FICHA 11 – resolução

DPC

1. O nível de ácido úrico X (mg/dl) numa determinada população de homens adultos e saudáveis tem uma distribuição normal de valor esperado μ mg/dl e desvio padrão $\sigma = 1$mg/dl.

   (a) Recolheu-se uma amostra de dimensão 100 cujos níveis de ácido úrico apresentaram valor médio igual a 5.5 mg/dl. Determine um intervalo de confiança para μ a 99%.

   (b) Pretendendo-se testar a hipótese $\mu = 5.5$mg/dl recolheu-se outra amostra de dimensão 100 que forneceu um valor médio igual a 6 mg/dl para o nível de ácido úrico. Classifique, quanto à veracidade, a afirmação seguinte: "A 95% (nível de significância 5%), esta amostra leva-nos a rejeitar a hipótese inicial".

   (c) Suponha agora que $X \sim N(5.5; 1)$. Calcule a dimensão da amostra de modo a que seja pelo menos igual a 0.9 a probabilidade de a média da amostra se situar entre 5 e 6.

   Resolução:

   (a) $z_{0.005} = 2.576$, $\mu \in [5.24,\ 5.76]$.

   (b) Um teste de hipóteses inclui sempre duas hipóteses mutuamente exclusivas (isto é, se uma é verdadeira, a outra é necessariamente falsa), chamadas hipótese nula (H_0) e hipótese alternativa (H_1). A ideia a seguir no teste resume-se no seguinte. Com base na amostra obtida, calcula-se uma estatística apropriada, cuja distribuição seja conhecida (ex. a média e/ou a variância da amostra). De seguida, começa-se por assumir que a hipótese nula é verdadeira. Assumindo H_0 verdadeira, calcula-se quão provável seria obter a estatística observada ou outra ainda mais extrema. Se essa probabilidade for "pequena", então o pressuposto de que H_0 era verdadeira é provavelmente falso, logo rejeita-se H_0 e aceita-se H_1. Caso contrário, não há evidência suficiente para rejeitar H_0, pelo que não se aceita H_1. Neste exercício, temos um teste bilateral para a média μ de uma distribuição $N(\mu, \sigma^2)$ (com σ conhecido):

   $$\begin{aligned} &H_0 : \mu = \mu_0 \quad \text{(hipótese nula)} \\ &H_1 : \mu \neq \mu_0 \quad \text{(hipótese alternativa)} \end{aligned}$$

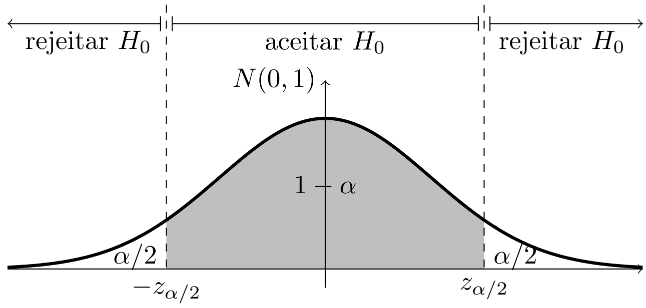
   A estatística a utilizar para este tipo de teste é a média amostral, $\bar{X}$, e, como já vimos, $\bar{X} \sim N(\mu, \sigma^2/n)$. Então, se assumirmos que H_0 é verdadeira, temos $\bar{X} \sim N(\mu_0, \sigma^2/n)$. Intuitivamente, se a média amostral observada ($\bar{x}$) estiver muito distante da média que H_0 assume como verdadeira (μ_0), então provavelmente H_0 será falsa e, por isso, rejeitá-la-emos. Concretamente, rejeitamos H_0 (e aceitamos H_1) se:

   $$P\left(|\bar{X} - \mu| \geq |\bar{x} - \mu| \,\middle|\, \mu = \mu_0\right) < \alpha,$$

