Capítulo 3. Variáveis aleatórias e distribuições discretas

Conceição Amado, Ana Pires e M. Rosário Oliveira

3.1 - Variáveis aleatórias. Função de distribuição: tipos de variáveis aleatórias.

Ao descrever-se o espaço de resultados de uma experiência aleatória os acontecimentos elementares não têm que ser numéricos, o que dificulta a aplicação de procedimentos matemáticos.
Em muitas situações não interessam todos os aspectos resultantes de uma dada experiência aleatória mas apenas um valor numérico.

Surge assim a necessidade de definição de **funções reais**, associadas à experiência aleatória, que permitam:

- associar a um valor real cada acontecimento elementar;
- preservar a sua distribuição de probabilidades.

Informalmente pode-se dizer que variável aleatória é uma função que associa um número real a cada um dos elementos do espaço de resultados de uma experiência aleatória.

Definição: Considere uma experiência aleatória e Ω o seu espaço de resultados. À função $X(\omega)$ tal que $X : \Omega \longrightarrow \mathbb{R}$ diz-se uma **v**ariável **a**leatória (v.a.) se para qualquer $x \in \mathbb{R}$, o conjunto de Ω, $\{\omega \in \Omega : X(\omega) \leq x\}$ é um acontecimento.

$$\begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ \omega & & X(\omega) \\ & X & \end{array}$$

Observações:

1. representam-se, geralmente, pelas últimas letras maiúsculas do alfabeto latino, $X, Y, Z, X_1, Y_1, Z_1, \cdots$;
2. um valor particular representa-se por uma letra minúscula;
3. $P(X = x) = P(X(\omega) = x) = P\{\omega \in \Omega : X(\omega) = x\}$
4. num mesmo espaço de resultados podem definir-se diversas variáveis aleatórias.

Definição: Define-se **função de distribuição (cumulativa)** de uma v.a. X à função real de valor real, $F_X : \mathbb{R} \longrightarrow [0,1]$, que a cada $x \in \mathbb{R}$ faz corresponder a probabilidade do acontecimento $\{\omega : X(\omega) \leq x\}$, representa-se

$$F_X(x) = P(X \leq x).$$

Algumas propriedades da função de distribuição:

1. $0 \leq F_X(x) \leq 1, \forall x \in \mathbb{R}$;
2. $F_X(x)$ é não decrescente;
3. $\lim_{x \to -\infty} F_X(x) = 0$ e $\lim_{x \to +\infty} F_X(x) = 1$;
4. $F_X(x)$ é contínua à direita;
5. $P(X = x_0) = F_X(x_0) - F_X(x_0^-) = F_X(x_0) - \lim_{x \to x_0} F_X(x)$;
6. $P(x_0 < X \leq x_1) = F_X(x_1) - F_X(x_0)$, com $x_0 < x_1$.

3.1 - Tipos de variáveis aleatórias.

D_X o conjunto (numerável) de pontos de descontinuidade de $F_X(x)$:

- A v.a. X diz-se **d**iscreta quando $P(X \in D_X) = 1$ $\rightarrow$ CAP. 3
- A v.a. X diz-se **c**ontínua quando $D_X = \emptyset$. $\rightarrow$ CAP. 4
- A v.a. X diz-se **m**ista quando $D_X \neq \emptyset$ e $P(X \in D_X) < 1$.

Observação:

- Em alternativa também se pode definir v.a. discreta se o conjunto dos seus possíveis valores for um conjunto finito ou infinito numerável.

Exemplo 3.1: Considere a experiência aleatória (E.A.):
$E.A. =$"Medição do tempo de funcionamento (tempo de vida), em horas, de uma componente", com $\Omega = [0, +\infty[$.
Definam-se as seguintes variáveis aleatórias:

- T - "v.a. que indica o tempo de vida (duração) da componente"

  Valores possíveis desta v.a. $(R_T) \longrightarrow$ coincidem com o espaço de resultados; $R_T = \mathbb{R}_0^+ \implies T$ **não é discreta**

- Z - "v.a. que indica se a duração é maior ou não que 100 horas."

$$Z = \begin{cases} 0, & T \leq 100 \\ 1, & T > 100 \end{cases}$$

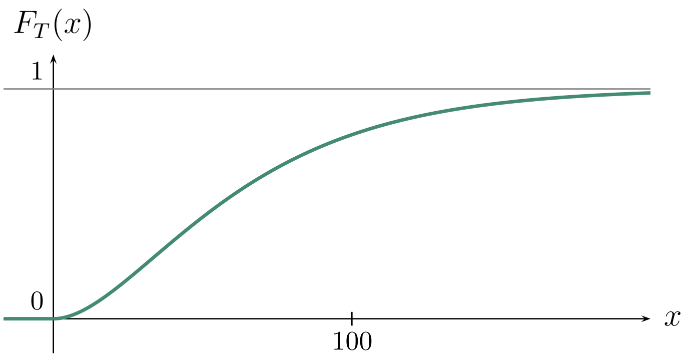
  $R_Z = \{0, 1\} \implies Z$ **é discreta**

- W - v.a. definida como:

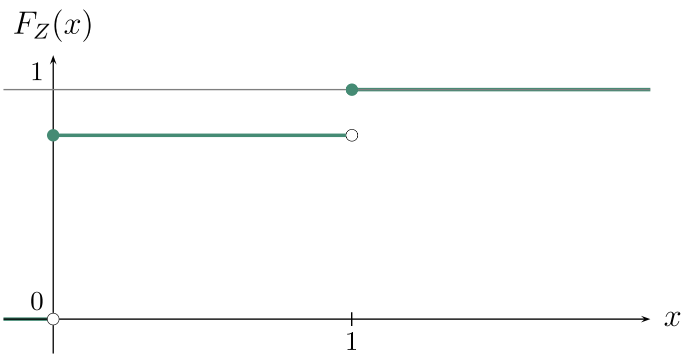
$$W = \begin{cases} 0, & T \leq 100 \\ T - 100, & T > 100 \end{cases}$$

