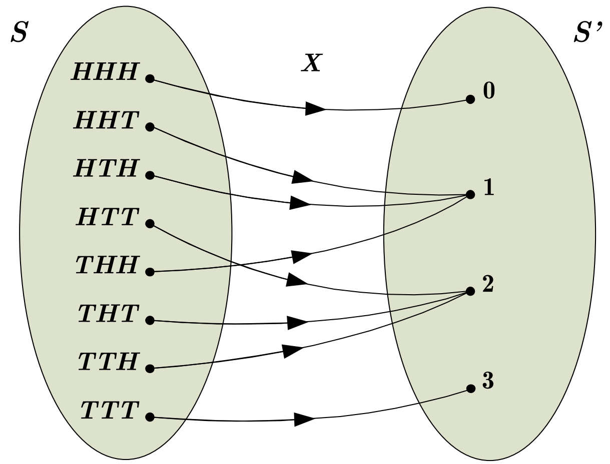
3.1. Variáveis Aleatórias

Variáveis que expressam os resultados de uma experiência aleatória.

- Uma variável aleatória é o resultado de uma **aplicação de um espaço amostral (S) num conjunto de chegada qualquer (S')**. **A cada resultado de S corresponde um único valor no conjunto de chegada** (variável aleatória), a diferentes valores de S pode corresponder o mesmo valor no conjunto de chegada.



X - nº de "caras" em 3 lançamentos de uma moeda

- Denotam-se, em geral, por uma **letra maiúscula** $(X,\ Y,\ Z,\ \ldots)$.
- As variáveis aleatórias expressas numa *escala quantitativa* classificam-se em **discretas ou contínuas, consoante o conjunto de valores que possam tomar seja numerável ou não numerável**. As variáveis do *primeiro tipo* encontram-se normalmente associadas a *contagens e as do segundo a medidas*.
- Em algumas situações, conjunto de valores que a variável aleatória toma *pode confundir-se* com o espaço amostral – S.

Exemplo

Experiência aleatória – medição do peso de uma pessoa
Espaço amostral – conjunto de todos os pesos
Variável aleatória – peso

- Noutras situações, **o conjunto de valores** que a variável aleatória toma **distingue-se claramente dos resultados**, correspondendo a uma transformação destes.

Exemplo

Experiência aleatória – três lançamentos de uma moeda equilibrada

Espaço amostral $-S = \{HHH,\ HHT,\ HTH,\ HTT,\ THH, THT,\ TTH,\ TTT\}$

Variável aleatória – número de "caras"

DGI 2019

3.2. Distribuições de Variáveis Discretas

3.2.1. *Função de Probabilidade*

Seja Y uma variável discreta. Chama-se função de Probabilidade da variável Y e designa-se por $p(y)$ a **função que associa a cada valor de y a probabilidade de Y ser igual a y.**

$$p(y) = P(Y = y)$$

Dos axiomas de probabilidade resulta que:

P1: $0 \leq p(y) \leq 1$

P2: $\sum_{y} p(y) = 1$

3.2.2. Função Distribuição de Probabilidade

Chama-se Função Distribuição de Probabilidade de Y e designa-se por $F(y)$ **a função que associa a cada valor de y a probabilidade de Y ser menor ou igual a y.**

$$F(y) = P(Y \leq y) = \sum_{Y \leq y} p(y)$$

Como resultado da sua definição, a Função Distribuição de Probabilidade satisfaz as seguintes propriedades:

P1: $0 \leq F(y) \leq 1$

P2: $P(a < Y \leq b) = F(b) - F(a)$

As Funções de Probabilidade e de Distribuição de Probabilidade podem representar-se através de tabelas ou diagramas de barras.

Exemplo

A variável Y traduz o número de dias com precipitação igual ou superior a 10 mm registados no mês de Agosto em Bragança, sendo conhecida a respectiva função de probabilidade $p(y)$.