   onde α é um valor dado, chamado *nível de significância* do teste. Os valores $\alpha = 5\%$ ou $\alpha = 1\%$ são os mais utilizados na prática. A probabilidade $P(|\bar{X} - \mu| \geq |\bar{x} - \mu| \mid \mu = \mu_0)$ é chamada de *valor-p*. Fazendo a padronização da v.a. ($Z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n}) \sim N(0, 1)$), a condição para rejeitar H_0 fica:

   $$\begin{aligned} &P\left(|Z| \geq \frac{|\bar{x} - \mu_0|}{\sigma/\sqrt{n}}\right) < \alpha \\ \Leftrightarrow\ &P\bigg(Z \geq \underbrace{\frac{|\bar{x} - \mu_0|}{\sigma/\sqrt{n}}}_{z_0}\bigg) + P\bigg(Z \leq \underbrace{-\frac{|\bar{x} - \mu_0|}{\sigma/\sqrt{n}}}_{-z_0}\bigg) < \alpha \\ \Leftrightarrow\ &2P(Z \geq z_0) < \alpha. \end{aligned}$$

Podemos representar este critério graficamente:



Pelo gráfico, torna-se óbvio que um critério equivalente para rejeitar H_0 é $z_0 \notin [-z_{\alpha/2},\ z_{\alpha/2}]$. Podemos ainda reescrevê-lo em função da média amostral observada, $\bar{x}$, obtendo o seguinte critério:

$$\bar{x} \notin \underbrace{\left[\mu_0 \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]}_{\text{região de aceitação}} \Leftrightarrow \text{rejeitar } H_0.$$

No exemplo dado, temos $H_0 : \mu = 5.5\text{mg/dl}$, $H_1 : \mu \neq 5.5\text{mg/dl}$, $\alpha = 0.05$, $\sigma = 1\text{mg/dl}$, $n = 100$ e $\bar{x} = 6\text{mg/dl}$. Assim, temos $z_0 = \frac{|6-5.5|}{1/\sqrt{100}} = 5.0$ e $z_{\alpha/2} = z_{0.025} = 1.960$, logo $z_0 > z_{\alpha/2}$, pelo que rejeitamos H_0. Concluímos, assim, que a afirmação dada no enunciado é verdadeira. Pelo cálculo do valor-p, chegaríamos necessariamente à mesma conclusão:

$$\text{valor-}p = P\left(|\bar{X}-\mu| \geq |\bar{x}-\mu| \,\middle|\, \mu = 5.5\right) = 2P(Z \geq z_0) = 2P(Z \geq 5.0) \approx 0 < 0.05 = \alpha,$$

logo rejeitamos H_0. Finalmente, a região de aceitação para este teste seria $\left[5.5 \pm 1.960\frac{1}{\sqrt{100}}\right] = [5.3,\ 5.7]$ (mg/dl). Como $\bar{x} = 6\text{mg/dl}$ não pertence a este intervalo, mais uma vez concluímos que devemos rejeitar H_0.

(c) $\mu = 5.5$, $\sigma = 1$,

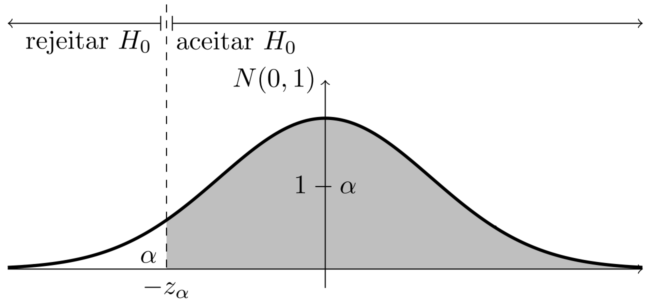
$$\begin{aligned} n: \quad & P(5 \leq \bar{X} \leq 6) \geq 0.9 \\ & \Leftrightarrow P\left(\frac{5-5.5}{1/\sqrt{n}} \leq Z \leq \frac{6-5.5}{1/\sqrt{n}}\right) \geq 0.9 \Leftrightarrow P\left(-0.5\sqrt{n} \leq Z \leq 0.5\sqrt{n}\right) \geq 0.9 \\ & \underset{\text{simetria}}{\Leftrightarrow} P\left(Z \geq 0.5\sqrt{n}\right) \leq 0.05 \Leftrightarrow 0.5\sqrt{n} \geq 1.645 \Leftrightarrow n \geq 11. \end{aligned}$$

