Departamento de Matemática

Introdução à Probabilidade e Estatística
2015/2016 - 2º Semestre

Capítulo 1

1:

(a) Vamos construir a tabela de frequências utilizando para tal a **regra de Sturges**

$$k = \lceil 1 + log_2(n) \rceil = \lceil 6.39 \rceil = 7$$

onde k é o número de classes e n a dimensão amostral. Tomando 7 classes, de igual amplitude, podemos organizar os dados na seguinte tabela de frequências:

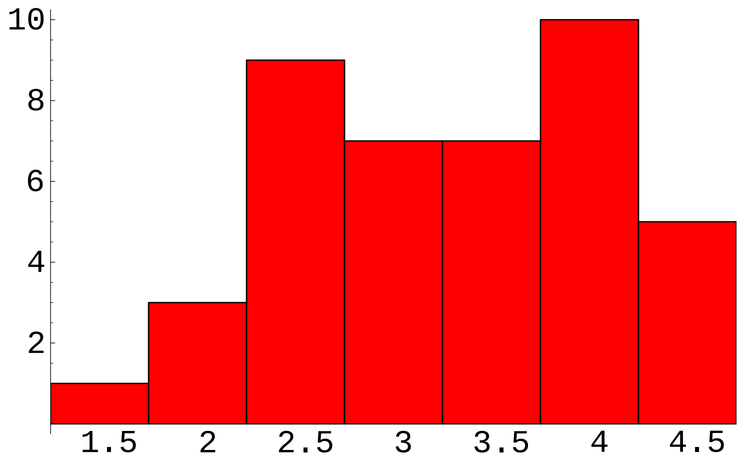
j	C_j	n_j	f_j	N_j	F_j	c_j
1	$[1.2, 1.7[$	1	$\frac{1}{42}$	1	$\frac{1}{42}$	1.45
2	$[1.7, 2.2[$	3	$\frac{3}{42}$	4	$\frac{4}{42}$	1.95
3	$[2.2, 2.7[$	9	$\frac{9}{42}$	13	$\frac{13}{42}$	2.45
4	$[2.7, 3.2[$	7	$\frac{7}{42}$	20	$\frac{20}{42}$	2.95
5	$[3.2, 3.7[$	7	$\frac{7}{42}$	27	$\frac{27}{42}$	3.45
6	$[3.7, 4.2[$	10	$\frac{10}{42}$	37	$\frac{37}{42}$	3.95
7	$[4.2, 4.7]$	5	$\frac{5}{42}$	42	1	4.45

Legenda

- C_j : j-ésima classe
- n_j : frequência absoluta da classe j
- f_j : frequência relativa da classe j
- N_j : frequência absoluta acumulada da classe j
- F_j : frequência relativa acumulada da classe j
- c_j : ponto médio da classe j

Nota: nesta resolução optou-se por se considerar 7 classes; outras soluções igualmente correctas foram propostas nas aulas práticas pelos respectivos docentes.

(b) O histograma construído com base na tabela de frequências anteriormente apresentada, sendo que em ordenadas figuram as frequências absolutas é da seguinte forma (se figurassem as frequências relativas ou as frequências relativas por unidade de amplitude das classes o gráfico manteria a sua forma gráfica, mudando apenas a escala):



(em abcissa figuram os pontos médios das classes)

(c) Classe modal $C_6 = [3.7, 4.2[$
Classe mediana $C_5 = [3.2, 3.7[$

(d) Média e desvio padrão dos dados agrupados e não agrupados

	$\bar{x}$	s
dados não agrupados	3.16	0.886
dados agrupados	3.235	0.813

Nota: enquanto que para dados não agrupados

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}{n-1}$$

para dados agrupados

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{k} c_j n_j = \sum_{j=1}^{k} \bar{x}_j f_j, \qquad s^2 = \frac{1}{n-1}\sum_{j=1}^{k}(c_j - \bar{x})^2 n_j.$$

Nota: quando no cálculo da variância em vez de se dividir por $n-1$ divide-se por n, temos a *variância corrigida*, que por vezes se denota por s_n^2.