y	$p(y)$
0	k
1	$2k$
2	$5k$
3	$5k$
4	$5k$
5	k
6	k
$\sum$	1

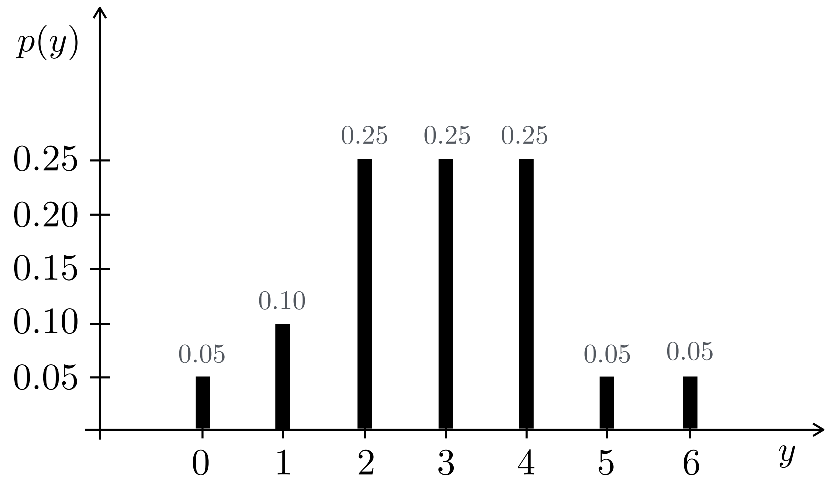
$$\sum_y p(y) = 1 \Leftrightarrow$$

$$\Leftrightarrow k + 2k + 5k + 5k + 5k + k + k = 1 \Leftrightarrow$$

$$\Leftrightarrow k = 0.05$$

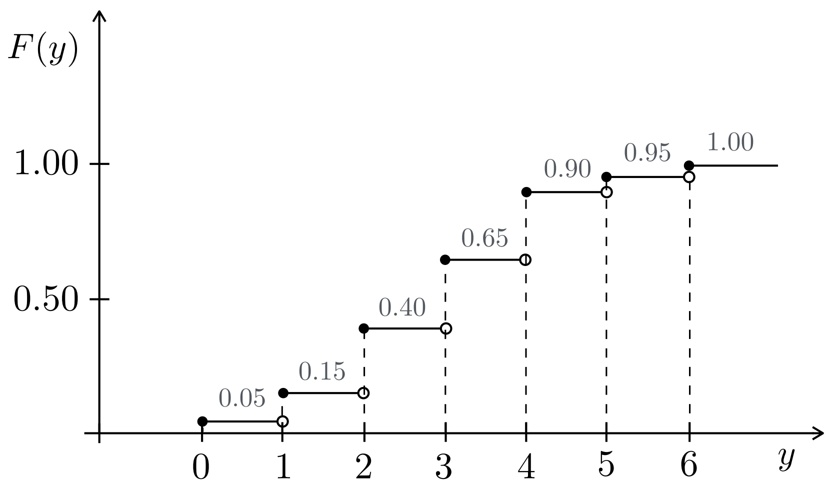
y	$p(y)$	$F(y)$
0	0.05	0.05
1	0.1	0.15
2	0.25	0.4
3	0.25	0.65
4	0.25	0.9
5	0.05	0.95
6	0.05	1
$\sum$	1	

Função de Probabilidade



$$p\left(y\right)=\begin{cases}0.05, & y=0\\0.10, & y=1\\0.25, & y=2\\0.25, & y=3\\0.25, & y=4\\0.05, & y=5\\0.05, & y=6\end{cases}$$

Função Distribuição de Probabilidade



$$F\left(y\right)=\begin{cases}0.00, & y<0\\0.05, & 0\leq y<1\\0.15, & 1\leq y<2\\0.40, & 2\leq y<3\\0.65, & 3\leq y<4\\0.90, & 4\leq y<5\\0.95, & 5\leq y<6\\1.00, & y\geq 6\end{cases}$$

3.3. Distribuições de Variáveis Contínuas

3.3.1. Função Densidade de Probabilidade

Seja X uma **variável aleatória que toma valores num espaço amostral contínuo**.