2. A partir de registros extensos, sabe-se que a duração do tratamento de uma doença por uma terapia tem uma média de 15 dias, com um desvio padrão de 3 dias. Alega-se que uma nova terapia pode reduzir o tempo de tratamento. Para testar essa afirmação, a nova terapia é testada em 70 pacientes, e o seu tempo de tratamento médio é 14.6 dias. Os resultados suportam a afirmação?

Resolução:
Sabe-se que $X \sim N(\mu, 3^2)$. Note que, neste caso, pretende-se averiguar se a nova terapia reduz o tempo de tratamento médio (que era de 15 dias na terapia anterior), pelo que a hipótese alternativa virá com o sinal de "menor"(<):

$$\begin{aligned} & H_0 : \mu = 15 \text{ dias} \quad \text{(a nova terapia tem o mesmo tempo de tratamento médio que a anterior)} \\ & H_1 : \mu < 15 \text{ dias} \quad \text{(a nova terapia reduz o tempo de tratamento médio)} \end{aligned}$$

Assim, este é um teste unilateral, onde a região de rejeição existe apenas na cauda da esquerda da distribuição, como se observa na figura seguinte.



Temos $\sigma = 3$, $n = 70$ e $\bar{x} = 14.6$, logo:

$$\text{valor-}p = P\left(\bar{X} - \mu \leq \bar{x} - \mu \mid \mu = \mu_0\right) = P\left(Z \leq \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right) = P\left(Z \leq \frac{14.6 - 15}{3/\sqrt{70}}\right) \approx 13.2\%.$$

Note que, neste exercício, o nível de significância do teste (α) não é especificado. No entanto, valores de α superiores a 10% não se utilizam na prática, pois implicam uma razoável probabilidade de se incorrer num *erro do tipo I*, isto é, de se rejeitar H_0 quando esta é verdadeira. Assim, como obtivemos valor-$p > 10\%$, teremos de <u>aceitar H_0</u> (e rejeitar H_1), pelo que <u>não há evidência suficiente</u> para afirmar que a nova terapia reduz a duração média do tratamento.

3. Suponha que X é uma variável aleatória uniforme com $\theta \leq x \leq \theta + 1$. Queremos testar a hipótese nula $H_0 : \theta = 0$ versus a hipótese alternativa $H_1 : \theta > 0$. Considere um teste com $\alpha = 0.05$. Determine (i) a região de aceitação e (ii) a potência do teste $(1 - \beta)$.

Resolução:
$X \sim U(\theta, \theta + 1) \Leftrightarrow f_X(x) = 1,\ \theta \leq x \leq \theta + 1$; $H_0 : \theta = 0$, $H_1 : \theta > 0$, $\alpha = 0.05$. Embora o enunciado não o refira explicitamente, subentende-se que este teste de hipóteses é efetuado com base numa única observação x da v.a. X (isto é, assume-se $n = 1$).

(i) Note que $E(X) = (2\theta + 1)/2$ cresce com θ, logo uma observação x com um valor elevado favorece a hipótese H_1 ($\theta > 0$). Assim, a região de rejeição de H_0 será do lado direito da distribuição (ver figura abaixo, lado esquerdo). Concretamente, rejeitamos H_0 se $P(X \geq x \mid \theta = 0) < \alpha$, ou seja, se $x > 1 - \alpha$. A região de aceitação é, assim, o intervalo $(-\infty,\ 1 - \alpha] = (-\infty,\ 0.95]$.

