

Capítulo II – Introdução à Teoria da Probabilidade

- **Não serão abordados** neste ano lectivo os assuntos expostos nos slides:
 - 65 a 75
 - 114
 - 137 a 139 e
 - 169
- **Serão tratados apenas na prática** os assuntos expostos nos slides 77 a 83.

Noções Preliminares

Definição 1

Fenómenos aleatórios são fenômenos sujeitos à influência do acaso e, como tal, fora do alcance do observador.

Fenómenos aleatórios são caracterizados pela sua:
imprevisibilidade e regularidade estatística

Definição 2

Experiência aleatória é todo o procedimento que verifica as seguintes propriedades:

- pode repetir-se um grande número de vezes nas mesmas condições ou pelo menos em condições semelhantes;
- a sua realização dá um resultado de entre um conjunto de resultados possíveis;
- cada um dos resultados da experiência é imprevisível mas é possível considerar “estabilidade na frequência da sua ocorrência”.

Exemplos de experiências aleatórias

- 1 lançamento de dois dados e registo do número de pontos que sai;
- 2 lançamento de uma moeda e observação da face que fica voltada para cima;
- 3 contagem do número mensal de acidentes de automóvel numa autoestrada;
- 4 registo do tempo de vida de uma pessoa, em anos;
- 5 registo do tempo de trabalho de uma máquina até à primeira avaria.

Definição 3

Espaço de resultados ou **espaço amostra** é o conjunto de todos os resultados possíveis associados a uma experiência aleatória – representa-se por Ω .

Para os exemplos anteriores tem-se

- 1 $\Omega = \{(1, 1), (1, 2), (1, 3), \dots, (6, 5), (6, 6)\}$;
- 2 $\Omega = \{\text{'face valor'}, \text{'face país'}\} = \{\text{'FV'}, \text{'FP'}\} = \{1, 0\}$;
- 3 $\Omega = \mathbb{N}_0$;
- 4 $\Omega = \mathbb{N}$;
- 5 $\Omega = \mathbb{R}^+$.

Definição 4

Acontecimento aleatório é qualquer subconjunto do espaço de resultados.

Seja Ω o espaço de resultados associado a uma experiência aleatória.

- Diz-se que $A \subset \Omega$ **se realizou** se o resultado, ω , da experiência é um elemento de A , i.e., $\omega \in A$.
- $A \subset B$, diz-se A **subacontecimento** de B , se e só se a realização de A implica a realização de B ;
- A^c ou \bar{A} diz-se **acontecimento complementar** ou **contrário** a A , é o conjunto de todos os elementos de Ω que não estão em A ;

Álgebra dos acontecimentos (cont.)

- $A \cup B$, diz-se **união** de A com B , é o acontecimento que consiste na realização de pelo menos um dos acontecimentos.
- AB ou $A \cap B$, diz-se **produto** ou **intersecção**, é o acontecimento que se realiza apenas quando ambos os acontecimentos se realizam.
- Os acontecimentos A e B dizem-se **mutuamente exclusivos** ou **incompatíveis** se e só se a realização de um implica a não realização do outro, i.e., **se e só se** $AB = \emptyset$.
- $A - B = A \cap \bar{B}$ diz-se **diferença** dos acontecimentos A e B é o acontecimento que se realiza se e só se A se realiza sem que B se realize.
- \emptyset diz-se **acontecimento impossível**.
- Ω diz-se **acontecimento certo**.

Álgebra dos acontecimentos

Vamos recordar algumas **propriedades** das operações sobre acontecimentos (procure mais algumas...):

Propriedade

Interpretação

Associatividade

$$(A \cap B) \cap C = A \cap (B \cap C)$$

$$(A \cup B) \cup C = A \cup (B \cup C)$$

Comutatividade

$$A \cap B = B \cap A$$

$$A \cup B = B \cup A$$

Distributividade

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

Leis de Morgan

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

...

...

Probabilidade de um acontecimento

Definição 5–Definição clássica: Laplace (séc. XIX)

Sob a hipótese de que **todos os casos são igualmente prováveis ou possíveis (princípio da simetria)**.

Probabilidade de realização de um acontecimento A

$$P = \frac{\text{número de casos favoráveis a } A}{\text{número total de casos possíveis}}$$

Definição 6–Definição frequencista

Considere-se n repetições de uma experiência aleatória; n_A o nº de vezes que se verificou A . Para n “grande” tem-se para as frequências relativas

$$f_n(A) = n_A/n \approx P$$

A probabilidade é então interpretada como frequência limite.

Probabilidade de um acontecimento

Ω — espaço de resultados associado a uma experiência aleatória.

Definição 7—Definição de Probabilidade: Axiomática de Kolmogorov

Probabilidade, P , é uma aplicação que a cada acontecimento de Ω associa um número real satisfazendo o seguinte conjunto de axiomas:

$$\mathbf{A1)} \quad P(A) \geq 0 \quad \forall A \subset \Omega;$$

$$\mathbf{A2)} \quad P(\Omega) = 1;$$

$$\mathbf{A3)} \quad P(A \cup B) = P(A) + P(B) \quad \text{se} \quad A \cap B = \emptyset. \quad (\text{Axioma das probabilidades totais}).$$

Se Ω é infinito,

$$\mathbf{A3^*)} \quad P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) \quad \text{se} \quad A_i \cap A_j = \emptyset, \quad i \neq j \quad (\text{Axioma completo das probabilidades totais}).$$

Propriedades da probabilidade

- 1 $P(\bar{A}) = 1 - P(A)$.
- 2 $P(\emptyset) = 0$.
- 3 $A \subset B \Rightarrow P(A) \leq P(B)$.
- 4 $P(A) \leq 1$.
- 5 $P(A - B) = P(A \cap \bar{B}) = P(A) - P(A \cap B)$.
- 6 Se $B \subset A \Rightarrow P(A - B) = P(A) - P(B)$.
- 7 Sejam A_1, \dots, A_n acontecimentos mutuamente exclusivos então
 $P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$
- 8 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Propriedades da probabilidade (cont.)

- 1 $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$
- 2 Generalização: Sejam A_1, A_2, \dots, A_n acontecimentos quaisquer
$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - \dots - P(A_{n-1} \cap A_n) + P(A_1 \cap A_2 \cap A_3) + \dots + P(A_{n-2} \cap A_{n-1} \cap A_n) + \dots + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_n).$$

Exercício 1

Sejam A , B e C acontecimentos definidos num espaço de resultados Ω tais que

$$P(A) = P(B) = P(C) = \frac{1}{4}; P(A \cap B) = P(B \cap C) = 0 \text{ e } P(A \cap C) = \frac{1}{8}.$$

Calcule, justificando, a probabilidade de se verificar pelo menos um dos acontecimentos A , B ou C .