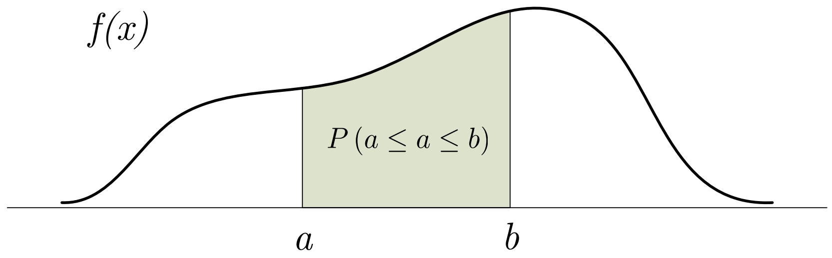
Chama-se Função Densidade de Probabilidade da variável X e designa-se por $f(x)$ uma função tal que:

$$P\left(x \in [x, x + dx]\right) = f(x)\, dx$$

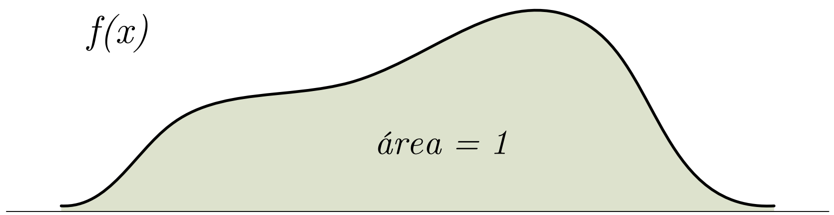
Dos axiomas da probabilidade e da definição de $f(x)$ resulta que:

P1: $f(x) \geq 0$

P2: $P(a \leq x \leq b) = P(a < x < b) = \int_a^b f(x)\, dx$



P3: $\int_{-\infty}^{+\infty} f(x)\, dx = 1$



3.3.2. *Função Distribuição de Probabilidade*

Chama-se Função Distribuição de Probabilidade de X e designa-se por $F(x)$ a função que associa a cada valor de x a probabilidade de X ser menor ou igual a x.

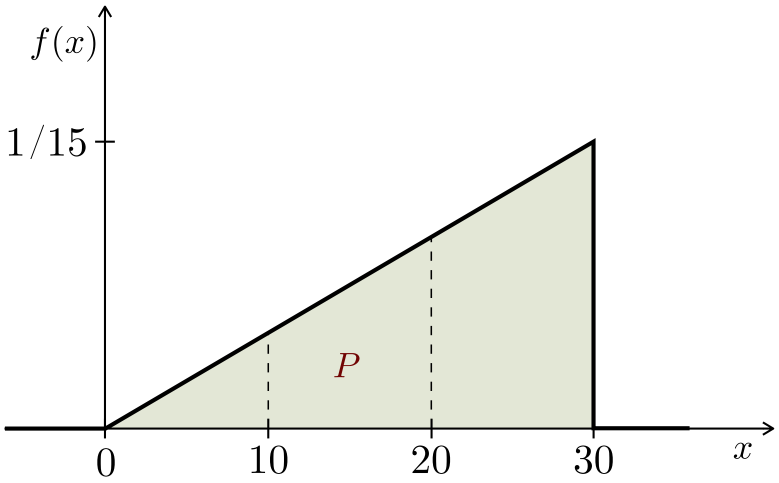
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\ du$$

Da definição de $F(x)$ resulta que:

