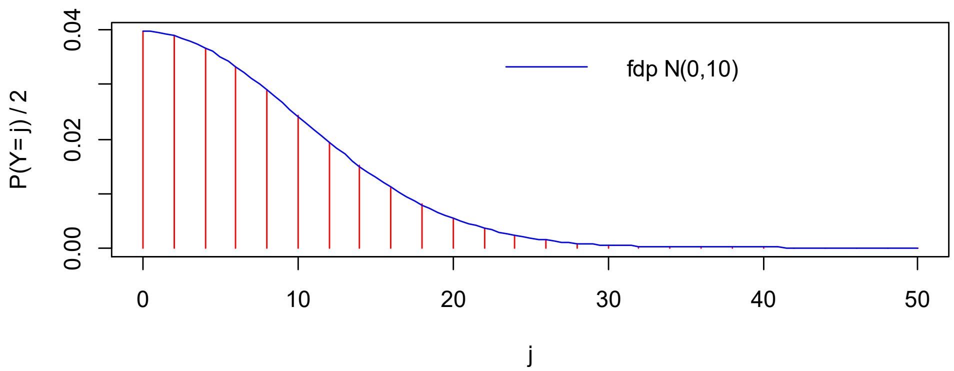
```
pnorm(51,0,10)-pnorm(-1,0,10)
[1] 0.5398277
```

Resolução (cont.): gráficos

(ii)



(iii)



Exercício: (nº 72) 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 $X_1, \ldots, X_n$ i.i.d. com valor médio $\mu = 15$ e desvio padrão $\sigma = 4.5$, com $n = 50$.

Pelo TLC, com a média $\overline{X} = \frac{1}{n}(X_1 + X_2 \ldots + X_n)$, temos $\dfrac{\overline{X} - \mu}{\sigma}\sqrt{n} \xrightarrow{d} Z \frown N(0,1)$

