

# CE-003: Estatística II - Turma: K/O, Prova Final (18/12/2013)

GRR: \_\_\_\_\_ Nome: \_\_\_\_\_ Turma: \_\_\_\_\_

1. A taxa de mortes por afogamento em finais de semana numa cidade praiana é de 2,5 para cada 100.000 habitantes. Faça as suposições necessárias e encontre a probabilidade de que ocorram quatro ou mais afogamentos em um determinado final de semana para o qual estima-se uma população de 350.000 habitantes.

**Solução:**

$X$  : número de afogamentos para 350.000 habitantes

$$X_B = B(n = 350.000, p = 2,5/100.000)$$

$$X_P \approx P(\lambda = n \cdot p = 8,75)$$

$$P[X \geq 4] = 1 - P[X \leq 3] = 0,975$$

2. Uma amiga sua acredita que ela tem 50% de chance de estar grávida, Ela decide fazer um teste de gravidez que resulta positivo. Enquanto isto voce lê informações sobre o teste que informam que 20% de mulheres não grávidas apresentam resultados positivos quando fazem o teste (falso positivos) e 10% das mulheres grávidas apresentam resultados negativos no teste. Com estas informações ajude sua amiga a reavaliar a chance de estar grávida.

**Solução:**

Eventos:  $G$ : estar grávida e  $P$ : teste positivo

Informações fornecidas:  $P[G] = 0,50$ ,  $P[P|\bar{G}] = 0,20$  e  $P[\bar{P}|G] = 0,10$  Solicitado:

$$P[G|P] = \frac{P[P \cap G]}{P[P]} = \frac{P[G \cap P]}{P[P \cap G] + P[P \cap \bar{G}]} = \frac{P[P|G] \cdot P[G]}{P[P|G] \cdot P[G] + P[P|\bar{G}] \cdot P[\bar{G}]} = \frac{0,90 \cdot 0,50}{0,90 \cdot 0,50 + 0,20 \cdot 0,50} = 0,818$$

Na forma de uma tabela:

	Positivo	Negativo	Total
Grávida	0,45	0,05	0,50
Não Grávida	0,10	0,40	0,50
Total	0,55	0,45	1,00

3. Os escores obtidos em um exame de proficiência se distribuem segundo a distribuição normal com média 400 e desvio padrão 45.
- qual a percentagem de pessoas com escores acima de 350?
  - qual a percentagem de pessoas com escores entre 450 e 500?
  - qual a percentagem de pessoas com escores que não se afastem da média mais do que 30?
  - qual valor deve ter 80% dos escores abaixo dele?
  - mantendo-se a mesma média, quanto deveria ser o desvio padrão ter 10% dos escores acima de 500?

**Solução:**

$X$  : escores no exame e proficiência

$$X \sim N(400, 45^2)$$

(a)  $P[X > 350] = P[Z > \text{Sexpr format}((350 - 400)/45, dig = 3)] = 0,8667 \rightarrow 86,67\%$

(b)  $P[450 < X < 500] = P[\text{Sexpr format}((450 - 400)/45, dig = 3) < Z < \text{Sexpr format}((500 - 400)/45, dig = 3)] = 0,1201 \rightarrow 12,01\%$

- (c)  $P[|X - 400| < 30] = P[370 < X < 430] = P[\text{Sexprformat}((370 - 400)/45, \text{dig} = 3) < Z < \text{Sexprformat}((430 - 400)/45, \text{dig} = 3)] = 0,495 \rightarrow 49,5\%$
- (d)  $P[X < a] = 0,8 \rightarrow P[Z < 0,842] \rightarrow a - 400/45 = 0,842 \rightarrow a = 437,9$
- (e)  $P[X > 500 | \mu = 400, \sigma] = 0,1 \rightarrow P[Z < 1,28] \rightarrow 500 - 400/\sigma = 1,28 \rightarrow \sigma = 78$

4. Os números abaixo mostram as notas de um grupo de alunos em duas avaliações

Aluno	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Prova 1	35	39	50	47	33	17	17	80	23	51	2	21	20	12	81	98	47	34
Prova 2	65	63	80	72	65	35	62	72	50	60	32	59	40	68	79	85	80	55

- (a) Calcule média, variância e coeficiente de variação das notas em cada avaliação
- (b) Calcule mediana, quantis, amplitude e amplitude interquartilica de cada avaliação
- (c) Faça um diagrama *box-plot* para comparar as notas das duas avaliações
- (d) Com as notas das duas provas juntas faça um único diagrama ramo-e-folhas sublinhando as notas da segunda prova.
- (e) Usando as medidas e gráficos acima compara o rendimento dos alunos nas duas provas.
- (f) Existe relação (associação) entre os resultados das duas provas? Faça um gráfico e calcule alguma(s) medida(s) estatística(s) para verificar se há associação.
- (g) Suponha agora que as provas possam ser consideradas uma amostra aleatória. Faça um testes estatísticos adequado para verificar:
- Se a média da Prova 1 está significativamente abaixo de 50.
  - Se houve aumento (significativo) de rendimento da primeira para segunda prova.