(ii) Num teste de hipóteses, podem ocorrer dois tipos de erros:

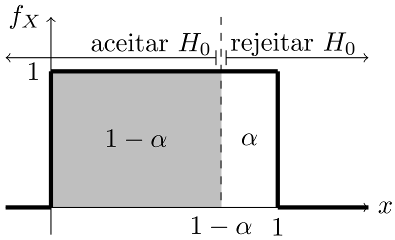
- Erro do tipo I – rejeitar H_0 quando esta é verdadeira. Prob(erro do tipo I) $= \alpha$. (nível de significância do teste)
- Erro do tipo II – aceitar H_0 quando esta é falsa (e H_1 verdadeira). Prob(erro do tipo II) $= \beta$.

Temos então $\beta = \text{Prob}(\text{aceitar } H_0 \mid H_1 \text{ verdadeira})$. Assumindo H_1 verdadeira, teremos $\theta = \theta_1 > 0$, onde θ_1 é um valor arbitrário mas que satisfaz H_1 (ver figura abaixo, lado direito). Por outro lado, aceitamos H_0 se a amostra pertencer à região de aceitação (calculada anteriormente), logo:

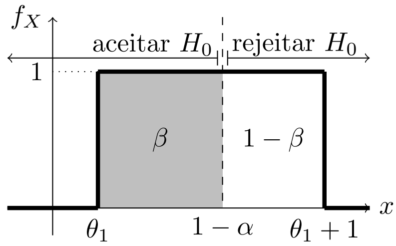
$$\begin{aligned}\beta &= P\left(X \in (-\infty,\ 0.95] \mid \theta = \theta_1\right) = P(X \leq 0.95 \mid \theta = \theta_1) \\ &= \begin{cases} \int_{\theta_1}^{0.95} 1\,dx = 0.95 - \theta_1, & \text{se } 0 < \theta_1 \leq 0.95 \\ 0, & \text{se } \theta_1 \geq 0.95 \end{cases}.\end{aligned}$$

A potência do teste é então $1 - \beta = \begin{cases} \theta_1 + 0.05, & \text{se } 0 < \theta_1 \leq 0.95 \\ 1, & \text{se } \theta_1 \geq 0.95 \end{cases}.$

(i) H_0 verdadeira (isto é, $\theta = 0$)



(ii) H_1 verdadeira (isto é, $\theta = \theta_1 > 0$)



4. Em 1000 lançamentos de moeda independentes, obtemos 560 caras e 440 coroas. É razoável concluir que a moeda é equilibrada?

Resolução:
Seja $X_i \in \{0,1\}$ a v.a. que representa sair "cara"($X_i = 1$) ou "coroa"($X_i = 0$) no i-ésimo lançamento ($i \in \{1, 2, \cdots, n\}$) e p a probabilidade de sair "cara"em cada lançamento (ou seja, $p = P(X_i = 1)$). A moeda é equilibrada se $p = 0.5$, logo o teste de hipóteses pretendido é $H_0 : p = 0.5$, $H_1 : p \neq 0.5$ (teste bilateral).

Método de resolução 1 (aproximação pela distribuição normal, válida para n grande):
Como já vimos anteriormente (ex. 8 da Ficha 10), pelo teorema do limite central, temos, para n elevado, $Z = \frac{\bar{X}-p}{\sqrt{p(1-p)/n}} \sim N(0,1)$. Temos $n = 1000$ lançamentos e a proporção observada na amostra foi $\bar{x} = 560/1000 = 0.560$, logo:

$$\text{valor-}p = 2P\left(Z \geq \frac{|\bar{x}-p|}{\sqrt{p(1-p)/n}} \,\middle|\, p = 0.5\right) = 2P\left(Z \geq \frac{|0.560-0.5|}{\sqrt{0.5(1-0.5)/1000}}\right) \approx 0 \implies \text{rejeitar } H_0.$$

Concluímos que não é razoável assumir-se que a moeda é equilibrada.