  $R_W = \{0\} \cup \mathbb{R}^+ \implies W$ **é mista**

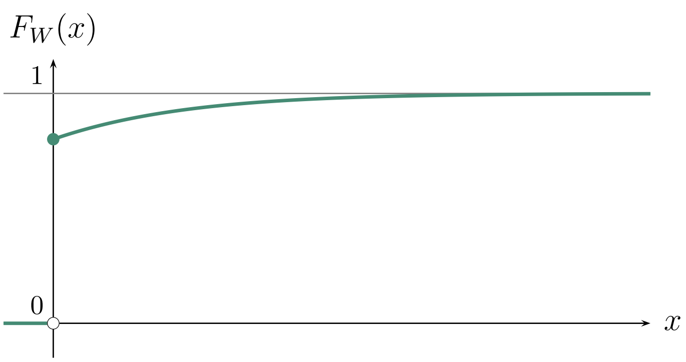
Esboço dos gráficos das funções de distribuição, F_T, F_Z e F_W:



T v.a. contínua



Z v.a. discreta



W v.a. mista

3.2 - Variáveis aleatórias discretas. Função (massa) de probabilidade

Exemplo 3.2: Lançamento de duas moedas equilibradas e observar o número de coroas.

$\Omega = \{(F,F),(C,F),(F,C),(C,C)\}$, onde $F =$"saída de Face" e $C =$"saída de Coroa";

Seja

Y - "v.a. número de coroas obtidas no lançamento."

Y é discreta e pode tomar os valores $y = 0,1,2$: $R_Y = \{0,1,2\}$

$$P(Y=0) = P(\{(F,F)\}) = \frac{1}{4}$$

$$P(Y=1) = P(\{(F,C),(C,F)\}) = \frac{1}{2}$$

$$P(Y=2) = P(\{(C,C)\}) = \frac{1}{4}$$

Exemplo 3.3: E.A. "Lançamentos sucessivos de um dado equilibrado até sair face 6."

$$\Omega = \{(6), (1,6), (2,6), \ldots, (5,6), (1,1,6), \ldots, (5,5,6), \ldots\}$$

X - "v.a. que indica o número de lançamentos até sair a face 6"

X é discreta, os valores possíveis são $y = 1, 2, 3, \ldots, R_X = \mathbb{N}$

$$P(X = 1) = \frac{1}{6}$$

$$P(X = 2) = \frac{5}{6} \times \frac{1}{6}$$

$$P(X = 3) = \left(\frac{5}{6}\right)^2 \times \frac{1}{6}, \ldots$$

$$P(X = x) = \left(\frac{5}{6}\right)^{x-1} \times \frac{1}{6}, \;\; x \in \mathbb{N}$$

3.2 - Função de probabilidade

Definição: Seja X v.a. discreta assumindo valores em $R_X = \{x_1, x_2, \ldots\}$. A **função (massa) de probabilidade** de X é dada por:

$$P(X = x) = \begin{cases} P(X = x_i), & x = x_i \in R_X \\ 0, & c.c. \end{cases}$$

e satisfaz:

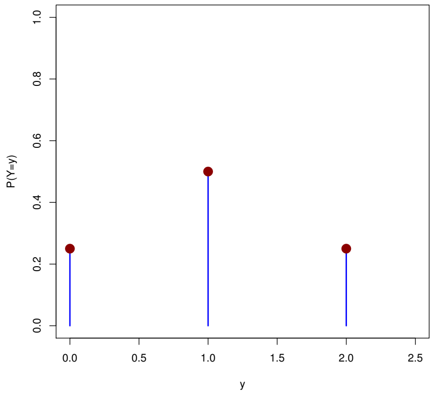
- $P(X = x) \geq 0, \forall x \in \mathbb{R}$;
- $\sum_{x \in \mathbb{R}} P(X = x) = \sum_{x \in R_X} P(X = x) = 1$.

Nota:
Esta função pode ser denotada por: $P(X = x)$, $P_X(x)$, $f_X(x)$ ou $f(x)$.

3.2 - Função de probabilidade

Exemplo 3.2 (cont.): A função de probabilidade de Y - v.a. "número de coroas obtidas no lançamento", é:

$$P(Y=y)=\begin{cases} 1/4, & y=0,2 \\ 1/2, & y=1 \\ 0, & y\notin\{0,1,2\} \end{cases}$$



Exemplo 3.3 (cont.): A função de probabilidade de
X - “v.a. número de lançamentos até sair a face 6”, é

$$f(x) = P(X = x) = \begin{cases} \left(\frac{5}{6}\right)^{x-1} \times \frac{1}{6}, & x \in \mathbb{N} \\ 0, & \text{c.c.} \end{cases}$$

É fácil verificar que

$$\sum_{x=1}^{\infty} \left(\frac{5}{6}\right)^{x-1} \times \frac{1}{6} = \frac{1/6}{1 - 5/6} = 1$$

(soma da série geométrica de razão $5/6$ e 1.$^{\text{o}}$termo $1/6$)

Observações:

Quando X é v.a. discreta:

1. $P(X \in A) = \sum_{x_i \in A} P(X = x_i), \forall A \subset \mathbb{R}$;
2. A **função de distribuição** (ou função de distribuição acumulada) de uma v.a. discreta, X, é

$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i), x \in \mathbb{R}$$

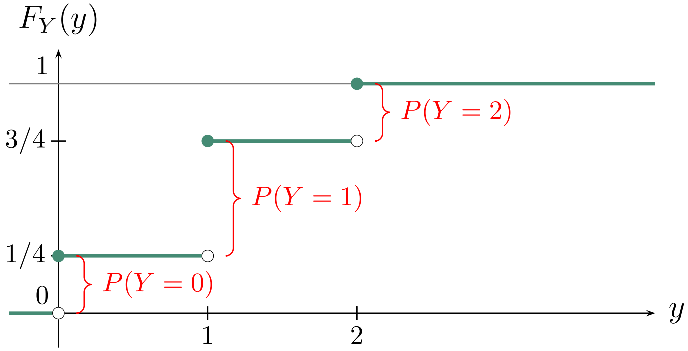
   é então uma função "em escada". Os "saltos" são o valor da função de probabilidade no ponto correspondente;
3. A função de distribuição é contínua à direita.