**Solução:**

(a)

$$\bar{x}_1 = 39,28 \quad s_1^2 = 670,68 \quad CV_1 = 65,93\%$$

$$\bar{x}_2 = 62,33 \quad s_2^2 = 237,53 \quad CV_2 = 24,73\%$$

(b)

$$md_1 = 34,5 \quad Q1_1 = 20 \quad Q3_1 = 50 \quad A_1 = 96 \quad AI_1 = 30$$

$$md_2 = 64 \quad Q1_2 = 55 \quad Q3_2 = 72 \quad A_2 = 53 \quad AI_2 = 17$$

(c)

(d) The decimal point is 1 digit(s) to the right of the |

```

0 | 2
1 | 277
2 | 013
3 | 234559
4 | 077
5 | 00159
6 | 023558
7 | 229
8 | 00015
9 | 8

```

- (e) Comentários sobre: valores centrais, variabilidade, assimetria e dados discrepantes
- (f) Coeficientes de correlação: Pearson  $r_P = 0,75$  e Spearman  $r_S = 0,732$   
Comentários: ...
- (g) • Teste de Hipótese para uma média

$$X_1 : \text{nota na Prova 1} \sim \text{Dist}(\mu_1, \sigma_1^2)$$

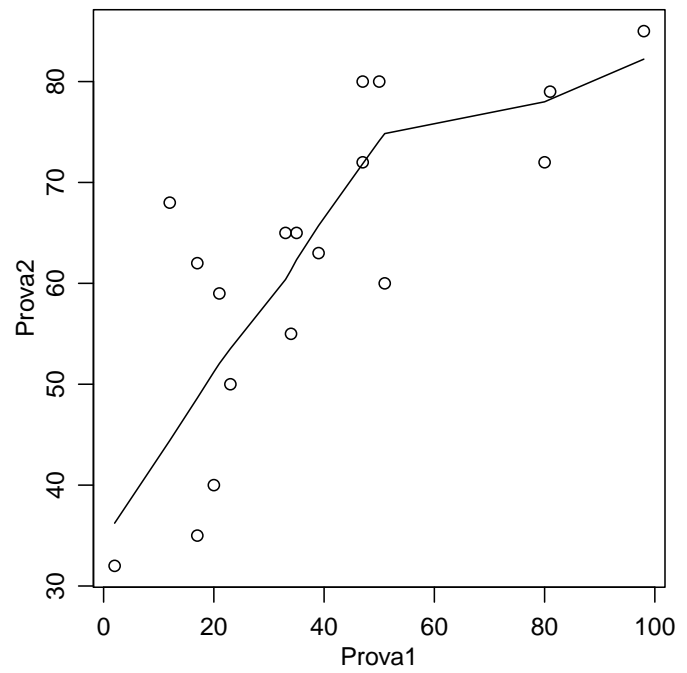
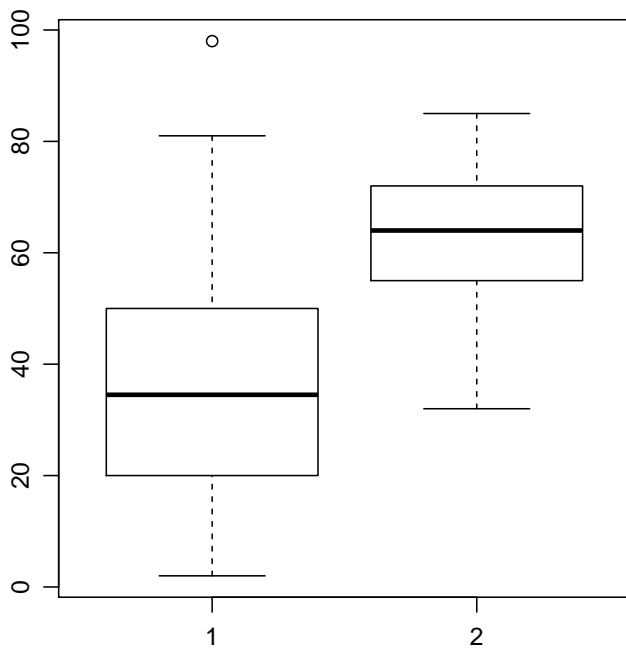
$$H_0 : \mu_1 \geq 50 \text{ vs } H_1 : \mu_1 < 50$$

$$\alpha = 0,05 \rightarrow t_t = -1,74 \rightarrow \hat{\mu}_1 = \bar{x}_1 < 39,381$$

$$t_c = \frac{39,3 - 50}{25,9/\sqrt{18}} = -1,757$$

$$p - \text{valor} = 0,04849$$

Conclusão: Rejeita-se  $H_0$  para  $\alpha = 0,05$  e portanto há evidências que a média esteja abaixo de 50.



- Teste de Hipótese para comparação de médias pareadas

$$X_1 : \text{nota na Prova 1} \sim \text{Dist}(\mu_1, \sigma_1^2)$$

$$X_2 : \text{nota na Prova 2} \sim \text{Dist}(\mu_2, \sigma_2^2)$$

$$D = X_2 - X_1 : \text{diferença de notas} \sim \text{Dist}(\mu_D = \mu_2 - \mu_1, \sigma_D^2)$$

$$H_0 : \mu_2 \leq \mu_1 \rightarrow \mu_D \leq 0 \text{ vs } H_1 : \mu_2 > \mu_1 \rightarrow \mu_D > 0$$

$$\alpha = 0,05 \rightarrow t_t = 1,74 \rightarrow \hat{\mu}_D = \bar{x}_D < 7,214$$

Amostra :

$$\hat{\mu}_D = \bar{x}_D = x_2 - x_1 = 23,056 \quad \hat{s}_D = 17,595$$

$$t_c = \frac{23,1 - 0}{17,6/\sqrt{18}} = 5,559$$

$$p\text{-valor} = 0,99998$$

Conclusão: Rejeita-se  $H_0$  para  $\alpha = 0,05$  e portanto há evidências que houve aumento de rendimento.