Definição 8–Definição de Probabilidade Condicional

Dados os acontecimentos A e B definidos em Ω , a probabilidade de A se realizar sabendo que B se realizou, ou seja, a probabilidade condicional de A dado B ou probabilidade de A se B representa-se por $P(A|B)$, com $P(B) > 0$ e define-se como

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(AB)}{P(B)}$$

Desta definição resulta o seguinte teorema:

Teorema 1–Teorema da probabilidade composta

Se $P(A) > 0$ e $P(B) > 0$,

$$P(A \cap B) \equiv P(AB) = P(A) P(B|A) = P(B) P(A|B)$$

Definição 9

Dois acontecimentos A e B dizem-se mutuamente **independentes** se e só se

$$P(A \cap B) = P(A) P(B).$$

Da definição 9 **conclui-se** que se A e B são independentes então $P(A|B) = P(A)$ se $P(B) > 0$ e $P(B|A) = P(B)$ se $P(A) > 0$.

Teorema 2

Se A e B são independentes
 A e \bar{B} , \bar{A} e B e \bar{A} e \bar{B} , também são independentes.

Nota: Independência não é equivalente a exclusividade mútua.

Resultado:

Se $P(A) > 0$ e $P(B) > 0$ e A e B independentes \Rightarrow A e B são não mutuamente exclusivos.

Obviamente o contra-recíproco é verdadeiro.

Generalização a três acontecimentos

Sejam A , B , C tais que $P(A) > 0$, $P(B) > 0$ e $P(C) > 0$, tem-se,

$$P(ABC) = P(A)P(B|A)P(C|AB) = P(B)P(C|B)P(A|BC) = \\ = P(C)P(A|C)P(B|AC).$$

Definição 10—Independência de três acontecimentos

Os acontecimentos A , B e C dizem-se **mutuamente independentes** ou apenas **independentes** se e só se

$$P(ABC) = P(A)P(B)P(C); \quad P(AB) = P(A)P(B); \\ P(AC) = P(A)P(C); \quad P(BC) = P(B)P(C).$$

Nota: A independência par a par não assegura independência de um conjunto de acontecimentos.

Exercício 2

Uma empresa produz concentrado de tomate recorrendo a três processos de fabrico e embalagem. Sabe-se que 20% da produção e embalagem de concentrado provém do processo A, 30% do processo B e 50% do processo C.

Nalgumas embalagens daquele concentrado tem-se verificado a ocorrência de deficiências. Sabe-se 1% das embalagens provenientes do processo A, 2% das provenientes do processo B e 8% das provenientes do processo C, respectivamente, têm deficiência.

- 1 Qual a percentagem de embalagens, produzidas naquela empresa, que apresentam deficiências?
- 2 Verifica-se que uma embalagem escolhida ao acaso apresenta deficiências. Qual a probabilidade de ter sido fabricada e embalada pelo processo A?

Teorema da probabilidade total

A resolução da [Pergunta 1](#). baseia-se no seguinte teorema

Teorema 3—Teorema da probabilidade total

Sejam A_1, A_2, \dots, A_n acontecimentos definindo uma **partição sobre Ω** , i.e.,

$$A_1 \cup A_2 \cup \dots \cup A_n = \Omega \quad \text{e} \quad A_i \cap A_j = \emptyset, \quad \forall i, j, \quad i \neq j.$$

Se $P(A_i) > 0$, então para qualquer acontecimento $B \subset \Omega$ tem-se

$$P(B) = \sum_{i=1}^n P(A_i) P(B|A_i).$$

Teorema de Bayes

Relativamente à [Pergunta 2](#). do exercício anterior, pretendemos *atualizar* a probabilidade de um acontecimento *a priori*, à custa da informação *a posteriori*.

O seguinte teorema formaliza a resposta à questão:

Teorema 4—Teorema de Bayes

Sejam A_1, A_2, \dots, A_n acontecimentos formando uma partição de Ω , onde $P(A_i) > 0$. Seja B um outro acontecimento de Ω , tal que $P(B) > 0$. Então para $k = 1, \dots, n$ tem-se

$$P(A_k|B) = \frac{P(A_k) \cdot P(B|A_k)}{\sum_{i=1}^n P(A_i) \cdot P(B|A_i)}$$

Variável aleatória

Muitas vezes o resultado de uma experiência aleatória não é numérico ou sendo-o não interessa lidar com os resultados possíveis de Ω , mas pretende-se associar-lhe uma quantidade numérica.

Exemplo - lançamento de dois dados e soma dos pontos das faces.

É então mais cómodo associar a cada acontecimento um número, definido de acordo com o objectivo do estudo.

Chama-se **variável aleatória** a esta correspondência.

Definição 11

Chama-se **variável aleatória (v.a.)** e costuma representar-se por **X** , a **uma função com domínio Ω e contradomínio em \mathbb{R}** , cujo valor é determinado pelo resultado de uma experiência aleatória, i.e.,

$$X : \Omega \rightarrow \mathbb{R}$$

$$X(\omega) = x$$

Tipos de variáveis aleatórias

Variáveis aleatórias **discretas** se assumem um conjunto finito ou infinito numerável de valores.

Exemplos:

- número de pintas que sai no lançamento de um dado;
- registo, a intervalos regulares, do número de pessoas em fila espera na caixa de um supermercado;

Variáveis aleatórias **contínuas** são as susceptíveis de tomar qualquer valor real num dado intervalo, que pode ser a recta real

(definição mais rigorosa será dada à frente)

Exemplos:

- o peso de um indivíduo;
- o comprimento de uma folha de uma planta.

Variáveis aleatórias

Mas ... aos valores de uma variável aleatória X pretendemos associar uma probabilidade P_X ou, mais simplesmente, P

Isto consegue-se muito facilmente definindo uma função real de variável real do seguinte modo:

Definição 12

Chama-se **função de distribuição cumulativa** ou apenas **função de distribuição** associada à variável aleatória X e representa-se por F ou F_X , à aplicação

$$F : \mathbb{R} \rightarrow [0, 1] \quad \text{tal que} \quad F(x) = P[X \leq x].$$

Propriedades da função de distribuição

1. $0 \leq F(x) \leq 1$
2. $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$
 $F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1.$
3. F é uma função monótona não decrescente, i.e., dados dois números reais x_1 e x_2 tais que $x_1 < x_2$, tem-se $F(x_1) \leq F(x_2)$
4. $F(x)$ é contínua à direita, i.e., $\lim_{x \rightarrow x_0^+} F(x) = F(x_0).$
5. $P(X = a) = F(a) - F(a^-)$ onde $F(a^-) = \lim_{x \rightarrow a^-} F(x)$

Função de distribuição e Probabilidade

O conhecimento da função de distribuição $F(\cdot)$ é equivalente ao conhecimento da lei de probabilidade $P_X = P$.

Como $F(x) = P[X \leq x] \rightarrow$ conhecer $P \Rightarrow$ conhecer $F(x)$.
Reciprocamente ... conhecer $F(x)$, permite calcular a probabilidade dos vários tipos de intervalos.

- $P(X < x) = P(X \leq x) - P(X = x) = F(x^-)$;
- $P(X \geq x) = 1 - P(X < x) = 1 - F(x^-)$;
- $P(X > x) = 1 - P(X \leq x) = 1 - F(x)$;
- $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$;
- $P(a < X < b) = P(X < b) - P(X \leq a) = F(b^-) - F(a)$;
- $P(a \leq X \leq b) = P(X \leq b) - P(X < a) = F(b) - F(a^-)$;
- $P(a \leq X < b) = P(X < b) - P(X < a) = F(b^-) - F(a^-)$.