2:

(a) Tome-se $\bar{x}_i$ e $s^{(i)}$ como sendo a média e o desvio padrão da turma i $(i = 1, 2, 3)$. Decorre do enunciado que $\bar{x}_1 = 13$, $\bar{x}_2 = 10$, $\bar{x}_3 = 9$, $s^{(1)} = 2$, $s^{(2)} = 2.2$, $s^{(3)} = 2.1$. Usando a ponderação adequada (uma vez que as amostras não têm igual dimensão), vem que

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3}{n_1 + n_2 + n_3} = 10.476.$$

Por outro lado, para determinar

$$s^2 = \frac{\sum_{i=1}^{30+35+40-1} x_i^2 - \bar{x}^2}{30+35+40-1}$$

temos de determinar as somas do tipo $\sum_i x_i^2$. Note-se que

$$\begin{cases} (s^{(1)})^2 = 4 \Longrightarrow \sum_{i=1}^{30} x_i^2 = 5186 \\ (s^{(2)})^2 = 2.2^2 \Longrightarrow \sum_{i=1}^{35} x_i^2 = 3664.56 \\ (s^{(3)})^2 = 2.1^2 \Longrightarrow \sum_{i=1}^{40} x_i^2 = 3411.99 \end{cases}$$

$$\Longrightarrow s^2 = \frac{5186 + 3664.56 + 3411.99 - 105 \times (10.476)^2}{104} = 7.11$$

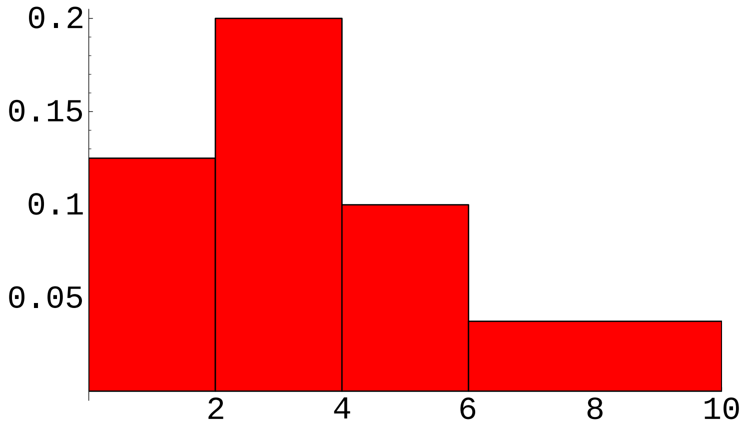
Em particular segue-se que $s = 2.665$.

(b) Seja $x^\star$ a nota do aluno na nova escala, sendo que na escala original a nota, x, é igual a 10 valores. Dado que o professor decide alterar linearmente a nota dos alunos, então a relação entre $x^\star$ e x é dada por $x^\star = ax + b = 10a + b$, onde a e b são constantes tais que $\bar{x}^\star = 12$ e $(s^\star)^2 = 2^2$ (respectivamente, média e variância das notas na nova escala). Dado que

$$\bar{x}^\star = a\bar{x} + b \quad \wedge \quad (s^\star)^2 = a^2 s^2 \Rightarrow a = 0.75, b = 4.143 \Rightarrow x^\star = 11.643$$

1.3:

(a) Neste caso em ordenadas (no histograma) figuram as frequências relativas por unidade de amplitude das classes.



(b) Dado que se trata de dados agrupados, teremos de utilizar as fórmulas correspondentes.

$$\bar{x} = \sum_{k=1}^{3} f_i c_i = 0.25 \times 1 + 0.40 \times 3 + 0.20 \times 5 + 0.15 \times 8 = 3.65$$

$$\begin{aligned} s^2 &= \frac{1}{119}(\sum_{k=1}^{3} n_k c_k^2 - 120 \times (3.65)^2) \\ &= \frac{1}{119}(0.25 \times 120 \times 1^2 + 0.40 \times 120 \times 3^2 + 0.20 \times 120 \times 5^2 + 0.15 \times 120 \times 8^2 - 120(3.65)^2) \\ &= 5.17 \end{aligned}$$

Em particular segue-se que $s = 2.27$.