4. O cálculo de probabilidades à custa de $F_X(x)$ é simples:

$$P(X = x) = F_X(x) - F_X(x^-);$$
$$P(X > x) = 1 - F_X(x);$$
$$P(x_0 < X \leq x_1) = F_X(x_1) - F_X(x_0), x_0 < x_1.$$
$$P(x_0 \leq X < x_1) = F_X(x_1^-) - F_X(x_0^-), x_0 < x_1.$$

Exemplo 3.2 (cont.):

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ 1/4 & 0 \leq y < 1 \\ 3/4 & 1 \leq y < 2 \\ 1 & y \geq 2 \end{cases}$$



3.3 Valor esperado, variância e algumas das suas propriedades. Moda e quantis.

Quer a função de distribuição quer a função de probabilidade caracterizam completamente uma v.a. discreta. No entanto, podemos estar interessados em outras medidas que caracterizem de forma parcial a v.a.. Entre essas medidas podem destacar-se:

- Medidas (parâmetros) de localização:
    - Valor esperado;
    - Moda;
    - Mediana;
    - Quantis.
- Medidas (parâmetros) de dispersão:
    - Variância;
    - Desvio padrão.

3.3 Valor esperado

Definição: O **valor esperado**, **valor médio**, ou **média** de uma v.a. discreta, X, com função de probabilidade $f_X(x) = P(X = x)$, representa-se por $E(X)$, μ_X ou μ e é dado por

$$E(X) = \sum_x x\, P(X = x).$$

Observações:

1. $E(X)$ é um valor numérico nas mesmas unidades que a variável X, mas não é necessariamente um dos valores que a variável pode tomar.
2. Quando X pode tomar um n.º infinito de valores (Exemplo 3.3 da secção anterior), $E(X)$ é dado pela soma de uma série, que pode não ser convergente, nesse caso diz-se que não existe valor esperado.
3. O valor esperado corresponde ao conceito físico de centro de gravidade (de um sistema unidimensional discreto em que os pontos têm coordenadas x_i e massas $P(X = x_i)$).

3.3 Valor esperado: propriedades

1) Valor esperado de uma função de X

$$\boxed{E\left[h(X)\right] = \sum_x h(x)\, P(X = x)}$$

Em geral tem-se $E\left[h(X)\right] \neq h\left[E(X)\right]$, excepto no casos em que h é uma função linear ou a v.a. é degenerada (ou seja, é uma constante).

2) Se X é v.a. inteira não negativa, então:

$$\boxed{E(X) = \sum_{x=0}^{+\infty} P(X > x) = \sum_{x=0}^{+\infty} \left[1 - F_X(x)\right]}$$

3.3 Valor esperado: propriedades

3) O operador **valor esperado é linear**

$$\boxed{E\,(a\,X+b)=a\,E(X)+b, \qquad \forall_{a,b\in\mathbb{R}}}$$

Demonstração:

$$\begin{aligned}
E\,(a\,X+b) &= \sum_x (a\,x+b)\,P(X=x) = \\
&= \left(\sum_x a\,x\,P(X=x)\right)+\left(\sum_x b\,P(X=x)\right) = \\
&= a\,\underbrace{\sum_x x\,P(X=x)}_{E(X)}+b\,\underbrace{\sum_x P(X=x)}_{1} = a\,E(X)+b
\end{aligned}$$

Caso particular: $E\,(b)=b, \qquad \forall_{b\in\mathbb{R}}$

Exemplo: Considere-se novamente a v.a. Y com $P(Y=0)=\frac{1}{4}$, $P(Y=1)=\frac{1}{2}$, $P(Y=2)=\frac{1}{4}$ que representa o número de coroas no lançamento de duas moedas equilibradas.

Imagine o seguinte jogo: paga-se um preço para lançar as duas moedas e recebe-se um prémio de Y^2 Euros. Qual o preço justo a pagar?

Como o valor equilibrado a longo prazo é $E(Y^2)$ e a função de probabilidade de Y^2 é dada por:

i	0	1	4
$P(Y^2=i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

logo $E(Y^2)=0\times\dfrac{1}{4}+1\times\dfrac{1}{2}+4\times\dfrac{1}{4}=1.5$ (Euros).

Notar que $E(Y^2)\neq[E(Y)]^2=\left(0\times\frac{1}{4}+1\times\frac{1}{2}+2\times\frac{1}{4}\right)^2=1$.

3.3 Moda

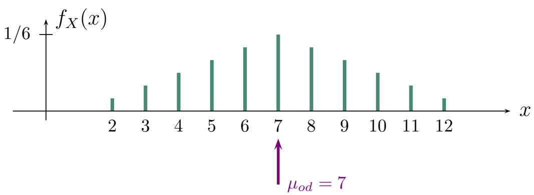
Definição: **Moda** de uma v.a. discreta, X com função de probabilidade $f_X(x)$ representa-se por μ_o ou $Mo(X)$ e é o valor, ou valores, com máxima probabilidade de ocorrência

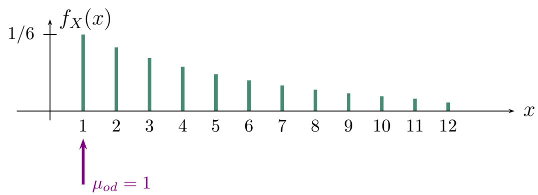
$$\mu_o = \text{moda de } X : \quad f_X(\mu_o) = \max_x f_X(x)$$

ou equivalente:

$$\mu_o = \arg\max_x f_X(x)$$

Observação: A moda pode não ser única. Podem definir-se modas relativas (correspondentes a máximos relativos).