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

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

(ii) $P(14.5 < \overline{X} < 15.5) =$

$= P(14.5 < Z^* < 15.5) = 0.5679$

```
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 73) Considere a lei $Laplace(\mu, \delta)$ com fdp $f(x) = \frac{1}{2\delta} e^{-|x-\mu|/\delta}$, $x \in \mathbb{R}$, com $\mu \in \mathbb{R}$ e $\delta > 0$. Seja $T \frown Laplace(0,1)$ e X e Y v.a. i.i.d. $Exp(1)$.

(a) Mostre que a lei $Laplace(\mu, \delta)$ é uma família de localização-escala.

(b) Determine a transformada de Laplace de T.

(c) Determine a t. Laplace de $X - Y$ e conclua que $X - Y$ e T têm igual distribuição.

(d) Determine a fd de T e a correspondente fd inversa (função quantil).

(e) Prove que $|T| \frown Exp(1)$.

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

(i) o método de inversão da f.d.
(ii) a alínea (c)
(iii) a alínea (e)

(g) Como pode simular $W \frown Laplace(\mu, \delta)$ a partir de uma simulação de T?

(h) Calcule o valor médio e variância da $Laplace(\mu, \delta)$.

Resolução:

(a) Recorrendo à nota do slide 180, basta reparar que a f.d.p. $f(x)=\frac{1}{2\delta}e^{-|x-\mu|/\delta}$, $x\in\mathbb{R}$, se escreve na forma $f(x)=\frac{1}{\delta}g(\frac{x-\lambda}{\delta})$ com $g(x)=\frac{1}{2}e^{-|x|}$, $x\in\mathbb{R}$. Note-se que g é a fdp de $T\frown Laplace(0,1)$ e que $W=\mu+\delta\ T\frown Laplace(\mu\,,\,\delta)$.

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

$$L_T(t)=\int_{-\infty}^{+\infty}\frac{1}{2}e^{-tx}e^{-|x|}\,dx=\int_{-\infty}^{0}\frac{1}{2}e^{-tx}e^{x}\,dx+\int_{0}^{+\infty}\frac{1}{2}e^{-tx}e^{-x}\,dx=$$

$$=\int_{0}^{+\infty}\frac{1}{2}e^{-(-t)y}e^{-y}\,dy+\frac{1}{2}\underset{\substack{\downarrow\\ \text{se } t>-1}}{L_X(t)}=\frac{1}{2}\underset{\substack{\downarrow\\ \text{se } -t>-1}}{L_X(-t)}+\frac{1}{2}L_X(t)=\frac{1}{2}\frac{1}{1-t}+\frac{1}{2}\frac{1}{1+t},$$

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

Resolução (cont.):

(c) Determine a t. Laplace de $X-Y$ e conclua que $X-Y$ e T têm igual distribuição

(d) 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,

$$L_{X-Y}(t) = \underbrace{L_X(t)L_{-Y}(t)}_{X \text{ e } -Y \text{ também são independentes}} = \underbrace{L_X(t)L_Y(-t)}_{L_{a+bY}(t)=e^{-at}L_Y(bt) \text{ (com } a=0 \text{ e } b=-1)} = \frac{1}{1+t}\,\frac{1}{1-t}$$

, se $t>-1$ e $-t>-1$, i.e., se $|t|<1$

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

$$X-Y \overset{d}{=} T$$

Resolução (cont.):

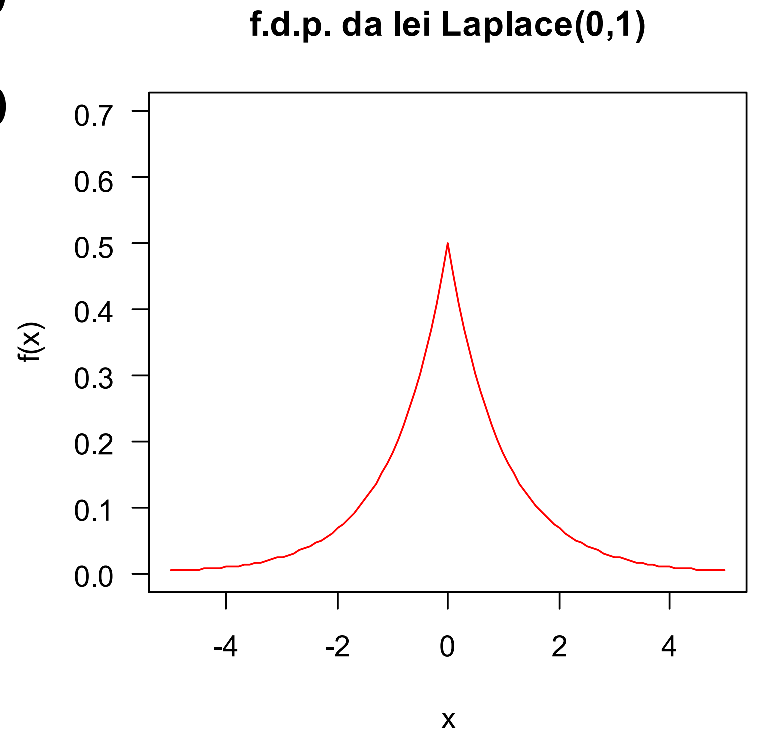
(d) Determine a fd de T e a correspondente fd inversa (função quantil).

(d) A fd de T é dada por

$$F_T(t) = P(T \le t) = \int_{-\infty}^{t} \frac{1}{2} e^{-|x|} dx = \begin{cases} \int_{-\infty}^{t} \frac{1}{2} e^{x} dx & , t < 0 \\ \underbrace{\frac{1}{2} + \int_{0}^{t} \frac{1}{2} e^{-x} dx}_{\frac{1}{2}+\frac{1}{2}(1-e^{-t})} & , t > 0 \end{cases}$$

$$\therefore \quad F_T(t) = \begin{cases} \frac{1}{2} e^{t} & , t < 0 \\ 1 - \frac{1}{2} e^{-t} & , t \ge 0 \end{cases}$$