Variáveis aleatórias

Vamos agora ver como calcular a função de distribuição cumulativa e consequentemente a probabilidade para cada um dos tipos de variáveis aleatórias caracterizados atrás:

- **variáveis aleatórias discretas** e
- **variáveis aleatórias contínuas**

Relembre-se que:

Uma variável aleatória diz-se **discreta** se toma um número finito ou uma infinidade numerável de valores.

Variáveis aleatórias discretas

Seja X uma v.a. tomando k valores, x_1, \dots, x_k , cada um deles com probabilidades p_1, \dots, p_k , respectivamente, i.e.,
 $p_i = P[X = x_i]$, ($i = 1, \dots, k$).

Definição 13

Chama-se **função massa de probabilidade** da v.a. X à **aplicação** que a cada valor $x_j \rightarrow p_j$, tal que

$$p_i = P[X = x_i]$$

A **função massa de probabilidade** satisfaz:

$$p_i \geq 0, i = 1, \dots, k \quad \sum_{i=1}^k p_i = 1.$$

Nota: Se a v.a. tomar uma infinidade numerável de valores tem-se

$$p_i \geq 0, \forall i \geq 1 \quad \sum_{i=1}^{\infty} p_i = 1.$$

Variáveis aleatórias discretas

Chama-se **distribuição de probabilidade** da v.a. X ao conjunto de pares $(x_i, p_i)_{i=1, \dots, k}$.

Habitualmente a **lei (distribuição) de probabilidade** da v.a. X dispõe-se na forma:

$$X = \begin{cases} x_1 & x_2 & \dots & x_k \\ p_1 & p_2 & \dots & p_k \end{cases} \quad \text{ou} \quad \frac{x_i}{P[X = x_i]} \mid \begin{array}{cccc} x_1 & x_2 & \dots & x_k \\ p_1 & p_2 & \dots & p_k \end{array}$$

A distribuição de probabilidade da v.a. discreta permite calcular facilmente a **função de distribuição cumulativa F_X**

$$F_X(x) = P[X \leq x] = \sum_{x_i \leq x} P[X = x_i],$$

ou seja temos a probabilidade cumulativa associada à variável X calculada em qualquer $x \in \mathbb{R}$.

Variáveis aleatórias contínuas

Definição 14

Uma variável aleatória diz-se **contínua** se existe uma função real de variável real, f , não negativa, tal que

$$F(x) = P[X \leq x] = \int_{-\infty}^x f(t) dt \quad -\infty < x < \infty$$

Nota:

$$\int_a^b f(x) dx = F(b) - F(a) = P(a < X \leq b) = P(a \leq X \leq b) \dots$$

Definição 15

A função f diz-se **função densidade de probabilidade** ou apenas **função densidade**. Deve verificar as seguintes condições:

$$f(x) \geq 0 \quad \forall x \in \mathbb{R}; \quad \int_{-\infty}^{+\infty} f(x) dx = 1$$

Exercício 3

O número de micro-ondas vendidos diariamente num estabelecimento é uma variável aleatória, X , com a seguinte distribuição de probabilidade

$$X = \begin{cases} 0 & 1 & 2 & 3 & 4 \\ 0.3 & 0.3 & 0.2 & 0.1 & 0.1 \end{cases}$$

- Determine a função de distribuição cumulativa de X ; represente-a graficamente.
- Determine $P[1 \leq X \leq 3]$. Interprete esta probabilidade.

Exercício 4

Seja X a v.a. que designa o tempo de vida (em anos) de um dado equipamento, cuja função densidade é

$$f(x) = \begin{cases} \frac{1}{5} e^{-x/5} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

- Mostre que f é de facto uma função densidade.
- Determine a função de distribuição cumulativa de X ; represente-a graficamente.
- Qual a probabilidade de esse equipamento durar entre 1 e 3 anos?

Variáveis aleatórias

Recordemos que:

- No caso de uma **variável aleatória discreta** a **função de distribuição cumulativa** é uma **função em escada**, onde os pontos de salto são os valores onde a v.a. está definida.
- No caso de uma **variável aleatória contínua** a **função de distribuição cumulativa** é uma **função contínua**.

Além de termos interesse em calcular probabilidades associadas a uma variável aleatória,

vamos agora calcular “indicadores” que a caracterizam – são valores reais habitualmente designados por **parâmetros**.

Definição 16

Dada uma v.a. X chama-se **valor médio, esperança matemática, valor esperado** ou **média** e representa-se por $E[X]$, μ_X ou simplesmente μ a

$$E[X] = \sum_{i=1}^n x_i p_i \quad X \text{ é v.a. discreta com distribuição } (x_i, p_i)$$

$$E[X] = \int_{-\infty}^{+\infty} x f(x) dx \quad X \text{ é v.a. contínua com f.d.p. } f(x)$$

Se X for v.a. discreta com uma infinidade numerável de valores tem-se $E[X] = \sum_{i=1}^{\infty} x_i p_i$. Neste caso só existe valor médio se “aquela soma infinita existir”.

Analogamente, no caso contínuo, só existe valor médio, $E[X] = \int_{-\infty}^{\infty} x f(x) dx$, se o integral for absolutamente convergente.

Valor Médio de uma função de X

Se X é uma v.a. e $Y = \varphi(X)$ é uma função real de variável real, define-se **valor médio de $\varphi(X)$** como

$$E[\varphi(X)] = \sum_i \varphi(x_i) p_i \quad X \text{ é v.a. discreta com distribuição } (x_i, p_i)$$

$$E[\varphi(X)] = \int_{-\infty}^{+\infty} \varphi(x) f(x) dx \quad X \text{ é v.a. contínua com f.d.p. } f(x)$$

Mais uma vez, para que exista valor médio exige-se que exista aquela “soma infinita” (no caso de se tratar de uma v.a. discreta com uma infinidade de valores) ou a convergência absoluta do integral.

1. Linearidade

- $E[a] = a$.
- $E[a + bX] = a + b E[X]$.
- $E[\varphi(X) + \psi(X)] = E[\varphi(X)] + E[\psi(X)]$

2. Positividade

Se $X \geq 0$, i.e. a variável toma apenas valores ≥ 0 ,
tem-se $E[X] \geq 0$.

3.
$$\inf(X) \leq E[X] \leq \sup(X)$$

Variância e Desvio Padrão

Definição 17

Chama-se **variância** de uma variável aleatória X e representa-se por $Var[X]$, σ_X^2 ou apenas σ^2 a

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

$\sigma_X = \sqrt{Var[X]}$ chama-se **desvio padrão**.

Exercício 5

Verifique que se pode escrever $Var[X] = E[X^2] - \mu^2$

Propriedades da variância e do desvio padrão

1. $Var[X] \geq 0$
2. $Var[a + b X] = b^2 Var[X]$.