P1: $\lim\limits_{x \to +\infty} F(x) = 1$

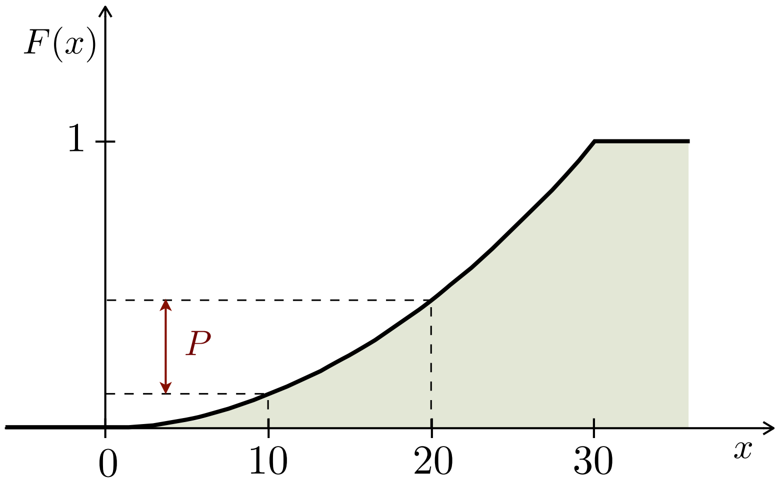
P2: $\dfrac{dF(x)}{dx} = f(x)$

Exemplo



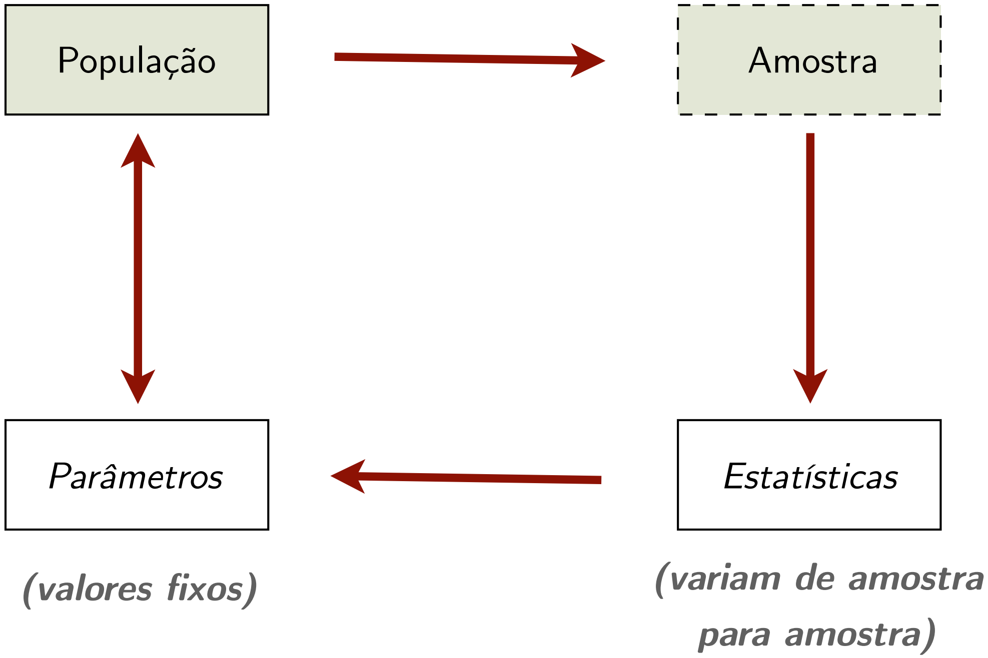
$$f(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{450}x, & 0 \le x \le 30 \\ 0, & x > 30 \end{cases}$$

$P = P(10 \le x \le 20)$



$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{900}x^2, & 0 \le x \le 30 \\ 1, & x > 30 \end{cases}$$

3.4. Parâmetros das Distribuições



3.4.1. Parâmetros de Localização

	Variáveis Discretas	Variáveis Contínuas
Valor esperado	$\mu_Y = E(Y) = \sum_y y \times p(y)$	$\mu_X = E(X) = \int_{-\infty}^{+\infty} x \times f(x)\, dx$
Mediana	$F[Y = \min(Y)] \geq 0.5$ $\eta_Y = \min(Y)$	$\int_{-\infty}^{\eta_X} f(x)\, dx = 0.5$
Moda	ξ_Y: valor(es) de Y para o(s) qual(is) $p(y)$ é máximo.	ξ_X:valor(es) de X para o(s) qual(is) $f(x)$ é máximo.

3.4.2. Parâmetros de Dispersão

	Variáveis Discretas	Variáveis Contínuas
Variância*	$\sigma^2{}_Y = VAR(Y) = \sum_{Y} (y - \mu_Y)^2 \times p(y)$	$\sigma^2{}_X = VAR(X) = \int_{-\infty}^{+\infty} (x - \mu_X)^2 \times f(x)\, dx$
Desvio Padrão	$\sigma_Y = \sqrt{\sigma^2{}_Y}$	$\sigma_X = \sqrt{\sigma^2{}_X}$

* Em alternativa podem utilizar-se as seguintes expressões simplificadas no cálculo da variância de variáveis discretas e contínuas:

$$\sigma^2{}_Y = \sum_{Y} y^2 \times p(y) - \mu_Y{}^2 \quad \text{e} \quad \sigma^2{}_X = \int_{-\infty}^{+\infty} x^2 \times f(x)\, dx - \mu_X{}^2$$