(c) $me = 3.25$

(d) Sejam $\bar{x}^\star$ e $(s^\star)^2$ a média e a variância, respectivamente, dos salários após aumento. Se for dado um aumento de 100%: $\bar{x}^\star = 2\bar{x}$, $\quad (s^\star)^2 = 2^2 s^2$ (note-se que neste caso $\bar{x}_i^\star = 2\bar{x}_i$, onde $\bar{x}_i^\star$ é o ponto médio da classe k após aumento de 100%).

(e) Se for dado um aumento de 2 unidades: $\bar{x}^\star = \bar{x} + 2$, $\quad (s^\star)^2 = s^2$ (note-se que neste caso $\bar{x}_i^\star = \bar{x}_i + 2$).

Capítulo 2

1: Considere-se a seguinte notaçãao (para designar os acontecimentos relevantes ao problema):

$$S = \{\text{programa seleccionado possui erro(s) de sintaxe}\}$$
$$I = \{\text{programa seleccionado possui erro(s) de input/output}\}$$
$$O = \{\text{programa seleccionado possui erro(s) de outro tipo}\}$$

Do enunciado do exercício resulta que: $P(S) = 0.20$, $P(I) = 0.10$, $P(O) = 0.05$, $P(S \cap I) = 0.06$, $P(S \cap O) = P(I \cap O) = 0.03$ e $P(S \cap I \cap O) = 0.02$.

(a) Nesta alínea pretende-se determinar $P(S \cap \bar{I} \cap \bar{O})$, donde

$$P(S \cap \bar{I} \cap \bar{O}) = P(S) - P(S \cap I) - P(S \cap O) + P(S \cap I \cap O) = 0.13.$$

(sugestão: faça um diagrama de Venn)

(b) Dado que a probabilidade pedida é $P(S \cup I \cup O)$, então

$$\begin{aligned} P(S \cup I \cup O) &= P(S) + P(I) + P(O) - P(S \cap I) - P(S \cap O) - P(I \cap O) + P(S \cap I \cap O) \\ &= 0.20 + 0.10 + 0.05 - 0.06 - 0.03 - 0.03 + 0.02 = 0.25. \end{aligned}$$

(fórmula da adição para 3 acontecimentos)

2: Considera-se que se regista o estado de cada um dos dois transístores amostra-dos pela ordem com que estes são retirados.

(a) O espaço de resultados é $= \{(D, D), (F, D), (F, D), (F, F)\}$, onde D (F) significa defeituoso (não defeituoso ou funcional) e: $P(\{(D, D)\}) = 0.04$, $P(\{(F, D)\}) = P(\{(D, F)\}) = 0.16$ e $P(\{(F, F)\}) = 0.64$.

(b) $P(A_1) = P(A_2) = 0.2$, $P(A_3) = 0.36$ e $P(A_4) = 0.32$.

(c) Considerando que não há reposição, o espaço de resultados mantém-se, mas:

$$P(\{(D, D)\}) = \frac{2}{90}, \quad P(\{(F, D)\}) = P(\{(D, F)\}) = \frac{16}{90} \quad \text{e} \quad P(\{(F, F)\}) = \frac{56}{90};$$

$$P(A_1) = P(A_2) = 0.2, \quad P(A_3) = \frac{17}{45} \quad \text{e} \quad P(A_4) = \frac{16}{45}.$$

3: Suponha-se que o teste é aplicado a um indivíduo escolhido ao acaso da população.

Dados: $P(C) = 0.005$, $P(T|C) = 0.99$, $P(\bar{T}|\bar{C}) = 0.95$, onde $C = \{$o indivíduo testado tem cancro$\}$ e $T = \{$o resultado do teste aplicado ao indivíduo é positivo$\}$.