Para o **desvio padrão** tem-se $\sigma_{(a+b X)} = |b| \sigma_X$

Voltemos ao Exercício 3

O número de micro-ondas, de uma dada marca, vendidos diariamente num estabelecimento é uma variável aleatória, X , com a seguinte distribuição de probabilidade

$$X = \begin{cases} 0 & 1 & 2 & 3 & 4 \\ 0.3 & 0.3 & 0.2 & 0.1 & 0.1 \end{cases}$$

- Qual o valor esperado do número de micro-ondas vendidos por dia?
- Se cada micro-ondas é vendido por 85 Euros qual é a distribuição de probabilidade da receita bruta da venda de micro-ondas por dia.
- Calcule a receita bruta esperada da venda de micro-ondas por dia.

Exercício 6

Considere X a v.a. que designa a duração (em minutos) de cada chamada telefónica efectuada num certo local, cuja função densidade é

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

- Calcule a duração média de uma chamada telefónica.
- Calcule a variância de X .
- Se o preço de cada minuto de conversação for 60 cêntimos, qual é, em média, o preço de cada chamada telefónica.

Quantis e Mediana de uma variável aleatória

Definição 18

Dada uma v.a. X chama-se **quantil de probabilidade p** e representa-se por χ_p o menor valor da variável aleatória X tal que $F_X(\chi_p) \geq p$.

Se $p = 0.5$, chama-se **mediana de X** , representa-se por $\chi_{0.5}$, e é o menor valor da variável tal que $F_X(\chi_{0.5}) \geq 0.5$.

Notas:

- Se X é v.a. contínua o quantil de probabilidade p é o valor χ_p tal que $F_X(\chi_p) = p$.
- Então se X é uma v.a. contínua a mediana $\chi_{0.5}$, é a solução de $F_X(x) = 0.5 \iff \int_{-\infty}^{\chi_{0.5}} f(t)dt = 0.5$.

Vectores aleatórios

Muitas vezes pretendemos associar a cada resultado de uma experiência aleatória $k \geq 2$ atributos numéricos. Obtemos então um vector (x_1, \dots, x_k) , realização do **vector aleatório** (X_1, \dots, X_k) .

Iremos referir-nos apenas ao caso $k = 2$, portanto trataremos de pares aleatórios

Exemplos Pretendemos registar:

- a quantidade de precipitado P e o volume V de gás numa experiência química
- para uma árvore seleccionada ao acaso, a altura e o diâmetro do tronco à altura do peito . . .

Definição 19

Chama-se **par aleatório** (X, Y) à aplicação

$$\begin{aligned}(X, Y) : \Omega &\rightarrow \mathbb{R}^2 \\ \omega &\rightarrow (x, y)\end{aligned}$$

Tipos de pares aleatórios que vamos estudar:

- Par aleatório **discreto** \Rightarrow componentes são ambas variáveis aleatórias discretas;
- Par aleatório **contínuo** \Rightarrow componentes são ambas variáveis aleatórias contínuas.

Pares aleatórios discretos

(X, Y) diz-se um par aleatório **discreto** se toma os valores (x_i, y_j) com probabilidades $p_{ij} = P[X = x_i, Y = y_j]$.

Definição 20

Chama-se **distribuição de probabilidades conjunta** do par (X, Y) aos valores (x_i, y_j) e respectivas probabilidades p_{ij}

p_{ij} é chamada **função massa de probabilidade conjunta** e deve verificar as seguintes condições:

$$p_{ij} \geq 0 \quad \forall (i, j) \quad \text{e} \quad \sum_i \sum_j p_{ij} = 1.$$

Pares aleatórios discretos

Um modo cómodo de representar a **distribuição de probabilidades conjuntas** de um par aleatório discreto (X, Y) é na forma de um quadro

X	Y	y_1	y_2	...	y_n	
x_1		p_{11}	p_{12}	...	p_{1n}	$p_{1\bullet}$
x_2		p_{21}	p_{22}	...	p_{2n}	$p_{2\bullet}$
.	
.	
.	
x_m		p_{m1}	p_{m2}	...	p_{mn}	$p_{m\bullet}$
		$p_{\bullet 1}$	$p_{\bullet 2}$...	$p_{\bullet n}$	1

$p_{i\bullet} = \sum_{j=1}^n p_{ij}$ e $p_{\bullet j} = \sum_{i=1}^m p_{ij}$ chamam-se **probabilidades marginais** de X e Y respectivamente.

Definição 21

A **probabilidade condicional** de X dado $Y = y_j$ (fixo) com $P[Y = y_j] > 0$ é definida como

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\bullet j}},$$

Definição 22

Do mesmo modo a **probabilidade condicional** de Y dado $X = x_i$ (fixo) com $P[X = x_i] > 0$ é definida como

$$P(Y = y_j | X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{p_{i\bullet}}.$$

Definição 23

Um par aleatório (X, Y) diz-se **contínuo** se existir uma função $f(x, y)$, chamada **função densidade (de probabilidade) conjunta**, que verifica as seguintes condições:

- $f(x, y) \geq 0$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1.$

Dado $A \subset \mathbb{R}^2$ tem-se $P[(X, Y) \in A] = \int \int_A f(x, y) dx dy.$

Definição 24

A **densidade marginal de X** é definida como $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$

e a

densidade marginal de Y como $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$

Definição 25

Define-se **densidade condicional** de X dado $Y = y$, fixo, como

$$f_{X|Y=y}(x) = \frac{f(x, y)}{f_Y(y)}, \quad f_Y(y) > 0$$

Definição 26

Define-se **densidade condicional** de Y dado $X = x$, fixo, como

$$f_{Y|X=x}(y) = \frac{f(x, y)}{f_X(x)}, \quad f_X(x) > 0$$

Definição 27

Dado o par aleatório (X, Y) diz-se que as variáveis X e Y são **independentes** se e só se

- $p_{ij} = p_{i\bullet} \times p_{\bullet j} \quad \forall i, j,$ no caso de (X, Y) ser um **par aleatório discreto**
- $f(x, y) = f_X(x) \times f_Y(y) \quad \forall (x, y) \in \mathbb{R}^2$ no caso de (X, Y) ser um **par aleatório contínuo**.

Definição 28

Dado o par aleatório (X, Y) , e $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, define-se

$$E[g(X, Y)] = \sum_i \sum_j g(x_i, y_j) p_{ij}, \quad \text{no caso discreto}$$

$$E[g(X, Y)] = \int \int_{\mathbb{R}^2} g(x, y) f(x, y) dx dy, \quad \text{no caso contínuo.}$$

Propriedades do Valor Médio

1. **Aditividade** $E[X \pm Y] = E[X] \pm E[Y]$

2. **Desigualdade de Schwarz** Se $E[X^2]$ e $E[Y^2]$ existem então
 $E^2[XY] \leq E[X^2]E[Y^2]$.

Corolário: $E^2[X] \leq E[X^2]$

Nota: se $E[X^2]$ existe \implies existe $E[X]$.