3.3 Quantis

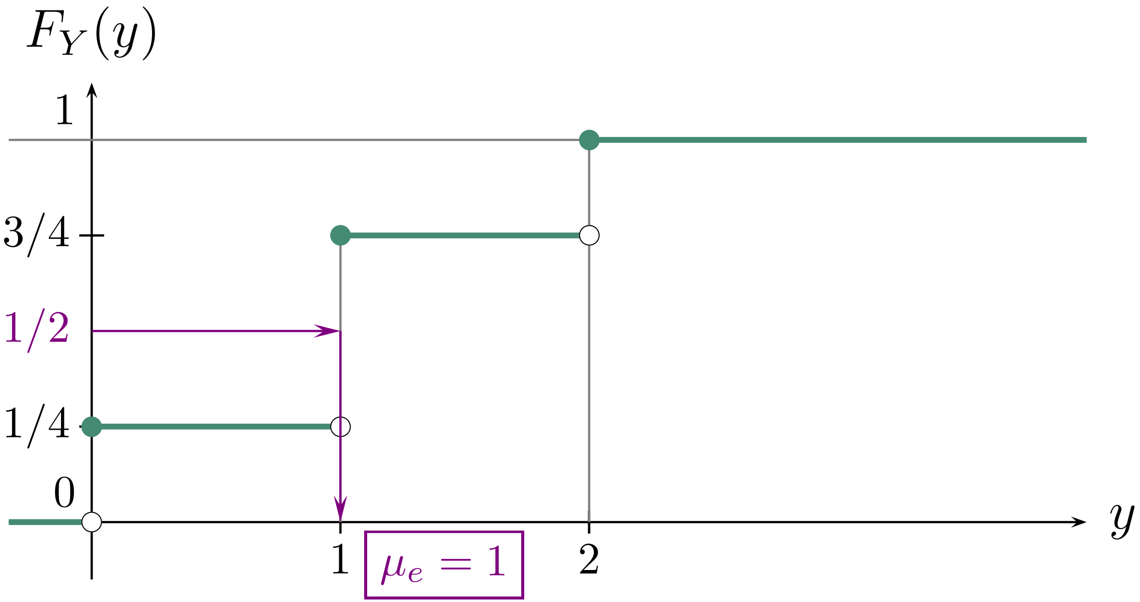
Definição: **Mediana** de uma v.a. discreta X com função de distribuição $F_X(x)$, representa-se por μ_e ou me_X e é um ponto central em termos de probabilidade, ou seja, tal que, se possível, $P(X \leq \mu_e) = P(X \geq \mu_e)$

$$\mu_e : \quad P(X \leq \mu_e) \geq \frac{1}{2} \quad \text{e} \quad P(X \geq \mu_e) \geq \frac{1}{2}$$

que é equivalente:

$$\mu_e : \quad \frac{1}{2} \leq F_X(\mu_e) \leq \frac{1}{2} + P(X = \mu_e)$$

Graficamente:



3.3 Quantis

Generalização da ideia de mediana:

Definição: **Quantil de ordem** p $(0 < p < 1)$ de uma v.a. discreta, X com função de distribuição $F_X(x)$, denota-se por $\chi_p = F_X^{-1}(p)$ e é dado por

$$\chi_p : \quad P(X \leq \chi_p) \geq p \quad \text{e} \quad P(X \geq \chi_p) \geq 1 - p$$

que é equivalente:

$$\chi_p : \quad p \leq F_X(\chi_p) \leq p + P(X = \chi_p)$$

Observações:

- A mediana da v.a. X corresponde ao quantil de ordem $p = \frac{1}{2}$
- Aos quantis cujo p é múltiplo de $\frac{1}{4}$ chamam-se quartis.

3.3 Variância

Definição: A **variância**, de uma v.a. discreta, X com função de probabilidade $f_X(x)$ e valor esperado $\mu_X \equiv E(X)$, representa-se por $V(X)$, $\mathrm{VAR}(X)$, σ_X^2 ou σ^2 e é dada por

$$V(X) = E\left[(X - \mu_X)^2\right] = \sum_x (x - \mu_X)^2 \, f_X(x)$$

Observações:

1. A variância é um valor numérico expresso no quadrado das unidades da v.a.X.
2. Assim como pode não existir valor esperado, também pode não existir variância.
3. Do mesmo modo que o valor esperado é análogo do conceito físico de centro de gravidade, $V(X)$ é análoga do momento de inércia em relação a um eixo que passa pelo centro de gravidade.

Definição: O desvio padrão, de uma v.a. X, σ_X, é a raiz quadrada positiva de $V(X)$.

Observação: O desvio padrão é expresso nas unidades da variável.

Propriedades da variância

1) $V(X) \geq 0, \ \forall_X$ óbvio: $V(X) = \sum_x \underbrace{(x - \mu_X)^2}_{\geq 0} \underbrace{f_X(x)}_{\geq 0}$

2) $V(X) = E(X^2) - [E(X)]^2 = E(X^2) - E^2(X) = E(X^2) - \mu^2$

Dem.:

$$\begin{aligned} V(X) &= E\left[(X-\mu)^2\right] = E\left(X^2 - 2\,\mu\,X + \mu^2\right) = \\ &= E(X^2) - 2\,\mu\,\underbrace{E(X)}_{\mu} + \mu^2 = E(X^2) - 2\,\mu^2 + \mu^2 \end{aligned}$$

Nota: 1) e 2) $\Rightarrow$ $E(X^2) \geq \mu^2$

Propriedades da variância (cont.)

3) $V(aX+b)=a^2\,V(X), \qquad \forall_{a,b\in\mathbb{R}}$ (a variância não é um operador linear!)

Demonstração:

$$\begin{aligned} V(aX+b) &= E\left[(aX+b-E(aX+b))^2\right] = \\ &= E\left[(aX+b-a\,E(X)-b)^2\right] = \\ &= E\left[(aX-a\,E(X))^2\right] = E\left[a^2\,(X-E(X))^2\right] = \\ &= a^2 E\left[(X-E(X))^2\right] = a^2\,V(X) \end{aligned}$$

Casos particulares:

- $V(b) = 0 \quad (a = 0)$
- $V(X + b) = V(X) \quad (a = 1)$
- $V(aX) = a^2\, V(X) \quad (b = 0)$

4) $V(X) = 0 \quad \Leftrightarrow \quad \exists_c :\ P(X = c) = 1$

(ou seja, X é constante — também se diz “v.a. degenerada”)

Exemplo Considerem-se duas v.a. X e Y cujas funções de probabilidade são, respectivamente:

$$X: \quad \begin{array}{c|ccc} x & -1 & 1 & c.c. \\ \hline f_X(x) & 1/2 & 1/2 & 0 \end{array} \qquad Y: \quad \begin{array}{c|ccc} y & -1000 & 1000 & c.c. \\ \hline f_Y(x) & 1/2 & 1/2 & 0 \end{array}$$

onde $E(X) = E(Y) = 0$. No entanto:

$E(X^2) = (-1)^2 \times \frac{1}{2} + 1^2 \times \frac{1}{2} = 1$ e

$E(Y^2) = (-1000)^2 \times \frac{1}{2} + 1000^2 \times \frac{1}{2} = 1000^2$

$V(X) = 1$ e $V(Y) = 1000^2$

$\sigma_X = 1$ e $\sigma_Y = 1000$

Definição: A v.a. X discreta diz-se ter distribuição uniforme discreta no conjunto $\{x_1,\ x_2,\ \ldots,\ x_N\}$, de cardinal N, se tiver probabilidade $f_X(x;N) = P(X = x) = \frac{1}{N}$, para $x = x_1, x_2, \ldots, x_N$ e $f_X(x) = 0$ nos restantes valores.