3.4.3. *Outros Parâmetros*

	Variáveis Discretas	Variáveis Contínuas
Momentos na Origem	$\mu'_i = \sum_y y^i \times p(y)$	$\mu'_i = \int_{-\infty}^{+\infty} x^i \times f(x)\, dx$
Momentos Centrados	$\mu_i = \sum_y (y-\mu_Y)^i \times p(y)$	$\mu_i = \int_{-\infty}^{+\infty} (x-\mu_X)^i \times f(x)\, dx$
Coeficiente de Assimetria	$\tau = \dfrac{\sum_Y (y-\mu_Y)^3 \times p(y)}{{\sigma_Y}^3} = \dfrac{\mu_3}{\sigma^3}$	$\tau = \dfrac{\int_{-\infty}^{+\infty} (x-\mu_X)^3 \times f(x)\, dx}{{\sigma_X}^3} = \dfrac{\mu_3}{\sigma^3}$
Coeficiente de Achatamento	$\kappa = \dfrac{\sum_Y (y-\mu_Y)^4 \times p(y)}{{\sigma_Y}^4} - 3 = \dfrac{\mu_4}{\sigma^4} - 3$	$\kappa = \dfrac{\int_{-\infty}^{+\infty} (x-\mu_X)^4 \times f(x)\, dx}{{\sigma_X}^4} - 3 = \dfrac{\mu_4}{\sigma^4} - 3$

Relações de interdependência entre momentos centrados (μ_i) e momentos na origem (μ'_i)

$$\mu_2 = \mu'_2 - (\mu'_2)^2$$

$$\mu_3 = \mu'_3 - 3 \times \mu'_2 \times \mu'_1 + 2 \times (\mu'_1)^3$$

$$\mu_4 = \mu'_4 - 4 \times \mu'_3 \times \mu'_1 + 6 \times \mu'_2 \times (\mu'_1)^2 - 3 \times (\mu'_1)^4$$

3.5. Variáveis Transformadas

3.5.1. *Transformação Linear*

Seja Z uma variável qualquer (discreta ou contínua) à qual se aplica uma transformação linear.

$W = a + b \times Z$	$\mathrm{E}(W) = a + b \times E(Z)$
	$\mathrm{VAR}(W) = b^2 \times VAR(Z)$

Exemplo

Calcule o valor esperado e a variância da variável Y, obtida a partir de uma outra variável X, de acordo com a seguinte expressão: $Y = 2X + 4$

sendo: $\mu_X = 4 \quad e \quad {\sigma^2}_X = 0.5$

R: $\mu_Y = 12 \quad e \quad {\sigma^2}_Y = 2$

DGI
2019

3.5.2. *Transformação Não Linear*

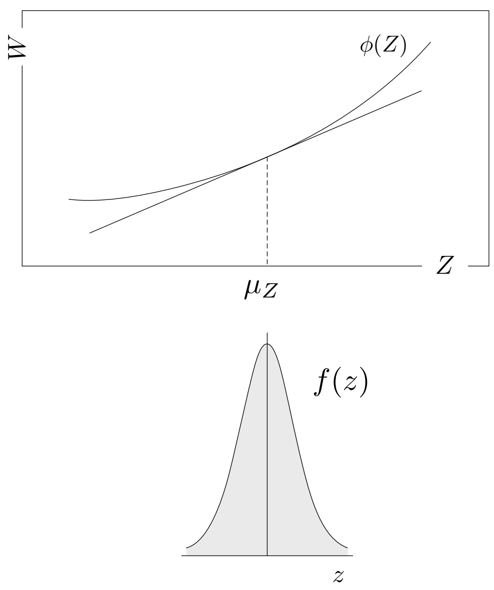
Seja Z uma variável qualquer (discreta ou contínua) à qual se aplica uma transformação não linear.

$W = \phi(Z)$	$E(W) = E\left[\phi(Z)\right] \approx \phi\left[E(Z)\right]$
$\phi(Z)$ função não linear	$VAR(W) = VAR\left[\phi(Z)\right] \approx \left(\frac{d\phi}{dz}\,\|_{Z=\mu_Z}\right)^2 \times VAR(Z)$