3. Se X e Y variáveis aleatórias independentes



$$E[XY] = E[X]E[Y]$$

Valor Médio - propriedades

Nota:

O recíproco da propriedade 3. não é verdadeiro:

Verifique que se X e Y são v. a.'s com a seguinte distribuição de probabilidades

X	Y	-1	0	1
0		0	1/3	0
1		1/3	0	1/3

tem-se $E[XY] = E[X] \times E[Y]$ e no entanto X e Y não são independentes. **Verifique!**

A covariância

Definição 29

Dado o par aleatório (X, Y) chama-se **covariância** de X e Y a

$$\mathbf{Cov}[X, Y] \equiv \sigma_{XY} = \mathbf{E}[(X - \mu_X)(Y - \mu_Y)]$$

Exercício 7

Verifique que $\mathbf{Cov}[X, Y] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$

Propriedades

1. Sejam X e Y variáveis aleatórias.

$$\text{Var}[X \pm Y] = \text{Var}[X] + \text{Var}[Y] \pm 2\text{Cov}[X, Y]$$

2. Se X e Y são variáveis aleatórias **independentes**

$$\text{Var}[X \pm Y] = \text{Var}[X] + \text{Var}[Y]$$

3. Se X e Y são v. a.'s independentes $\implies \text{Cov}[X, Y] = 0$.

Nota: O recíproco não é verdadeiro.

4. $\text{Cov}[a + bX, c + dY] = bd \text{Cov}[X, Y]$.

5. $|\text{Cov}[X, Y]| \leq \sigma_X \sigma_Y$.

O coeficiente de correlação; propriedades

Definição 30

Chama-se **coeficiente de correlação** de X e Y e representa-se por ρ ou $\rho_{X,Y}$ a

$$\rho \equiv \rho_{X,Y} = \frac{\text{Cov}[X, Y]}{\sigma_X \sigma_Y}$$

($\sigma_X > 0$ e $\sigma_Y > 0$).

Propriedades do coeficiente de correlação

1. $-1 \leq \rho_{X,Y} \leq 1$
2. Se X e Y são v. a. independentes $\implies \rho_{X,Y} = 0$.
3. $\rho_{a+bX, c+dY} = \begin{cases} \rho_{X,Y} & \text{se } bd > 0 \\ -\rho_{X,Y} & \text{se } bd < 0 \end{cases}$

Momentos e função geradora de momentos

O cálculo do valor médio e da variância de uma v.a. X e ainda propriedades de pares aleatórios (ou genericamente vectores aleatórios) podem ser abordados de forma uniformizadora usando uma função adequada (quando ela está definida).

Considere-se uma função associada à v.a. X que vamos representar por M_X

$$M_X : \mathbb{R} \longrightarrow \mathbb{R}$$

tal que

$$M_X(t) = E[e^{tX}] \quad (t \in \mathbb{R})$$

Exercício 8

Considere as seguintes variáveis aleatórias:

- X , variável aleatória discreta, associada ao lançamento de uma moeda equilibrada.
- X , variável aleatória contínua, com função densidade

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

Para cada uma calcule $M_X(t)$, com $t \in \mathbb{R}$:

Função geradora de momentos

Tem-se o seguinte resultado:

$$M'_X(0) \equiv \left. \frac{dM_X}{dt} \right|_{t=0} = E[X] \quad \text{e} \quad M''_X(0) \equiv \left. \frac{d^2 M_X}{dt^2} \right|_{t=0} = E[X^2]$$

Nota:

Esta função, a que se chama **função geradora de momentos**, pode ser então usada para determinar $E[X]$ e $\text{Var}[X]$, calculando a primeira e segunda derivadas em $t = 0$ (se existirem).

Para as variáveis aleatórias indicadas no **exercício** do slide anterior, calcule $E[X]$ e $\text{Var}[X]$, com recurso a M_X .

Propriedades da função geradora de momentos

1. $M_{a+b X}(t) = e^{at} M_X(bt)$.

2. **Teorema 5–Teorema da unicidade**

Se para duas v.a. X e Y se verifica $M_X(t) = M_Y(t)$ então X e Y têm a mesma função de distribuição.

Reciprocamente, se existir a função geradora de momentos, ela é única.

3. Se X e Y são variáveis aleatórias independentes

$$M_{X+Y}(t) = M_X(t) \times M_Y(t)$$

Nota: Mais adiante esta propriedade será de grande utilidade.

Principais Modelos (Distribuições) Discretos

- Distribuição uniforme discreta
- Distribuição de Bernoulli e binomial
- Distribuição geométrica
- Distribuição hipergeométrica
- Distribuição de Poisson

A distribuição uniforme discreta

Definição 31

Uma v.a. X diz-se ter **distribuição uniforme discreta** se $P(X = x_i) = 1/k$, $i = 1, \dots, k$, i.e., se toma os valores

com probabilidades x_1, x_2, \dots, x_k
 $1/k, 1/k, \dots, 1/k$

Valor médio, variância e função geradora de momentos

$$E[X] = \frac{1}{k} \sum_{i=1}^k x_i; \quad \text{Var}[X] = \frac{1}{k} \sum_{i=1}^k (x_i - \mu)^2; \quad M_X(t) = \frac{1}{k} \sum_{i=1}^k e^{tx_i}.$$

A distribuição uniforme discreta

Caso particular

$$\text{Se } X = \begin{cases} 1 & 2 & \dots & n \\ 1/n & 1/n & \dots & 1/n \end{cases}$$

$$E[X] = \frac{n+1}{2}; \quad \text{Var}[X] = \frac{n^2-1}{12} \quad \text{e} \quad M_X(t) = \frac{e^t(1-e^{nt})}{n(1-e^t)}, \quad t \neq 0$$

A distribuição de Bernoulli

Considere-se que se **realiza uma experiência aleatória** na qual o objectivo é apenas registar se ocorreu:

{	realização de um acontecimento	<u>sucesso</u>
	não realização do acontecimento	<u>insucesso</u>

Exemplos:

- o teste de uma dada droga num rato e o registo da reacção positiva ou negativa;
- a inspecção dos items numa linha de fabrico para observar se cada um é defeituoso ou não.

Cada uma das repetições sucessivas da experiência – **prova**.

Provas de Bernoulli

Diz-se que estamos perante **provas de Bernoulli independentes** se realizámos uma sucessão de provas satisfazendo:

- cada prova tem apenas um de dois resultados possíveis: **sucesso** ou **insucesso**.
- em cada prova a probabilidade de sucesso, p , permanece constante, sendo $q = 1 - p$, a probabilidade de insucesso.
- o resultado de cada prova é independente do resultado das restantes.

A distribuição de Bernoulli e a distribuição binomial

Definição 32

Chama-se **variável aleatória de Bernoulli** à variável X , associada ao resultado de cada prova de Bernoulli e considera-se

- $X = 1$, com probabilidade p , se há sucesso;
- $X = 0$, com probabilidade $1 - p = q$, se há insucesso.

Definição 33

A v.a. X que conta o número de sucessos em n provas de Bernoulli independentes chama-se **variável aleatória binomial**, diz-se ter **distribuição binomial** e representa-se por $X \sim \mathcal{B}(n, p)$.

A distribuição de binomial

Exemplo

Numa experiência colocam-se 5 bolbos de junquillo a germinar, de um pacote com uma garantia de germinação de 40% dos bolbos. Qual a probabilidade de, desses 5 bolbos, 3 germinarem?