Notação: $X \sim Unif\{x_1, \ldots, x_N\}$

Para esta distribuição tem-se:

$$E(X) = \frac{1}{N}\sum_{i=1}^{N} x_i$$

e

$$V(X) = E(X^2) - (E(X))^2 = \frac{1}{N}\sum_{i=1}^{N} x_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)^2$$

Caso Particular:

Se $X \sim \{1, 2, \ldots, N\}$,

tem-se:

$$\boxed{\mu_X = E(X) = \frac{N+1}{2} =}$$

$$= \sum_{i=1}^{N} i \frac{1}{N} = \frac{1}{N} \sum_{i=1}^{N} i = \frac{1}{N} \frac{N(N+1)}{2}$$

$$\boxed{\sigma_X^2 = V(X) = \frac{N^2-1}{12}}$$

Nota: $\sum_{i=1}^{N} i^2 = \frac{N(N+1)(2N+1)}{6}$

✓**Exemplo:** X-"v.a. que indica a face do dado (equilibrado) que ocorre em um lançamento"

$X \sim Unif\{1, \ldots, 6\}, E(X) = 3.5, V(X) = 35/12$

Distribuição Bernoulli

Definição: Uma experiência aleatória chama-se prova ou experiência de Bernoulli quando só tem dois resultados possíveis: A ou $\bar{A}$.

- a ocorrência do acontecimento A tem probabilidade p: $P(A) = p$ $(0 < p < 1)$;
- a ocorrência do acontecimento $\bar{A}$ tem probabilidade $1 - p$: $P(\bar{A}) = 1 - p$ $(0 < p < 1)$.

Nesta experiência designa-se por:

- sucesso à ocorrência do acontecimento A.
- insucesso à ocorrência do acontecimento $\bar{A}$.

Distribuição de Bernoulli

A partir de experiências de Bernoulli define-se uma variável aleatória X, com distribuição de Bernoulli, do seguinte modo: $X = 1$ se ocorre A e $X = 0$ se ocorre $\bar{A}$, com $R_X = \{0, 1\}$

Definição: A v.a. X discreta tem distribuição de Bernoulli de parâmetro p, $0 < p < 1$ (escreve-se, $X \sim Ber(p)$) quando

$$P(X = x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \\ 0, & c.c. \end{cases} \equiv \begin{cases} p^x(1-p)^{1-x}, & x = 0, 1 \\ 0, & c.c. \end{cases},$$

Para esta distribuição tem-se:

- $E(X) = E(X^2) = p$;
- $V(X) = E(X^2) - E^2(X) = p - p^2 = p(1-p)$;
- $F_X(x)$ tabelado para alguns valores de p.

Distribuição de Bernoulli

Exemplo da Urna: Considere-se uma urna com N bolas, das quais:

M são Brancas - característica associada ao "sucesso" (A)

$(N - M)$ são Pretas - característica associada ao "insucesso" $(\bar{A})$

E.A.- "Extrair 1 bola ao acaso"
X - "v.a.que indica a ocorrência do sucesso (bola Branca)",
$X \sim Ber(p = \frac{M}{N})$

Admitindo que $N = 100$, $M = 20$, $(N - M) = 80$ tem-se $X \sim Ber(0.2)$.

- A distribuição de Bernoulli é muito importante para a definição das distribuições Binomial, Geométrica e Hipergeométrica.

3.6 Distribuição Binomial

O modelo probabilístico adequado para descrever uma experiência aleatória que consiste na repetição de n provas de Bernoulli independentes é a distribuição Binomial.

Definição: A v.a. X discreta que indica o n$^{\text{o}}$de sucessos observados em n realizações independentes de provas de Bernoulli e com probabilidade de sucesso p em cada prova, diz-se ter distribuição de Binomial de parâmetros, $n \in \mathbb{N}_0$ e $p \in]0, 1[$ (escreve-se, $X \sim Bin(n, p)$) e tem função de probabilidade igual a:

$$P\left(X = x\right) = \begin{cases} \binom{n}{x} p^x \left(1 - p\right)^{n-x}, & x = 0, 1, 2, \cdots n \\ 0, & c.c. \end{cases}$$

Para esta distribuição tem-se:

- $E(X) = np$
- $V(X) = np(1 - p)$
- $F_X(x)$ tabelado para alguns valores de n e p.

3.6 Distribuição Binomial

Observações:

- Recorrendo ao binómio de Newton:

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x\, b^{n-x}$$

com $a = p$ e $b = 1 - p$ é imediato mostrar que $\sum_{x=0}^{n} P(X = x) = 1$, pois

$$\sum_{x=0}^{n} P(X = x) = \sum_{x=0}^{n} \binom{n}{x} p^x\,(1-p)^{n-x} = 1 \qquad (\forall_{0<p<1}) \qquad (1)$$

- Se $X \sim Bin(n, p) \Rightarrow Y = n - X \sim Bin(n, 1-p)$
- Se $X \sim Bin(1, p) \Leftrightarrow X \sim Ber(p)$

Observações(cont.)– Cálculo de probabilidades para v.a. com distribuição binomial

- Se n pequeno: usar a fórmula de $P(X = k)$
- Se n grande:
  - Programas em computador ou calculadora
  - Tabelas, mas notar que:
    - Número limitado de situações cobertas: $n = 1, \ldots, 20$ e $p = 0.01, 0.02, \ldots, 0.09, 0.10, 0.15, 0.20, \ldots, 0.50$
    - O que está tabelado é a função de distribuição. Recordar que a função de probabilidade se pode obter neste caso fazendo $P(X = k) = P(X \leq k) - P(X \leq k - 1) = F_X(k) - F_X(k - 1)$
  - Aproximações (a ver posteriormente).

