49. Sejam:
X – v.a. que representa o número de empregados, em 100, contagiados
p – proporção de empregados, em todo o hospital, que estão contagiados
É necessário realizar um teste de hipóteses sobre p, com o objectivo de averiguar se pode concluir-se que $p < 0.25$; sendo esta a proposição que motiva a realização do teste, será expressa na hipótese alternativa, H_1. Para que a proposição seja falsa, basta que $p = 0.25$, o que será expresso na hipótese nula, H0.
Hipóteses a testar:
H_0: $p = 0.25$ *vs* H_1: $p < 0.25$
Estatística de teste (E.T.):
$n = 100$ é suficientemente "grande" $\Rightarrow Z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} \underset{H_0}{\overset{a}{\cap}} \mathcal{N}(0, 1)$, onde $\hat{p} = X/n$ e p_0 é o valor especificado nas hipóteses.
Valor observado da E.T. (V.O.):
$$x = 23 \Rightarrow \hat{p} = 23/100 = 0.23 \Rightarrow z_0 = \frac{0.23 - 0.25}{\sqrt{\dfrac{0.25 \times 0.75}{100}}} = \frac{-0.2}{\sqrt{0.1875}} \approx -0.46$$
Valor-p:
$P = P(\mathcal{N}(0, 1) \leq -0.46) = \Phi(-0.46) = 1 - \Phi(0.46) = 1 - 0.67724 = 0.32276$
Decisão:
Rejeita-se H_0 para níveis de significância $\alpha \geq 0.32276 \Rightarrow$ Não se rejeita H_0 para qualquer nível usual
Conclusão:
Os dados recolhidos não fornecem evidência de que a proporção de contagiados seja inferior a 25%.

50. Sejam:
X – v.a. que representa o número de jovens, em 40, que pretendem ingressar no ensino superior
p – proporção de jovens, em todo a região, que pretendem ingressar no ensino superior
Hipóteses a testar:
H_0: $p = 0.7$ *vs* H_1: $p > 0.7$
E.T.:
Admitindo que $n = 40$ é suficientemente "grande" $\Rightarrow Z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} \underset{H_0}{\overset{a}{\cap}} \mathcal{N}(0, 1)$, onde $\hat{p} = X/n$ e p_0 é o valor especificado nas hipóteses.
V.O.:
$$x = 30 \Rightarrow \hat{p} = 30/40 = 0.75 \Rightarrow z_0 = \frac{0.75 - 0.7}{\sqrt{\dfrac{0.7 \times 0.3}{40}}} = \frac{0.05}{\sqrt{0.00525}} \approx 0.69$$
Valor-p:
$P = P(\mathcal{N}(0, 1) \geq 0.69) = 1 - \Phi(0.69) = 1 - 0.7549 = 0.2451$
Decisão:
Rejeita-se H_0 para níveis de significância $\alpha \geq 0.2451 \Rightarrow$ Não se rejeita H_0 para qualquer nível usual
Conclusão:
Os dados recolhidos não fornecem evidência de que a afirmação dos autarcas corresponda à verdade.

51. Seja X – v.a. que representa a pressão arterial diastólica (em mmHg) de um homem escolhido ao acaso, com μ = E(X)
Hipóteses a testar:
H_0: $\mu = 84$ *vs* H_1: $\mu < 84$
E.T.:
Nada se sabe sobre a distribuição de X; admitindo que n = 38 é suficientemente "grande" ⇒
$\Rightarrow Z = \frac{\bar{X}-\mu_0}{S}\sqrt{n} \underset{H_0}{\overset{a}{\cap}} \mathcal{N}(0, 1)$, onde μ_0 é o valor especificado nas hipóteses.
V.O.:
$\sum_i x_i = 3129;\ \sum_i x_i^2 = 260335 \Rightarrow \bar{x} \approx 82.342;\ s \approx 8.521 \Rightarrow z_0 = \frac{82.342-84}{8.501}\sqrt{38} \approx -1.2$
Valor-p:
P = P($\mathcal{N}(0, 1) \le -1.2) = \Phi(-1.2) = 1 - \Phi(1.2) = 1 - 0.88493 = 0.11507$
Decisão:
Rejeita-se H_0 para níveis de significância $\alpha \ge 0.10935$ ⇒ Não se rejeita H_0 para qualquer nível usual
Conclusão:
Os dados recolhidos não fornecem evidência para afirmar que o valor médio da pressão arterial diastólica seja inferior a 84 mmHg.