Método de resolução 2 (resolução exata, válida para qualquer n, mas pouco prática para n grande):
Seja $Y = \sum_{i=1}^{n} X_i$ o número de vezes que se obteve "cara"nos n lançamentos. Temos $Y \sim \text{Binom}(n,p)$ e $E(Y) = np$. Fizeram-se $n = 1000$ lançamentos e observou-se $y = 560$, logo:

$$\begin{aligned}
\text{valor-}p &= P\left(|Y - E(Y)| \geq |y - E(Y)| \,\middle|\, p = 0.5\right) = P\left(|Y - np| \geq |y - np| \,\middle|\, p = 0.5\right) \\
&= P\left(|Y - 500| \geq 60 \,\middle|\, p = 0.5\right) = P(Y \geq 560 \mid p = 0.5) + P(Y \leq 440 \mid p = 0.5) \\
&\approx 0 \implies \text{rejeitar } H_0.
\end{aligned}$$

Concluímos que não é razoável assumir-se que a moeda é equilibrada.

5. O número de falhas num metro de um determinado tipo de fio tem uma distribuição Poisson. Alega-se que o número médio de falhas é de 0.02 por metro. Uma amostra aleatória 100 fios com um comprimento de um metro cada um revela um total de 6 falhas.

   (a) Prove que a soma de duas variáveis aleatórias independentes de Poisson produz uma variável aleatória de Poisson, isto é, se $X \sim Pois(\lambda_a)$, e $Y \sim Pois(\lambda_b)$ então $Z = X + Y \sim Pois(\lambda_a + \lambda_b)$.

   (b) Determine a hipótese nula como a média de falhas em 100 fios de um metro.

   (c) Use o resultado em (a) para discutir se a informação apoia a afirmação se desejamos um $\alpha = 0.05$.

Resolução:

(a) Sejam f_X, f_Y e f_Z as funções de probabilidade de X, Y e Z, respetivamente, e seja $f_{X,Y}$ a função de probabilidade conjunta do par (X,Y). Temos X e Y independentes, logo $f_{X,Y}(x,y) = f_X(x)f_Y(y)$

para todo o (x,y). Temos ainda $Z = X + Y$, pelo que podemos obter f_Z da seguinte forma:

$$\begin{aligned}
f_Z(z) = P(Z = z) = P(X + Y = z) &= \sum_{(x,y):\, x+y=z} P(X = x \wedge Y = y) \\
&= \sum_{(x,y):\, x+y=z} f_{X,Y}(x,y) = \sum_{(x,y):\, x+y=z} f_X(x) f_Y(y) \\
&= \sum_{\forall x} f_X(x) f_Y(z-x). \qquad \text{(ver nota no final do exercício)}
\end{aligned}$$

Como $X \sim \text{Pois}(\lambda_a)$ e $Y \sim \text{Pois}(\lambda_b)$, temos $f_X(x) = e^{-\lambda_a}\lambda_a^x/x!$, $x \geq 0$, e $f_Y(z-x) = e^{-\lambda_a}\lambda_a^{z-x}/(z-x)!$, $z - x \geq 0$. Assim, $f_Z(z) = 0$ para $z < 0$. Finalmente, para $z \geq 0$,

$$\begin{aligned}
f_Z(z) &= \sum_{x:\, x\geq 0 \wedge z-x\geq 0} \frac{e^{-\lambda_a}\lambda_a^x}{x!}\frac{e^{-\lambda_b}\lambda_b^{z-x}}{(z-x)!} = e^{-\lambda_a-\lambda_b}\sum_{x=0}^{z}\frac{\lambda_a^x\lambda_b^{z-x}}{x!(z-x)!} \\
&= \frac{e^{-\lambda_a-\lambda_b}}{z!}\sum_{x=0}^{z}\frac{z!}{x!(z-x)!}\lambda_a^x\lambda_b^{z-x} \\
&= \frac{e^{-\lambda_a-\lambda_b}}{z!}\sum_{x=0}^{z}\binom{z}{x}\lambda_a^x\lambda_b^{z-x} \qquad \textstyle\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} = (a+b)^n \ \text{(binómio de Newton)} \\
&= \frac{e^{-(\lambda_a+\lambda_b)}(\lambda_a+\lambda_b)^z}{z!},
\end{aligned}$$