3.6 Distribuição Binomial: Exemplo 1

Exemplo da Urna: N bolas das quais M são Brancas e $(N-M)$ são Pretas

E.A.-"Extrair n bolas ao acaso e com reposição"

$\Longrightarrow$ n provas de Bernoulli independentes

X - "v.a. que indica o n$^{\text{o}}$de sucessos (n$^{\text{o}}$de bolas Brancas) nas n bolas"

Como $P(A_i) = \frac{M}{N} = p, i = 1, \ldots, n$, e $P(A_i|A_{i-1}) = \frac{M}{N} = p$ tem-se que:

$$X \sim Bin(n, p = \tfrac{M}{N})$$

Para $N = 100$, $M = 20$, $(N-M) = 80$ e $n = 3$ temos que $X \sim Bin(3, 0.2)$.

Qual é a probabilidade de o n$^{\text{o}}$de bolas brancas ser no máximo 1?

$P(X \leq 1) = F_X(1) = 0.896$ (consultando a tabela)

3.6 Distribuição Binomial: Exemplo 2

Exemplo 2: Numa população muito grande, sabe-se que 50% das pessoas são a favor de jogos electrónicos infantis.

Se forem escolhidas 10 pessoas, ao acaso, qual é a probabilidade de encontrar 6 que são a favor de jogos electrónicos infantis?

E qual é a probabilidade de encontrar pelo menos 6 nas mesmas condições?

Resolução: Seja X - v.a. que indica o n.$^{\text{o}}$de pessoas a favor de jogos electrónicos infantis nas 10 escolhidas

X representa o n.$^{\text{o}}$de sucessos em 10 provas de Bernoulli que se podem considerar independentes (dado que a população é muito grande) com $p = 0.5$, logo $\quad X \sim Bin(10, 0.5)$.

$P(X = 6) =?$

$P(X \geq 6) =?$

3.6 Distribuição Binomial: Exemplo 2

$$P(X = 6) = \binom{10}{6} \times \left(\frac{1}{2}\right)^{10} = \frac{210}{1024} = \frac{105}{512}$$

Em R: `> dbinom(6,10,0.5)` $\longrightarrow$ `[1] 0.2050781`

Em Excel: `=BINOMDIST(6,10,0.5,FALSE)` $\longrightarrow$ `0.205078125`

Tabela (p. 1, canto inferior direito):
$F_X(6) - F_X(5) = 0.8281 - 0.6230 = 0.2051$

$$P(X \geq 6) = \sum_{x=6}^{10} \binom{10}{x} \left(\tfrac{1}{2}\right)^{10} = \ldots = 1 - P(X < 6) = 1 - P(X \leq 5) = 1 - F_X(5)$$

Em R: `> 1-pbinom(5,10,0.5)` $\longrightarrow$ `[1] 0.3769531`

Em Excel: `=1 - BINOMDIST(5,10,0.5,TRUE)` $\longrightarrow$ `0.376953125`

Tabela (p. 1, canto inferior direito): $1 - 0.6230 = 0.3770$

23 Mar 2020

3.7 Distribuição geométrica

Nas distribuições Binomial e Hipergeométrica o número de provas de Bernoulli (n) é fixo. Considere-se agora que temos uma sucessão infinita de provas de Bernoulli independentes até à realização do 1º sucesso.

Definição: Seja X a v.a. que indica o nº de provas de Bernoulli independentes e com probabilidade de sucesso constante (p), realizadas até à obtenção do primeiro sucesso (inclusive). Diz-se que X tem distribuição geométrica de parâmetro p, $X \sim Geo(p)$ e a sua função de probabilidade é:

$$P(X = x) = \begin{cases} (1-p)^{x-1}p, & x = 1, 2, 3, \ldots \\ 0, & c.c. \end{cases}$$

Valor esperado e variância:

$$E(X) = 1/p;\ V(X) = (1-p)/p^2 \quad , \quad F_X(x) = \begin{cases} 1 - (1-p)^{\lfloor x \rfloor}, & x \geq 1 \\ 0, & cc \end{cases}$$

Exemplo da Urna: Qual é a probabilidade de ser necessário extraír (com reposição) no mínimo 3 bolas até encontrar a 1ª branca?

Dem Função distribuição:

$$x \geq 1, \quad F_X(x) = P(X \leq x) = \sum_{k=1}^{\lfloor x \rfloor} (1-p)^{k-1} p \;=\; p \sum_{n=0}^{\lfloor x \rfloor - 1} (1-p)^{n}$$

$n = k-1$ e
$1 \leq k \leq +\infty$
então $0 \leq k-1 \leq +\infty$

$$F_X(x) = p \, \frac{1-(1-p)^{\lfloor x \rfloor}}{1-(1-p)} = 1-(1-p)^{\lfloor x \rfloor} \qquad \text{QED}$$

Propriedade Falta de memória da Geom(p):

$\boxed{\text{Teo:}}$ Se $X \sim Geom(p)$ então $P(X > k+x \mid X > k) = P(X > x)$, $\forall x, k \in \mathbb{N}$

Dem:
$$P(X > k+x \mid X > k) = \frac{P(X > k+x, \, X > k)}{P(X > k)} = \frac{P(X > k+x)}{P(X > k)}$$

$$= \frac{(1-p)^{k+x}}{(1-p)^{k}} = (1-p)^{x} = 1 - \left[1-(1-p)^{x}\right] = 1 - F_{Geom(p)}(x)$$

$$= P(X > x) \qquad \text{QED}$$

observação:
- Se X é va discreta e verifica a prop. falta de memória então $X \sim Geom(p)$
- A prop. de falta de memória da geométrica implica que:

$$\boxed{X \sim Geom(p) \quad \Rightarrow \quad X-k \mid X>k \sim Geom(p), \; \forall k \in \mathbb{N}}$$

3.5 Distribuição Hipergeométrica

O modelo probabilístico adequado para descrever uma experiência aleatória que consiste na repetição de n provas de Bernoulli dependentes (extrações sem reposição) é a distribuição Hipergeométrica.