52. Seja X – v.a. que representa o tempo (em minutos) que uma ambulância, escolhida ao acaso, demora a chegar ao destino; $X \cap \mathcal{N}(\mu, \sigma^2=9)$
Hipóteses a testar:
H_0: $\mu = 10$ *vs* H_1: $\mu > 10$
E.T.:
$X \cap$ Normal e σ^2 é conhecido $\Rightarrow Z = \frac{\bar{X}-\mu_0}{\sigma}\sqrt{n} \underset{H_0}{\cap} \mathcal{N}(0, 1)$, onde μ_0 é o valor especificado nas hipóteses.
V.O.:
$\bar{x} = 13 \Rightarrow z_0 = \frac{13-10}{3}\sqrt{16} = 4$
Região de Rejeição (R.R.):
$\alpha = 0.05 \Rightarrow z_{1-\alpha} = z_{0.95} = 1.645 \Rightarrow R_{0.05} = [1.645, +\infty)$
Decisão:
$z_0 \in R_{0.05}$ ⇒ Rejeita-se H_0 ao nível de significância de 5%
Conclusão:
Os dados recolhidos fornecem evidência para afirmar que o tempo médio é superior a 10 minutos.
Alternativamente:
Valor-p:
P = P($\mathcal{N}(0, 1) \ge 4) \approx 0$
Decisão:
Rejeita-se H_0 para níveis de significância $\alpha \ge 0$ ⇒ Rejeita-se H_0 para qualquer nível de significância

53. Seja X – v.a. que representa o valor da amilase (em unidades por 100ml) presente no sangue de um indivíduo escolhido ao acaso, com $\mu = E(X)$

Hipóteses a testar:

H_0: $\mu = 120$ *vs* H_1: $\mu \neq 120$

E.T.:

Nada se sabe sobre a distribuição de X e $n = 15$ é demasiado "pequeno";

Admitindo que $X \cap \mathcal{N}(\mu, \sigma^2)$, sendo σ^2 desconhecido $\Rightarrow$ $T = \frac{\bar{X} - \mu_0}{S}\sqrt{n} \underset{H_0}{\cap} t_{(n-1)}$, onde μ_0 é o valor especificado nas hipóteses

V.O.:

$\bar{x} = 96;\ s = 35 \Rightarrow t_0 = \frac{96 - 120}{35}\sqrt{15} \approx -2.66$

Valor-p:

$P = 2\times\min\{ P(t_{(14)} \leq -2.66), P(P(t_{(14)} \geq -2.66)\} = 2\times P(t_{(14)} \leq -2.66) = 2[1 - P(t_{(14)} \leq 2.66)]$

Com base nos valores tabelados dos quantis da distribuição t:

$2.62449 < 2.66 < 2.97684 \Rightarrow P(t_{(14)} \leq 2.62449) < P(t_{(14)} \leq 2.66) < P(t_{(14)} \leq 2.97684) \Leftrightarrow$

$\Leftrightarrow 0.99 < P(t_{(14)} \leq 2.66) < 0.995 \Leftrightarrow 2(1 - 0.995) < 2[1 - P(t_{(14)} \leq 2.66)] < 2(1 - 0.99) \Leftrightarrow 0.01 < P < 0.02$

Alternativamente:

Com base no cálculo numérico da função de distribuição da t:

$P(t_{(14)} \leq 2.66) = 0.99067 \Rightarrow P = 2[1 - 0.99067] \Leftrightarrow P = 2\times 0.00933 = 0.01866$

Decisão:

Rejeita-se H_0 para níveis de significância $\alpha \geq P \Rightarrow$ Ao nível usual de 1% não se rejeita H_0 mas aos níveis usuais de 5% e 10% já se rejeita

Conclusão:

Ao nível de significância usual de 1%, os dados recolhidos não fornecem evidência para afirmar que o valor médio da amílase seja diferente de 120 unidades por 100ml de sangue, mas aos níveis de significância de 5% e 10% já fornecem essa evidência.