Como a germinação é independente de bolbo para bolbo, a probabilidade de **germinarem 3 bolbos** de entre os 5 é então

$$\binom{5}{3} (0.4)^3 (0.6)^2$$

A distribuição binomial

Então sendo X a v.a. que conta o número de sucessos em n provas de Bernoulli independentes, $X \sim \mathcal{B}(n, p)$, temos a

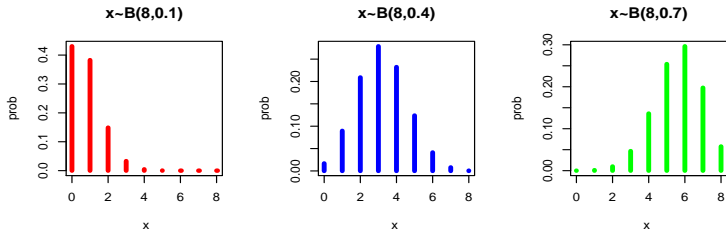
Caracterização da v.a. $X \sim \mathcal{B}(n, p)$:


$x = 0, 1, 2, \dots, n$ $\longrightarrow n^{\circ}$ de “sucessos” nas n provas

$P[X = x] = \binom{n}{x} p^x (1 - p)^{n-x}$ \longrightarrow probabilidade de se observarem x “sucessos”

A distribuição binomial—Exercício 9

Para $n = 8$ e vários valores de p , veja a função massa de probabilidade.



Sugestão: Consulte as folhas de Introdução ao software  e use os comandos (por exemplo, para obter o primeiro gráfico):

```
> x <- 0:8  
> plot(x,dbinom(x,size=8,prob=0.1),type="h", col = "red",  
lwd=4,xlab="x",main="X ~ B(8, 0.1)",ylab="prob")
```

A distribuição binomial

Valor médio, variância e função geradora de momentos
 $X \sim \mathcal{B}(n, p)$

$$E[X] = np; \quad \text{Var}[X] = npq; \quad M_X(t) = (p e^t + q)^n$$

Relação entre a distribuição do número de sucessos e de insucessos

$$X \sim B(n, p) \Rightarrow (n - X) \sim B(n, 1 - p).$$

Para valores de $n \leq 20(25)$, existem tabelas para o cálculo das probabilidades.

As tabelas que temos à disposição apresentam os valores da função de distribuição cumulativa.

A distribuição geométrica

Considere-se de novo que temos **provas de Bernoulli independentes**, **mas agora . . .**

o número de provas não é fixo pois ... **pretendemos ir realizando provas até ocorrer** pela **primeira vez** o “sucesso”.

Seja então **X** o **número de provas necessárias até que ocorra pela primeira vez o “sucesso”**. Diz-se que **X** tem **distribuição geométrica** e costuma representar-se por **$X \sim \mathcal{G}(p)$** .

Caracterização da v.a. $X \sim \mathcal{G}(p)$

$$P[X = x] = pq^{x-1} \quad x = 1, 2, \dots \quad 0 < p < 1 \quad q = 1 - p$$

A distribuição geométrica

Valor médio, variância e função geradora de momentos

$$M_X(t) = \frac{p e^t}{1 - qe^t} \quad (qe^t < 1); \quad E[X] = 1/p; \quad \text{Var}[X] = q/p^2$$

Observação:

Interpretando a distribuição geométrica como o número de provas que se vão realizando até se observar um "sucesso":

Se tiverem decorrido mais de m provas sem que se tenha verificado um "sucesso", a probabilidade de se ter de esperar mais de n provas para se observar um "sucesso" é a mesma caso se estivesse no início da experiência.

A distribuição geométrica

A observação anterior é apresentada no seguinte teorema:

Teorema 6 - Propriedade da falta de memória da distribuição geométrica

Se $X \sim \mathcal{G}(p)$ então sendo m e n inteiros positivos

$$P[X > m + n | X > m] = P[X > n]$$

Este teorema é muito fácil de provar, bastando recorrer ao conceito de **probabilidade condicional** e ainda ao seguinte resultado, fácil de provar

Resultado:

Se $X \sim \mathcal{G}(p)$ tem-se $P[X > n] = (1 - p)^n$, $n = 0, 1, 2, \dots$

A distribuição hipergeométrica

Mas ... há experiências nas quais a probabilidade de sucesso não se mantém constante, não sendo as provas independentes.

Exemplo

Num lote de 20 pneus enviados a um fornecedor sabe-se que há 6 defeituosos. Um cliente vai a esse fornecedor comprar 5 pneus. Qual a probabilidade de levar 2 defeituosos?

- O total de modos de seleccionar 5 pneus quaisquer do lote é $\binom{20}{5}$
- Há $\binom{6}{2}$ modos de seleccionar 2 defeituosos e, para cada um destes há $\binom{14}{3}$ modos de escolher 3 bons, para completar os 5.

Portanto ... a probabilidade de, dos 5 pneus escolhidos ao acaso, 2 serem defeituosos (e portanto 3 bons) é: $\frac{\binom{6}{2} \binom{14}{3}}{\binom{20}{5}}$

A distribuição hipergeométrica

Definição 34

Diz-se que temos uma **experiência hipergeométrica** se

dada uma população de dimensão

N com $\begin{cases} K \text{ "sucessos"} \\ N - K \text{ "insucessos"} \end{cases} \rightarrow$ extraímos, sem reposição n

Definição 35

A v.a. X que conta o número de sucessos numa experiência hipergeométrica **é uma v.a. hipergeométrica** de parâmetros N , n e K e costuma representar-se por $X \sim \mathcal{H}(N, n, K)$

A distribuição hipergeométrica

Qual a probabilidade de $\begin{cases} \text{dos } K & \text{seleccionar } x \\ \text{dos } N - K & \text{seleccionar } n - x \end{cases}$?

Seja $X \sim \mathcal{H}(N, n, K)$

$$P[X = x] = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}}, \quad \max(0, n - N + K) \leq x \leq \min(n, K)$$

Valor médio e variância de $X \sim \mathcal{H}(N, n, K)$

$$E[X] = n \frac{K}{N}; \quad \text{Var}[X] = n \frac{K}{N} \left(1 - \frac{K}{N}\right) \frac{N-n}{N-1}.$$

A distribuição hipergeométrica

Observação: Quando $N \gg n$, a probabilidade de sucesso em cada tiragem sem reposição varia muito pouco de prova para prova, então .

..

→ pode considerar-se a distribuição binomial como uma aproximação da distribuição hipergeométrica com $p = K/N$, i.e.,

Resultado:

Se N bastante maior que n tem-se

$$\mathcal{H}(N, n, K) \approx \mathcal{B}(n, p), \quad \text{com } p = K/N.$$

Como **regra prática**, pode considerar-se **boa a aproximação** para $n < N/10$.