```
curve(0.5*exp(-abs(x)),-5,5,ylim=c(0,0.7),
       ylab="f(x)",las=1,col=2)
title("f.d.p. da lei Laplace(0,1)")
```



Resolução (cont.):

(d) Determine a fd de T e a correspondente fd inversa (função quantil).

(d) A função inversa da fd F_T calcula-se resolvendo $p = F_T(x)$ em ordem a x :

Se $p < 0.5$, resolvemos $p = \frac{1}{2}e^{x}$

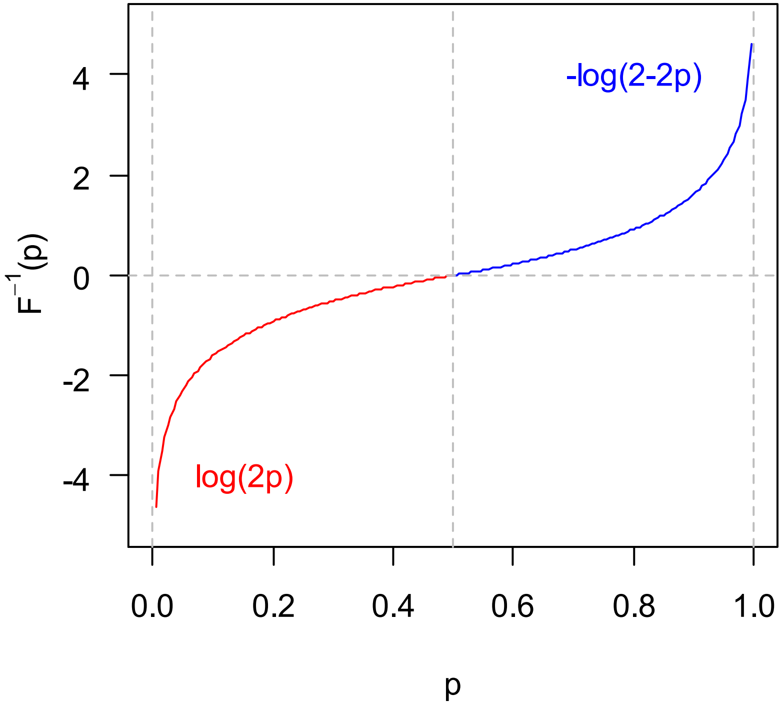
cuja solução é $x = \log(2p)$

Se $p > 0.5$, resolvemos $p = 1 - \frac{1}{2}e^{-x}$

cuja solução é $x = -\log(2(1-p))$

Conclui-se que

$$F_T^{-1}(p) = \begin{cases} \log(2p) & , 0 < p < 0.5 \\ -\log(2-2p) & , 0.5 \leq p < 1 \end{cases}$$



Resolução (cont.): (e) Prove que $|T| \frown Exp(1)$.

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

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

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

Conclui-se portanto que $|T| \frown Exp(1)$.

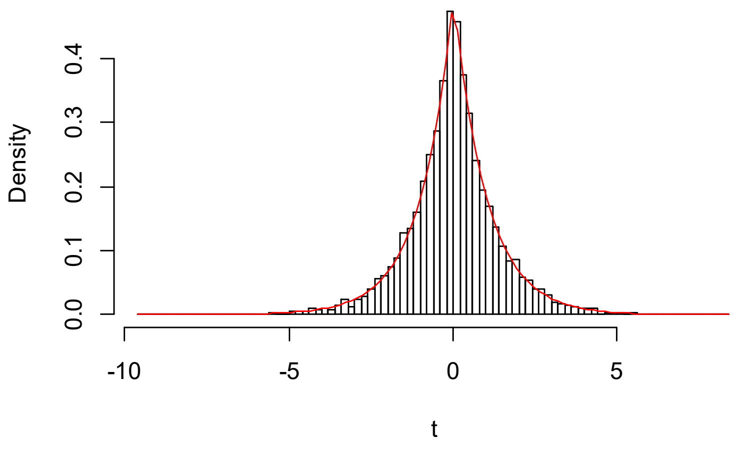
(f) Simulam-se amostras da v.a. T utilizando

(i) o método de inversão da fd : $U \frown U[0,1] \Rightarrow F^{-1}(U)$ tem fd F

Simula-se $U \frown U[0,1]$ e calcula-se $F^{-1}(U)$, com $F^{-1}(u) = \begin{cases} \log(2u) & ,u < 0.5 \\ -\log(2-2u) & ,u \geq 0.5 \end{cases}$

Note que $F^{-1}(u) = s \log(\ (1-s) + 2\ s\ u\ )$, com $s = \text{sgn}(0.5-u) = \begin{cases} +1 & ,u < 0.5 \\ -1 & ,u > 0.5 \end{cases}$