Definição: (**Exemplo da Urna**) N bolas das quais M são Brancas e $(N-M)$ são Pretas $(M \leq N)$

E.A.-"Extrair n bolas (com $n \leq N$) ao acaso e sem reposição" $\Longrightarrow$ n provas de Bernoulli dependentes

X - "v.a. que indica o n^{o}de sucessos (n^{o}de bolas Brancas) nas n bolas"

Como $P(A_i) = \frac{M}{N} = p, i = 1, \ldots, n,$ (probabilidade do sucesso constante) mas $P(A_i|A_{i-1}) \neq \frac{M}{N} = p$ (provas de Bernoulli dependentes) X tem distribuição hipergeométrica com parâmetros N, M e n, $X \sim Hip(N, M, n)$, com função de probabilidade:

$$P(X = x) = \begin{cases} \dfrac{\dbinom{M}{x}\dbinom{N-M}{n-x}}{\dbinom{N}{n}}, & x = \max\{0, n-(N-M)\}, \ldots, \min\{M, n\} \\ 0, & c.c. \end{cases}$$

Valores possíveis:

Se $n \leq M$ e $n \leq N - M \qquad x = 0, 1, 2, \ldots, n$

Se $n > M$ e $n \leq N - M \qquad x = 0, 1, 2, \ldots, M$

Se $n \leq M$ e $n > N - M \qquad x = n - (N - M), \ldots, n$

Se $n > M$ e $n > N - M \qquad x = n - (N - M), \ldots, M$

ou, englobando todos os casos,

$$x = \max\{0, n - (N - M)\}, \ldots, \min\{M, n\}$$

Valor esperado e variância:

$$E(X) = n\frac{M}{N}$$

$$V(X) = n\frac{M}{N}(1 - \frac{M}{N})\frac{N-n}{N-1}$$

Observações:

- Como $p = \dfrac{M}{N}$ tem-se $E(X) = np$ e $V(X) = np(1-p)\frac{N-n}{N-1}$ assim, $Bin(n, p)$ e $Hip(M, N, n)$ têm o mesmo valor médio e as variâncias apenas se distinguem pelo factor $\frac{N-n}{N-1}$.
- Quando N é grande e n pequeno comparado com N, esbate-se a diferença entre extracções com e sem reposição, i.e. $\frac{N-n}{N-1} \approx 1$.
- Se N grande e n pequeno em relação a N (regra prática: $n < 0.1N$), pode usar-se a distribuição binomial para calcular valores aproximados das probabilidades, i.e.

$$X \sim Hip(N, M, n) \approx \tilde{X} \sim Bin\left(n, p = \frac{M}{N}\right)$$

3.5 Distribuição Hipergeométrica: Exemplo 1

Exemplo da Urna: $N = 100$ bolas das quais $M = 20$ são Brancas e $(N - M) = 80$ são Pretas

E.A.-"Extrair 3 bolas ao acaso e sem reposição"

$\Longrightarrow$ 3 provas de Bernoulli dependentes

Y - "v.a. que indica o n^{o}de sucessos (n^{o}de bolas Brancas) nas 3 bolas"

$Y \sim Hip(N = 100, M = 20, n = 3)$ e a sua função de probabilidade é:

$$P(Y = y) = \begin{cases} \dfrac{\binom{20}{x}\binom{80}{3-x}}{\binom{100}{3}}, & y = 0, 1, 2, 3 \\ 0 & , cc \end{cases}$$

Qual é a probabilidade de o n^{o}de bolas brancas ser no máximo 1?

$P(Y \leq 1) = F_Y(1) = \frac{\binom{20}{0}\binom{80}{3}}{\binom{100}{3}} + \frac{\binom{20}{1}\binom{80}{2}}{\binom{100}{3}}$ (não temos tabela da função de distribuição)

3.5 Distribuição Hipergeométrica: Exemplo 2

Exemplo:
Caixa com 100 peças das quais 10 são defeituosas. São efectuadas 5 extracções ao acaso, sem reposição.

Seja X - número de peças defeituosas encontradas nas 5 extracções;

Valores possíveis: $x = 0, 1, 2, 3, 4, 5$

$X \sim Hip(N = 100, M = 10, n = 5)$ e a sua função de probabilidade é:

$$P(X = x) = \begin{cases} \dfrac{\binom{10}{x}\binom{90}{5-x}}{\binom{100}{5}}, & x = 0, 1, 2, 3, 4, 5 \\ 0, & cc \end{cases}$$

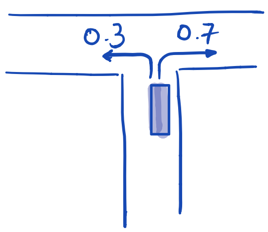
Nota:
Se as extracções forem efectuadas com reposição (o que em termos práticos é estranho!...) têm-se 5 provas de Bernoulli independentes com $P(\text{sucesso}) = 0.1$, ou seja $\quad X \sim Bin(5, 0.1)$.

Comparação das probabilidades:

	Sem reposição	**Com** reposição
$P(X = 0) \simeq$	0.584	0.590
$P(X = 1) \simeq$	0.339	0.328
$P(X = 2) \simeq$	0.070	0.073
$P(X = 3) \simeq$	0.006	0.008
$P(X = 4) \simeq$	2.5×10^{-4}	4.5×10^{-4}
$P(X = 5) \simeq$	3.3×10^{-6}	1.0×10^{-5}

A situação **com reposição**, embora irrealista, dá probabilidades próximas das calculadas para o caso **sem reposição**, esta aproximação é razoável já que $n < 0.1N$. i.e. $n = 5$, $0.1\ N = 0.1(100) = 10$ e $5 < 10$.

Example 3. Ao chegarem a um cruzamento os condutores de viaturas decidem, de forma independente, virar à esquerda ou à direita, sendo a prob. de virar à esquerda 0.3 para qualquer condutor.



(a) Determine a prob. de pelo menos 10 dos próximos 15 condutores virarem à esq.

(b) Calcule a prob. de chegar no total mais de 3 viaturas ao entronca/ até se que se observe a primeira viragem à esquerda.