A distribuição de Poisson

Considere que pretende contar, por exemplo, o número de:

- chamadas telefónicas recebidas numa central telefónica num certo intervalo de tempo;
- chegadas de clientes a uma bilheteira durante um certo período;
- chegadas de sinistrados a um banco de um hospital durante um certo período;
- dias que uma dada escola fecha durante o inverno;
- erros de tipografia por página;

Se a contagem do número de “sucessos” que ocorrem num dado intervalo de tempo ou num domínio específico, satisfaz as seguintes condições:

A distribuição de Poisson

- o número de “sucessos” que ocorrem num dado intervalo de tempo ou domínio é independente do número que ocorre em qualquer outro intervalo ou domínio disjunto do anterior;
- a probabilidade que o “sucesso” se verifique uma vez em qualquer intervalo muito curto (ou região muito pequena), de amplitude δ , é proporcional a δ , i.e, é igual a $\lambda\delta$ e não depende do número de sucessos que ocorrem fora desse intervalo ou região;
- a probabilidade de que o “sucesso” se verifique mais do que uma vez num intervalo ou domínio de amplitude muito pequena é ≈ 0 .

diz-se que estamos perante **experiências de Poisson** ou **um processo de Poisson**

A distribuição de Poisson

Definição 36

A v.a X que conta o número de sucessos numa experiência de Poisson diz-se ter **distribuição de Poisson** e depende apenas do parâmetro $\lambda \rightarrow$ número médio de sucessos que ocorrem no intervalo de tempo (ou na região especificada).

Representa-se por $X \sim \mathcal{P}(\lambda)$ e a lei de probabilidade é:

$$P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots, \quad \lambda > 0.$$

Nota: Facilmente se verifica que $P[X = x] \geq 0 \quad \forall x = 0, 1, 2, \dots$, mas para mostrar que $\sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} = 1$, são necessários conhecimentos sobre séries de funções que actualmente os alunos não possuem.

A distribuição de Poisson

Valor médio, variância e função geradora de momentos

$$M_X(t) = e^{\lambda(e^t - 1)} \quad E[X] = \lambda \quad \text{Var}[X] = \lambda.$$

Teorema 7–Teorema da estabilidade da soma

Se as v.a. X_i $i = 1, \dots, k$ são independentes e $X_i \sim \mathcal{P}(\lambda_i)$ então

$$\sum_{i=1}^k X_i \sim \mathcal{P}\left(\sum_{i=1}^k \lambda_i\right).$$

Existem tabelas da Poisson para consulta \rightarrow função de distribuição cumulativa.

A distribuição de Poisson

A distribuição de Poisson surge ainda como o limite da distribuição binomial quando $n \rightarrow \infty$ e $p \rightarrow 0$.

Teorema 8

Quando $n \rightarrow \infty$ e $p \rightarrow 0$, mantendo-se constante o produto np tem-se

$$X \sim B(n, p) \Rightarrow X \sim \mathcal{P}(\lambda) \quad \text{com } \lambda = np.$$

Regra prática Em geral, a distribuição de Poisson fornece uma **boa aproximação da distribuição binomial** quando $n \geq 20$ e $p \leq 0.05$

Principais Distribuições Contínuas

- Distribuição uniforme contínua
- Distribuição de Gauss ou normal
- Distribuição exponencial

A distribuição uniforme contínua

Definição 37

Uma v.a. contínua diz-se ter **distribuição uniforme** ou **rectangular** no intervalo (a, b) e representa-se por $X \sim \mathcal{U}(a, b)$ se a função densidade de probabilidade (f.d.p.) é da forma:

$$f(x) = \begin{cases} 1/(b-a) & a < x < b \\ 0 & x \leq a \text{ ou } x \geq b. \end{cases}$$

Valor médio, variância e função geradora de momentos

$$E[X] = \frac{a+b}{2}; \quad \text{Var}[X] = \frac{(b-a)^2}{12} \quad \text{e} \quad M_X(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}, \quad t \neq 0$$

A distribuição uniforme contínua

Caso particular:

Considere a distribuição $\mathcal{U}(0, 1)$

Exercício 10

Escreva a função densidade, a função distribuição cumulativa, valor médio, variância e função geradora de momentos.

A distribuição normal ou de Gauss

Surge século XVIII → ligada ao estudo dos erros de medições repetidas de uma mesma quantidade.

Papel fulcral nas Probabilidades e Estatística, porque:

- muitas variáveis biométricas têm uma distribuição muito próxima da normal;
- por vezes uma variável que não é normal pode ser transformada de um modo simples numa outra com distribuição normal;
- a parte central de muitos modelos não normais é por vezes razoavelmente bem aproximada por uma distribuição normal.

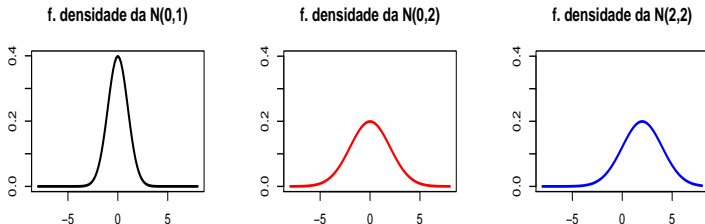
Definição 38

Uma v.a. contínua X diz-se ter **distribuição normal** ou **de Gauss** com parâmetros μ e σ e representa-se por $X \sim \mathcal{N}(\mu, \sigma)$ se a sua f.d.p. é da forma:

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right]$$

$$-\infty < x < +\infty, \quad -\infty < \mu < +\infty, \quad 0 < \sigma < +\infty$$

A distribuição normal ou de Gauss



Gráficos da função densidade normal para alguns valores de μ e σ .

Propriedades da curva densidade da variável com distribuição normal

1. É simétrica relativamente a μ .
2. É uma curva unimodal, a moda é μ .
3. Tem pontos de inflexão em $\mu + \sigma$ e $\mu - \sigma$.

A distribuição normal ou de Gauss

Valor médio, variância e função geradora de momentos

$$E[X] = \mu; \quad \text{Var}[X] = \sigma^2 \quad \text{e} \quad M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}} \quad \forall t \in \mathbb{R}$$

Definição 39

Se $\mu = 0$ e $\sigma = 1$ a variável aleatória com distribuição $\mathcal{N}(0, 1)$ chama-se **normal reduzida**.

A distribuição normal reduzida

Notações para a normal reduzida

$$Z \sim \mathcal{N}(0, 1); \quad \varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad \text{e} \quad \Phi(z) = P[Z \leq z]$$

Propriedade – consequência da simetria

$$\Phi(-z) = 1 - \Phi(z)$$

Tabelas \rightarrow dão o valor da função de distribuição cumulativa da normal reduzida.

A distribuição normal ou de Gauss

Alguns teoremas de grande importância no estudo da normal.

Teorema 9

Seja $X \sim \mathcal{N}(\mu, \sigma)$ a v.a. $Y = a + bX$ é também normal e tem-se $Y \sim \mathcal{N}(a + b\mu, |b|\sigma)$.

Corolário - muito importante

Seja $X \sim \mathcal{N}(\mu, \sigma)$, então a v.a. $Z = \frac{X - \mu}{\sigma}$ tem distribuição normal reduzida, i.e., $Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$.