```
#### simulação de T pelo método de inversão da fd:
u <- runif(10000)
s <- sign(0.5-u)
t <- s*log((1-s)+ 2*s*u)
hist(t,100, freq=F)
curve(0.5*exp(-abs(x)),col=2,add=T)
## variante com um ciclo "for":
t <- 0
for (i in 1:10000)
   {p <- runif(1);
    if (p<0.5) t[i]<-log(2*p) else t[i]<- -log(2-2*p)}
hist(t,70,freq=F)
curve(0.5*exp(-abs(x)),col=2,add=T)
```



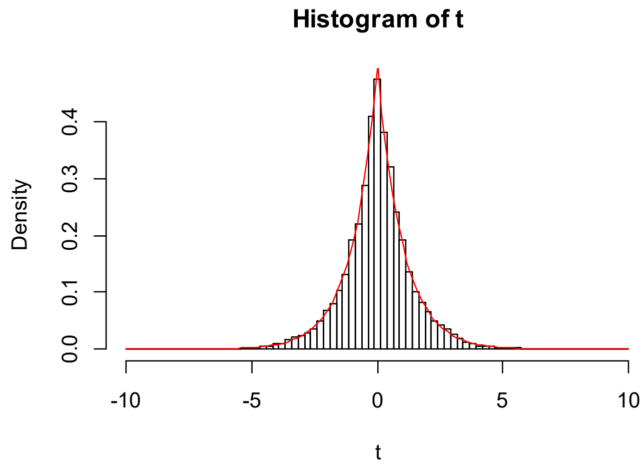
Resolução (cont.):

(f) Simule amostras da v.a. T utilizando (ii) a alínea (c) (iii) a alínea (e)

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

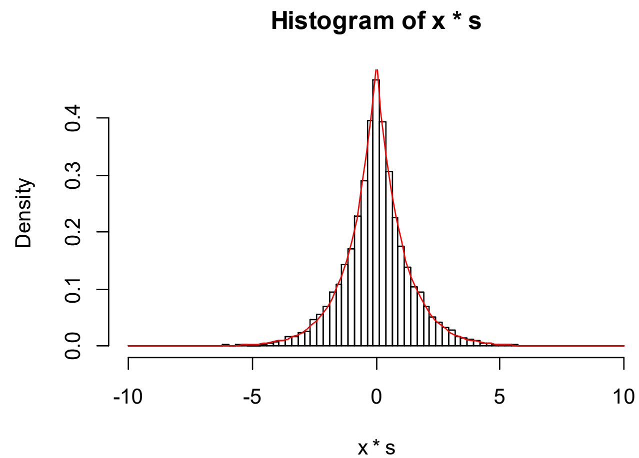
(ii) a alínea (c), i.e., $X - Y \stackrel{d}{=} T$

```
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 (e), i.e., $|T| \frown Exp(1)$.

```
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.):

(g) Como pode simular $W \frown Laplace(\mu, \delta)$ a partir de uma simulação de T?

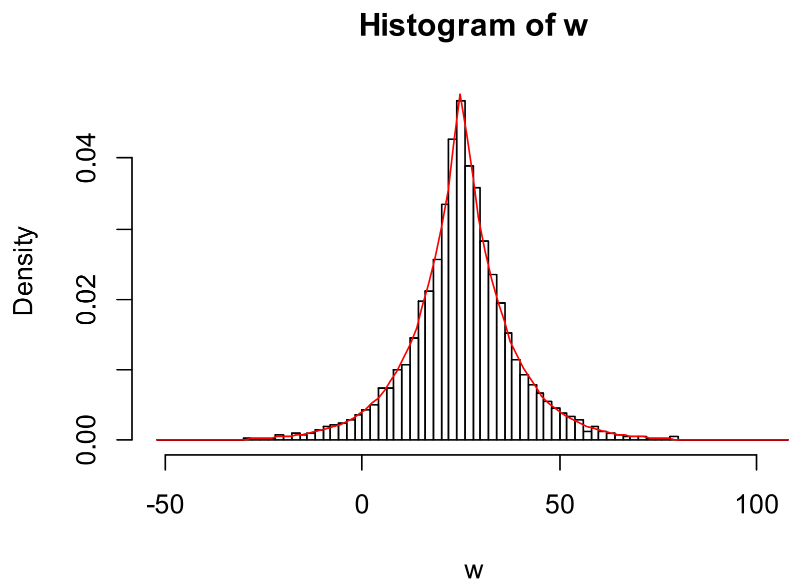
(g) Simulação de $W \frown Laplace(\mu, \delta)$ à custa de $T \frown Laplace(0,1)$: $W = \mu + \delta\, T$

Por exemplo, para simular $W \frown Laplace(25, 10)$, a partir dos dados `t` simulados em (f) :

```
m <- 25; d <- 10
w <- m + d*t
hist(w,70,freq=F)
curve(0.5/d*exp(-abs(x-m)/d),col=2,add=T)
```

Ou com o package `extraDistr`, usando o comando `rlaplace(n, mu = 0, sigma = 1)`:

```
library(extraDistr)
w <- rlaplace(10000, m, d)
hist(w,70,freq=F)
curve(0.5/d*exp(-abs(x-m)/d),col=2,add=T)
```



Resolução (cont.): (h) Calcule o valor médio e variância da $Laplace(\mu, \delta)$.

(h) Temos $E(W) = E(\mu + \delta\, T) = \mu + \delta\, E(T)$ e $Var(W) = Var(\mu + \delta\, T) = \delta^2\, Var(T)$.

Basta então calcular os momentos $E(T)$ e $Var(T)$, com $T \frown Laplace(0,1)$.

Qual a maneira mais simples de resolver?

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

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

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

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 (c)...

Solução: $E(T) = 0$ e $Var(T) = 2$