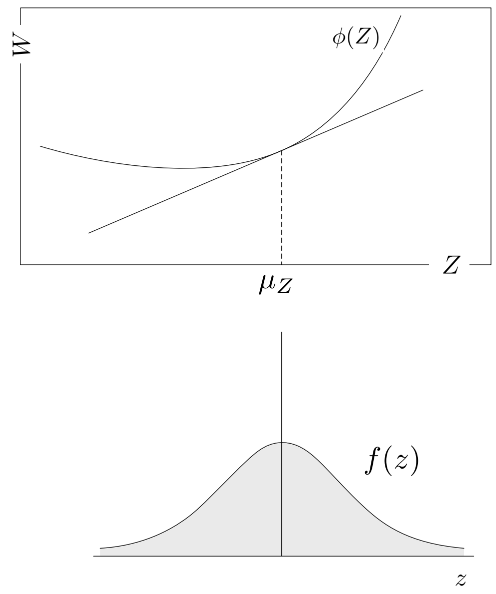
Desenvolvimento em Série de Taylor da função $\phi(Z)$ em torno de $E(Z)$:

$$\phi(Z) \approx \phi\left[E(Z)\right] + (Z - E(Z)) \times \left[\frac{\partial \phi}{\partial Z}\right]_{Z=E(Z)}$$

A aproximação será tanto *melhor* quanto *menor* for a *curvatura da função* $\phi(Z)$ *em torno de* μ_Z e quanto *menor for a dispersão da distribuição da variável* Z.



(aproximação linear mais adequada)



(aproximação linear menos adequada)

Exemplo

Considere a variável X com $\mu_X = 4 \quad e \quad {\sigma^2}_X = 0.5$. Calcule o valor esperado e a variância de: $T = 2X^2$

R: $\mu_T = 32 \quad e \quad {\sigma^2}_T = 128$

3.6. Combinação Linear de Variáveis Independentes

Considere-se a variável W definida através de uma combinação linear de duas variáveis independentes X e Y, contínuas ou discretas:

$$W = a + b \times X + c \times Y$$

Os parâmetros da sua distribuição de probabilidade podem ser calculados de duas formas distintas:

- Diretamente a partir da sua distribuição de probabilidade;
- Com base nas distribuições de probabilidade das variáveis que lhe deram origem e na combinação linear dessas variáveis.

$$\mu_W = a + b \times \mu_X + c \times \mu_Y$$

$$\sigma_W^2 = b^2 \times \sigma_X^2 + c^2 \times \sigma_Y^2$$

Generalização

Seja W uma variável aleatória definida a partir de uma combinação linear de N variáveis independentes

$$W = a + b_1 \times X_1 + b_2 \times X_2 + \ldots + b_N \times X_N$$

Os parâmetros da sua distribuição de probabilidade são dados por:

$$\mu_W = a + b_1 \times \mu_{X_1} + b_2 \times \mu_{X_2} + \ldots + b_N \times \mu_{X_N}$$

$$\mu_W = b_1^2 \times \sigma_{X_1}^2 + b_2^2 \times \sigma_{X_2}^2 + \ldots + b_N^2 \times \sigma_{X_N}^2$$

3.7. Geração de Observações Aleatórias com recurso à Técnica de Monte Carlo

3.7.1. *Introdução*

A geração de observações aleatórias envolve duas operações a realizar sequencialmente:

(1) *Geração de números aleatórios* seguindo uma distribuição uniforme $\mathrm{U}(0, 1)$

- *Método da "semente"*
- *Recurso a tabelas de números aleatórios*

(2) *Transformação daqueles números* noutros igualmente aleatórios, seguindo a distribuição em causa.

3.7.2. Geração de observações que envolvem Variáveis Discretas

- Gerar n números aleatórios Z_i $(i = 1, 2, \ldots, n)$ distribuídos segundo $U(0, 1)$;
- Dividir o intervalo $[0, 1]$ em sub intervalos de dimensões proporcionais aos valores de $p(x)$;
- Fazer corresponder a cada valor gerado Z_i, o valor X_i em conformidade com o intervalo a que Z_i pertença.

Exemplo

Pretende-se gerar uma amostra aleatória de dimensão $n = 8$ a partir de uma população com distribuição definida através da função de probabilidade seguinte:

x	$p(x)$
0	0.25
1	0.50
2	0.25

DGI
2019

3.7.3. Geração de observações que envolvem Variáveis Contínuas

- Gerar *n* números aleatórios Z_i $(i = 1, 2,..., n)$ distribuídos segundo $\mathrm{U}(0, 1)$;
- Transformação dos valores recorrendo à inversa de $F(x)$.

Exemplo

Pretende-se gerar uma amostra aleatória de dimensão $n = 3$ a partir de uma população com distribuição exponencial negativa $(\beta = 2)$