(a) Nesta alínea pretende-se determinar $P(C|T)$. Note-se que decorre do teorema de Bayes que

$$\begin{aligned} P(C|T) &= \frac{P(T|C)P(C)}{P(T)} \\ &= \frac{P(T|C)P(C)}{P(T|C)P(C) + P(T|\bar{C})P(\bar{C})} \\ &= \frac{0.99 \times 0.005}{0.99 \times 0.005 + (1 - 0.95) \times (1 - 0.005)} \\ &= 0.0905. \end{aligned}$$

(b) Seja $D = \{$os resultados de dois testes aplicados sucessivamente ao indivíduo são ambos positivos$\}$. Note-se que se T_i designar o acontecimento

$$T_i = \{i - \text{ésimo teste dá resultado positivo}\}$$

então segue-se que $D = T_1 \cap T_2$. Dos dados do problema, e uma vez que há independência condicional dos testes dado o estado do indívíduo, vem que

$$P(D|C) = P(T_1 \cap T_2|C) = P(T_1|C)P(T_2|C) = P(T|C)^2$$

e

$$P(D|\bar{C}) = P(T_1 \cap T_2|\bar{C}) = P(T_1|\bar{C})P(T_2|\bar{C}) = \left(P(T|\bar{C})\right)^2.$$

Aplicando um raciocínio idêntico ao da alínea anterior, vem que a probabilidade pedida é $P(C|D) = 0.6633$.

Note que no enunciado não é dito que os resultados dos testes são independentes pelo que

$$P(T_1 \cap T_2) \quad \textbf{não é necessariamente igual a } P(T_1)P(T_2).$$

Departamento de Matemática

Introdução à Probabilidade e Estatística
2015/2016 - 2º Semestre

Capítulo 4

1: (a) $F_X(x) = P(X \leq x) = \begin{cases} 0 & x < 0 \\ 5x^4 - 4x^5 & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}.$

(b) $P(X < \frac{2}{3}) = P(X \leq \frac{2}{3}) = 0.462.$

(c) Seja Y a variável aleatória que designa o lucro por litro. Então

i) $P(Y = C_1 - C_3) = P(1/3 < X < 2/3) = 101/243$ e $P(Y = C_2 - C_3) = 142/243$. Assim, supondo $C_2 < C_1$:

$$F_Y(y) = \begin{cases} 0 & y < C_2 - C_3 \\ \frac{142}{243} & C_2 - C_3 \leq y < C_1 - C_3 \\ 1 & y \geq C_1 - C_3 \end{cases}.$$

ii) $E[Y] = \frac{101}{243}C_1 + \frac{142}{243}C_2 - C_3$.

2: Sejam $T_B = \{$a peça escolhida é do tipo $B\}$, $T_A = \bar{T}_B$ e $X =$ duração da peça escolhida.

$$P(T_B|X < 0.9) = \frac{P(X < 0.9|T_B)P(T_B)}{P(X < 0.9|T_A)P(T_A) + P(X < 0.9|T_B)P(T_B)} = 0.4502.$$

3: (a) Se $P(X < 2.5) = 0.5$ então $E[X] = \mu_X = 2.5$. Logo $P(X < 3.42) = 0.975 \Leftrightarrow \Phi(\frac{3.42-2.5}{\sigma_X}) = 0.975 = \ ^{-1}(1.96)$ pelo que $\sigma_X = 0.4694$.

(b) $P(|X - \mu_X| < \sigma_X) = \Phi(1) - \Phi(-1) = 0.6926$.

Departamento de Matemática

Introdução à Probabilidade e Estatística
2015/2016 - 2º Semestre

Capítulo 5

1: (a) $P(X=x)=\begin{cases}0.2 & x=0\\ 0.65 & x=1\\ 0.15 & x=2\\ 0 & c.c.\end{cases}$ e $P(Y=y)=\begin{cases}0.5 & y=0\\ 0.36 & y=1\\ 0.14 & y=2\\ 0 & c.c.\end{cases}$.

(b) $F_X(x)=\begin{cases}0 & x<0\\ 0.2 & 0\le x<1\\ 0.85 & 1\le x<2\\ 1 & x\ge 2\end{cases}$.

(c) $P(Y>X)=0.18$.

(d) $\mathrm{E}[X+Y]=1.59$, $\mathrm{Var}[X+Y]=0.8619$.