Exercício 11

Uma vacaria tem uma produção diária de leite que se admite seguir uma lei normal com $\mu = 950$ l e $\sigma = 50$ l

- a) Qual a probabilidade de se ter uma produção inferior a 1000 litros?
- b) Qual a percentagem de dias em que a produção ultrapassa a produção média em mais de 100 litros?
- c) Se na região existe outra vacaria, com uma produção diária que se admite normal com $\mu = 900$ l e $\sigma = 40$ l, funcionando independentemente da primeira, qual a probabilidade de num dado dia a produção total das duas vacarias ser superior a 1800 litros?

A distribuição normal ou de Gauss

Para respondermos à alínea c) necessitamos do seguinte [Teorema](#)

Teorema 10

Sejam X_1, \dots, X_n , v.a. normais independentes, tais que $X_1 \sim \mathcal{N}(\mu_1, \sigma_1)$, $X_2 \sim \mathcal{N}(\mu_2, \sigma_2)$, \dots , $X_n \sim \mathcal{N}(\mu_n, \sigma_n)$.

A v.a. $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \dots + \mathbf{X}_n$ tem distribuição normal de parâmetros (μ, σ) , com

$$\mu = \mu_1 + \mu_2 + \dots + \mu_n \quad \text{e} \quad \sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}$$

A distribuição normal ou de Gauss

Teorema 11—Generalização do teorema anterior

Mostre que, sendo X_1, \dots, X_n v.a. nas condições do teorema 10, $a_1 X_1 + a_2 X_2 + \dots + a_n X_n$ tem distribuição normal de parâmetros (μ, σ) , com

$$\mu = a_1 \mu_1 + a_2 \mu_2 + \dots + a_n \mu_n \quad \text{e} \quad \sigma = \sqrt{a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \dots + a_n^2 \sigma_n^2}.$$

Corolário

Sejam X_i n v.a. normais independentes e semelhantes, i.e., tendo todas o mesmo valor médio μ e a mesma variância σ^2 .

As variáveis aleatórias **soma** e **média**, definidas respectivamente como

$$S_n = \sum_{i=1}^n X_i \quad \text{e} \quad \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

têm distribuição normal assim definida

$$S_n \sim \mathcal{N}(n\mu, \sigma\sqrt{n}) \quad \text{e} \quad \bar{X}_n \sim \mathcal{N}(\mu, \sigma/\sqrt{n}).$$

O Teorema Limite Central

Provámos que a soma de NORMAIS independentes é ainda uma normal. **Mas temos mais ...**

a **distribuição aproximada da SOMA de n variáveis aleatórias com QUALQUER lei**, mas independentes, identicamente distribuídas e verificando certas condições é também **normal**.

Teorema 12–Teorema limite central

Sejam X_1, \dots, X_n variáveis aleatórias independentes e identicamente distribuídas, com valor médio μ e variância σ^2 (finita).

A v.a. $S_n = \sum_{i=1}^n X_i$ verifica quando n é “grande”:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \sim \mathcal{N}(0, 1)$$

Aplicações do Teorema Limite Central

Note que também se tem $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$.

Teorema 13–Teorema de De Moivre

Seja X uma v.a. com **distribuição binomial** com valor médio $\mu = np$ e variância $\sigma^2 = npq$. Então quando $n \rightarrow \infty$,


$$\frac{X - np}{\sqrt{npq}} \sim \mathcal{N}(0, 1)$$

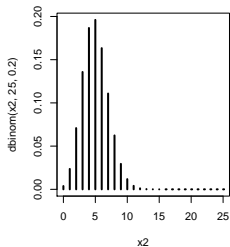
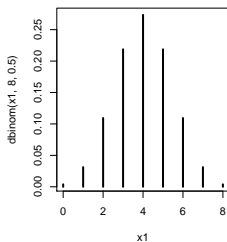
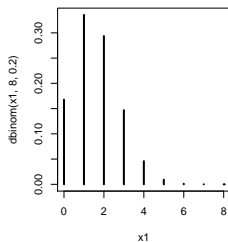
Aplicações do Teorema Limite Central

Recorde-se que se, na distribuição binomial, n grande e $p \approx 0$ (ou 1) uma boa aproximação é dada pela distribuição de Poisson.

E agora para valores de $p \approx 1/2$ o teorema limite central oferece muito boa aproximação para a normal.

Aplicações do Teorema Limite Central - Exercício 12

Utilizando o , obtenha os seguintes gráficos da função massa de probabilidade de $X \sim B(8, 0.2)$, $X \sim B(8, 0.5)$ e $X \sim B(25, 0.2)$.



O que observa?

Regra prática

Se na **distribuição binomial** $np > 5$ e $nq > 5 \implies$ a **aproximação pela distribuição normal é boa.**

Teorema 14

Seja $X \sim \mathcal{P}(\lambda)$. Quando $\lambda \rightarrow \infty$ então $\frac{X - \lambda}{\sqrt{\lambda}} \sim \mathcal{N}(0, 1)$.

Regra prática:

A aproximação é considerada boa para $\lambda \geq 20$.

Observação: Quando considerámos a aproximação da distribuição binomial pela Poisson, ambas eram distribuições discretas. Os dois teoremas acabados de enunciar dão-nos uma aproximação de uma v.a. discreta por uma v.a. contínua.

Neste caso é necessário fazer-se o que se designa por **correção de continuidade** que consiste em considerar todo o inteiro k representado pelo intervalo $(k - 1/2, k + 1/2)$.

A distribuição exponencial

Uma variável aleatória diz-se ter **distribuição exponencial** de parâmetro β e representa-se por $X \sim \text{Exp}(\beta)$ se a função densidade é

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} & x > 0, \beta > 0 \\ 0 & x \leq 0 \end{cases}$$

Valor médio, variância e função geradora de momentos

$$M_X(t) = \frac{1}{1-\beta t}, \quad (t \leq 1/\beta); \quad E[X] = \beta; \quad \text{Var}[X] = \beta^2$$

Aplicações:

Duração de vida, teoria da fiabilidade, tempos de espera, etc.

A distribuição exponencial: observações

Propriedade

A distribuição exponencial goza da propriedade da **falta de memória**, que podemos formalizar como: Se $X \sim \text{Exp}(\beta)$ tem-se $P[X > t + h | X > t] = P[X > h]$, para $t > 0, h > 0$

Exercício

Mostre que se verifica a propriedade referida acima.

Uma interpretação da propriedade da **falta de memória**

Se um objecto tem tempo de vida com distribuição exponencial, então, qualquer que seja a sua idade o tempo de vida que resta não é afectado pelo tempo já vivido.

A distribuição exponencial: observações

Relação entre a distribuição exponencial e a distribuição de Poisson:

Considere-se contagens de sucessos em intervalos de tempo. O tempo ao fim do qual se verifica o primeiro sucesso é uma variável aleatória contínua.

Teorema 15

Se X , número de sucessos num intervalo de tempo, é tal que $X \sim \mathcal{P}(\lambda)$ então W a v.a. que designa o tempo de espera pelo primeiro sucesso (ou o tempo entre a ocorrência de dois sucessos consecutivos) satisfaz

$$W \sim \text{Exp}(\beta = 1/\lambda).$$