54. Seja X – v.a. que representa o peso (em Kg), ao fim de 5 meses, de um porco da espécie em questão, com $\mu = E(X)$ e $\sigma^2 = var(X)$

Hipóteses a testar:

H_0: $\mu = 75$ *vs* H_1: $\mu \neq 75$

E.T.:

Admitindo que $n = 50$ é suficientemente "grande" $\Rightarrow$ $Z = \frac{\bar{X} - \mu_0}{S}\sqrt{n} \overset{a}{\underset{H_0}{\cap}} \mathcal{N}(0, 1)$, onde μ_0 é o valor especificado nas hipóteses.

V.O.:

$\sum_i x_i = 3730;\ \sum_i (x_i - \bar{x})^2 = 330.57 \Rightarrow \bar{x} = 74.6;\ s \approx 2.597 \Rightarrow z_0 \approx \frac{74.6 - 75}{2.597}\sqrt{50} \approx -1.09$

Valor-p:

$P = 2\min\{P(\mathcal{N}(0, 1) \leq -1.09), P(\mathcal{N}(0, 1) \geq -1.09)\} = 2\Phi(-1.09) = 2(1-\Phi(1.09)) = 2(1-0.86214) = 0.27572$

Decisão:

Rejeita-se H_0 para níveis de significância $\alpha \geq 0.27572 \Rightarrow$ Não se rejeita H_0 para qualquer nível usual

Conclusão:

Os dados recolhidos fornecem evidência para afirmar que o produtor tem razões para estar satisfeito.

55. Seja X – v.a. que representa o QI de um indivíduo, escolhido ao acaso, com idade superior a 16 anos, consumidor de drogas e que está preso; $X \cap \mathcal{N}(\mu, \sigma)$

Hipóteses a testar:

H$_0$: $\mu = 107$ *vs* H$_1$: $\mu \neq 107$

E.T.:

$X \cap$ Normal e σ^2 é desconhecido $\Rightarrow$ $T = \frac{\bar{X} - \mu_0}{S}\sqrt{n} \underset{H_0}{\cap} t_{(n-1)}$, onde μ_0 é o valor especificado nas hipóteses.

V.O.:

$\sum_i x_i = 1650$; $\sum_i x_i^2 = 183722 \Rightarrow \bar{x} = 110$; $s \approx 12.598 \Rightarrow t_0 = \frac{110 - 107}{12.598}\sqrt{15} \approx 0.922$

Valor-p:

$P = 2\times\min\{ P(t_{(14)} \le 0.922), P(P(t_{(14)} \ge 0.922)\} = 2\times P(t_{(14)} \ge 0.922) = 2[1 - P(t_{(14)} \le 0.922)]$

Com base nos valores tabelados dos quantis da distribuição t:

$0.86805 < 0.922 < 1.07628 \Rightarrow P(t_{(14)} \le 0.86805) < P(t_{(14)} \le 0.922) < P(t_{(14)} \le 1.07628) \Leftrightarrow$

$\Leftrightarrow 0.8 < P(t_{(14)} \le 0.922) < 0.85 \Leftrightarrow 2(1 - 0.85) < 2[1 - P(t_{(14)} \le 0.922)] < 2(1 - 0.8) \Leftrightarrow 0.3 < P < 0.4$

Alternativamente:

Com base no cálculo numérico da função de distribuição da t:

$P(t_{(14)} \le 0.922) = 0.81393 \Rightarrow P = 2[1 - 0.81393] \Leftrightarrow P = 2\times 0.18607 = 0.37214$

Decisão:

Rejeita-se H$_0$ para níveis de significância $\alpha \ge P \Rightarrow$ Não se rejeita H$_0$ para qualquer nível usual

Conclusão:

Os dados recolhidos não fornecem evidência para afirmar que o valor médio do QI na nova área seja diferente de 107.