2:

(a) i) $P(X=x)=\begin{cases}2/9 & x=1\\ 1/2 & x=2\\ 5/18 & x=3\\ 0 & c.c.\end{cases}$.

ii) $F_Y(y)=\begin{cases}0 & y<1\\ 1/6 & 1\le y<2\\ 11/18 & 2\le y<3\\ 1 & y\ge 3\end{cases}$.

iii) $P(X+Y\le 4)=11/18$.

iv) $P(X=x|Y=1)=\begin{cases}2/3 & x=1\\ 1/3 & x=3\\ 0 & c.c.\end{cases}$ e $P(X=x|Y=3)=\begin{cases}2/7 & x=1\\ 3/7 & x=2\\ 2/7 & x=3\\ 0 & c.c.\end{cases}$.

v) $E[X|Y=1]=5/3$.

(b) $E[X|Y]$ assume valores no conjunto $\{5/3, 9/4, 2\}$ e possui função de probabilidade

$$P(E[X|Y]=z)=\begin{cases}0.16 & z=\frac{5}{3}\\ 0.48 & z=\frac{9}{4}\\ 0.36 & z=2\\ 0 & c.c.\end{cases}.$$

(c) i) $Var[X|Y=1]=8/9$.

ii) $P(X.Y \text{ ser par })=11/18$.

iii) $P(Y=2|X.Y\leq 4)=6/11$.

iv) $F_{Y|X=3}(y)=\begin{cases}0 & y<1\\ 1/5 & 1\leq y<2\\ 3/5 & 2\leq y<3\\ 1 & y\geq 1\end{cases}.$

(d) Não são independentes pois, por exemplo, $P(X=1,Y=2)\neq P(X=1)P(Y=2)$.

3:

(a) i)

$X\backslash Y$	0	1	2	3	4	$P(X=x)$
0	0.04	0	0.3	0	0.06	0.4
1	0	0.22	0	0.28	0	0.5
2	0	0	0.1	0	0	0.1
$P(Y=y)$	0.04	0.22	0.4	0.28	0.06	1

ii) $P(X=x)=\begin{cases}0.4 & x=0\\ 0.5 & x=1\\ 0.1 & x=2\\ 0 & c.c.\end{cases}$ e $P(Y=y)=\begin{cases}0.04 & y=0\\ 0.22 & y=1\\ 0.4 & y=2\\ 0.28 & y=3\\ 0.06 & y=4\\ 0 & c.c.\end{cases}.$

iii) $F_X(x)=\begin{cases}0 & x<0\\ 0.4 & 0\leq x<1\\ 0.9 & 1\leq x<2\\ 1 & x\geq 2\end{cases}.$

iv) $P(X=x|Y=2)=\begin{cases}0.75 & x=0\\ 0.25 & x=2\\ 0 & c.c.\end{cases}.$

v) Não são independentes pois, por exemplo, $P(X=0, Y=1) \neq P(X=0)P(Y=1)$.

(b) i) $P(Y=y|X=0) = \begin{cases} 0.1 & y=0 \\ 0.75 & y=2 \\ 0.15 & y=4 \\ 0 & \text{c.c.} \end{cases}$, $\quad P(Y=y|X=1) = \begin{cases} 0.44 & y=1 \\ 0.56 & y=3 \\ 0 & c.c. \end{cases}$

e $\quad P(Y=y|X=2) = \begin{cases} 1 & y=2 \\ 0 & c.c. \end{cases}$.

ii) $E[Y|X=2]=2, \quad \text{Var}[Y|X=2]=0$.

iii) $F_{Y|X=0}(y) = \begin{cases} 0 & y<0 \\ 0.1 & 0 \leq y < 2 \\ 0.85 & 2 \leq y < 4 \\ 1 & y \geq 4 \end{cases}$.

iv) $P(Y=2|X.Y=0) = 0.75$.

v) $P(X+Y \text{ ser ímpar}) = 0$.

4: $P(X=0, Y=0) \neq P(X=0)P(Y=0)$. Logo X e Y não são independentes.