logo $Z \sim \text{Pois}(\lambda_a + \lambda_b)$. $\square$

Nota: Como resultado intermédio, demonstrámos que, dadas duas v.a. discretas e independentes X e Y (com qualquer função de probabilidade), a v.a. $Z = X + Y$ tem função de probabilidade dada por $f_Z(z) = \sum_{\forall x} f_X(x) f_Y(z-x)$. Esta expressão deverá ser-lhe familiar, pois corresponde à *convolução* (discreta) de f_X com f_Y. O resultado análogo mantém-se verdadeiro para v.a. contínuas, devendo aí considerar-se as respetivas funções densidade de probabilidade e a operação de convolução para o caso contínuo:

$$X \text{ e } Y \text{ v.a. contínuas e independentes, } Z = X + Y \implies f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx.$$

A demonstração para o caso contínuo é um pouco mais elaborada, mas pode tentar fazê-la como exercício. (Sugestão: Comece por encontrar a função de distribuição acumulada F_Z e, a partir desta, obtenha a densidade de probabilidade desejada, $f_Z = dF_Z/dz$.)

(b) $X_i \sim \text{Pois}(\lambda)-$ número de falhas no fio i ($i \in \{1, 2, \cdots, n\}$), que tem comprimento de 1m. $Y = \sum_{i=1}^{n} X_i-$ número total de falhas nos n fios. $E(Y) = nE(X_i) = n\lambda$. Alega-se que $\lambda = 0.02$ falhas e tem-se $n = 100$ fios. Assim, $H_0 : E(Y) = 100 \times 0.02 = 2$ falhas.

(c) Pelo resultado em (a), conclui-se que $Y \sim \text{Pois}(n\lambda)$. Queremos verificar se os fios cumprem a especificação dada pelo fabricante, pelo que será razoável definir a hipótese alternativa como $H_1 : E(Y) > 2$ falhas (teste unilateral). Na amostra obtida, observaram-se $y = 6$ falhas. Assim,

$$\begin{aligned}
\text{valor-}p &= P(Y \geq y \mid E(Y) = 2) = P(Y \geq 6 \mid n\lambda = 2) = 1 - P(Y \leq 5 \mid n\lambda = 2) \\
&\approx 1.7\% < 5\% = \alpha \implies \text{rejeitar } H_0.
\end{aligned}$$

A amostra obtida não apoia a afirmação de que $E(Y) = 2$ falhas, ou seja, de que $\lambda = 0.02$ falhas/m (conclui-se que λ é provavelmente superior).

6. Suponha que o número de acidentes no metro de Paris ao longo de um determinado ano segue uma distribuição de Poisson. O número médio de acidentes nos últimos anos é de 15. Este ano, apenas 10 indivíduos estiveram envolvidos em incidentes. É provável que a segurança no metro tenha aumentado significativamente?

Resolução:

$X \sim \text{Pois}(\lambda)$. Tinha-se um número médio de $\lambda = 15$ acidentes/ano, mas no último ano registaram-se apenas $x = 10$ acidentes, pelo que se quer testar a hipótese de que a média da distribuição (λ) tenha diminuído. Assim, temos $H_0 : \lambda = 15$, $H_1 : \lambda < 15$ (teste unilateral) e:

$$\text{valor-}p = P(X \leq 10 \mid \lambda = 15) \approx 11.85\% > 10\% \implies \text{não rejeitar } H_0.$$

Concluímos, assim, que a amostra observada não é suficiente para assumir que o número médio de acidentes por ano tenha diminuído.