$\boxed{\text{Solução:}}$

Processo de Poisson

A distribuição de Poisson pode ser vista como um limite "especial" da distribuição binomial ou então como o número de ocorrências aleatórias de um acontecimento que se repete no tempo ou no espaço (por exemplo, o n$^{\text{o}}$de clientes que chegam a um banco ou o n$^{\text{o}}$de defeitos numa peça de tecido,...).

Definição: Considere-se a contagem do n$^{\text{o}}$de ocorrências aleatórias de um acontecimento durante um intervalo de tempo (comprimento, área, etc). Se esse intervalo puder ser dividido em sub-intervalos de comprimento suficientemente pequeno de modo a que se verifique que:

1. A probabilidade de ocorrer mais do que um acontecimento num sub-intervalo é zero.
2. A probabilidade de uma ocorrência num sub-intervalo é constante e proporcional ao comprimento desse sub-intervalo.
3. O n$^{\text{o}}$de ocorrências desse acontecimento registadas nos diversos sub-intervalos são independentes entre si.

Então esta experiência aleatória chama-se Processo de Poisson.

3.8 Distribuição de Poisson

Definição: Seja X uma v.a. que indica o n$^{\text{o}}$de ocorrências de um acontecimento por unidade de tempo ou de espaço (comprimento, área,etc). Diz-se que X tem distribuição de Poisson de parâmetro $\lambda > 0$, $X \sim Poisson(\lambda)$ se a sua função de probabilidade é dada por:

$$P(X = x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!}, & x = 0, 1, 2, \ldots \\ 0, & c.c. \end{cases}$$

Valor esperado e variância:

$$E(X) = V(X) = \lambda$$

Nota - Verificar que a soma das probabilidades é 1:

$$\sum_{x=0}^{+\infty} P(X = x) = e^{-\lambda} \sum_{x=0}^{+\infty} \frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1$$

Demonstração para o valor esperado:

$$\begin{aligned} E(X) &= \sum_{x=0}^{+\infty} x\, e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{+\infty} x\, \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{+\infty} x\, \frac{\lambda\, \lambda^{x-1}}{x\,(x-1)!} = \\ &= e^{-\lambda} \lambda \sum_{x-1=0}^{+\infty} \frac{\lambda^{x-1}}{(x-1)!} = e^{-\lambda} \lambda\, e^{\lambda} = \lambda \end{aligned}$$

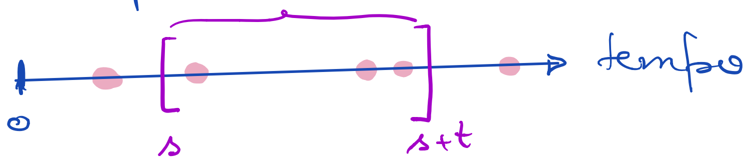
Cálculo de probabilidades:

- Fórmula
- Programas
- Tabelas (Função de distribuição para alguns valores de $\lambda \leq 40$)

Nota:
Podemos também usar a distribuição de Poisson para calcular probabilidades aproximadas da distribuição binomial quando n grande e p pequeno (regra prática: $n > 20$ e $p < 0.1$).

Processo de Poisson:

Suponha que se registam os instantes onde ocorrem certos fenómenos ao longo do tempo t unidades de tempo



Este denomina-se um processo estocástico.

Considere-se um intervalo de tempo t, que se pode dividir em subintervalos tão pequenos q.to se queira. Seja

$N(t)$ = n.º ocorrências em $[0,t]$

$N(t)$ diz-se um Processo de Poisson de taxa λ, $PP(\lambda)$, $\lambda > 0$ se:

1. $N(0)=0$
2. N.º acontecimentos que ocorrem em intervalos disjuntos são independentes (incrementos independentes)
3. A dist. do n.º de ocorrências num intervalo só depende da amplitude do intervalo e não da sua localização (incrementos estacionários)
4. $\lim_{h\to 0} \frac{P(N(h)=1)}{h} = \lambda$
5. $\lim_{h\to 0} \frac{P(N(h)\geq 2)}{h} = 0$

Então $N(t) \sim Po(\lambda t)$

Exemplos de situações em que <u>pode</u> ser aplicada(o) a(o) distribuição (processo) de Poisson:

- número de acidentes por semana num determinado cruzamento ou secção de estrada (não tendo em conta os acidentes em cadeia...)
- número de clientes que chegam a uma loja ou serviço num determinado intervalo de tempo (não tendo em conta as chegadas em grupo...)
- número de defeitos em peças ou materiais produzidos continuamente (tecidos, fios, etc.)

3.8 Distribuição de Poisson: Exemplo

Exemplo: sabemos (por observação anterior) que o número médio de defeitos por m^2 de um certo tipo de tecido é 2. Se admitirmos que estamos nas condições do Processo de Poisson, então:

- n.ºde defeitos em 10 m^2: $X_{10} \sim Poisson(20)$
- n.ºde defeitos em 100 m^2: $X_{100} \sim Poisson(200)$

  $\vdots$
- n.ºde defeitos em t m^2: $X_t \sim Poisson(2\,t)$

Qual é a probabilidade de não haver defeitos numa peça com 5 m^2? E de haver pelo menos 10 defeitos?

$X_5 \sim Poisson(10)$

$P(X_5 = 0) = e^{-10} \simeq 0$

$P(X_5 \geq 10) = 1 - P(X_5 < 10) = 1 - P(X_5 \leq 9) = 1 - F_{X_5}(9) =$
$1 - 0.4579 = 0.5421$ (tabela)

Exercício

TPC:

Exercício: Um canal de comunicação digital tem uma taxa de erro de 10^{-7}, o que significa que a probabilidade de um bit (zero ou um) chegar trocado ao destino é 10^{-7}. Em relação a uma transmissão de 10 milhões de bits (1.25 MB) calcular:

a) a probabilidade de a transmissão ser efectuada sem erros.

b) a probabilidade de no máximo ocorrerem 4 erros na transmissão.

c) o valor esperado, o valor mais provável, e a mediana do número de erros que ocorrem nessa transmissão.

Soluções: a) 0.3679 b) 0.9963 c) 1; 0 e 1; 1