56. Seja X – v.a. que representa a pressão arterial sistólica (em mmHg) de uma mulher, com idade compreendida entre os 40 e os 80 anos, escolhida ao acaso; $X \cap \mathcal{N}(\mu, \sigma)$

Hipóteses a testar:

H$_0$: $\mu = 160$ *vs* H$_1$: $\mu < 160$

E.T.:

$X \cap$ Normal e σ^2 é desconhecido $\Rightarrow$ $T = \frac{\bar{X} - \mu_0}{S}\sqrt{n} \underset{H_0}{\cap} t_{(n-1)}$, onde μ_0 é o valor especificado nas hipóteses.

V.O.:

$\sum_i x_i = 2977$; $\sum_i x_i^2 = 445611 \Rightarrow \bar{x} = 148.85$; $s \approx 11.435 \Rightarrow t_0 = \frac{148.85 - 160}{11.435}\sqrt{20} \approx -4.36$

Valor-p:

$P = P(t_{(19)} \le -4.36) \approx 0$

Decisão:

Rejeita-se H$_0$ para níveis de significância $\alpha \ge 0 \Rightarrow$ Rejeita-se H$_0$ para qualquer nível de significância

Conclusão:

Os dados recolhidos fornecem evidência para afirmar que o valor médio da pressão sistólica é inferior a 160 mmHg.

57.

a) Sejam:

X – v.a. que representa o número de jovens, em 400 com idades compreendidas entre 18 e 21 anos, que estão desempregados

p_X – proporção de jovens, em toda a população com idades compreendidas entre 18 e 21 anos, que estão desempregados

$n_X = 400$ é suficientemente "grande" $\Rightarrow$ $IC_{95\%}(p_X) = \left(\hat{p}_X \mp z_{0.975}\sqrt{\hat{p}_X(1-\hat{p}_X)/n_X}\right)$, onde $\hat{p}_X = X/n_X$ e $z_{0.975}$ é o quantil de probabilidade 0.975 da distribuição Normal padrão

x = 80 $\Rightarrow$ $\hat{p}_X = 80/400 = 0.2$;

$z_{0.975}\sqrt{\hat{p}_X(1-\hat{p}_X)/n_X} = 1.96\sqrt{0.2\times 0.8/400} = 1.96\times 0.02 = 0.0392$

Então:

$$IC_{95\%}(p_X) = (\,0.2 - 0.0392,\ 0.2 + 0.0392) = (\,\underline{0.1608},\ \underline{0.2392})$$

b) -----

58.

a) Seja X_D – v.a. que representa a taxa de colesterol (cg/l) de um indivíduo, escolhido ao acaso, na população dos doentes; $X_D \cap \mathcal{N}(\mu_D, \sigma_D)$

<u>Hipóteses a testar</u>:

H_0: $\mu_D = 200$ *vs* H_1: $\mu_D < 200$

<u>E.T.</u>:

$X_D \cap$ Normal e σ_D^2 é desconhecido $\Rightarrow$ $T = \dfrac{\bar{X}_D - \mu_0}{S_D}\sqrt{n_D} \underset{H_0}{\cap} t_{(n_D-1)}$, onde μ_0 é o valor especificado nas hipóteses.

<u>V.O.</u>:

$$t_0 = \frac{197-200}{\sqrt{2000}}\sqrt{50} \approx -0.47$$

<u>R.R.</u>:

$\alpha = 0.05$: $-\,t_{0.95;49} \approx -z_{0.95} = -1.645$, pois os quantis da distribuição $t_{(49)}$ já estão muito próximos dos quantis da distribuição Normal padrão

$R_{0.05} = (-\infty, -1.645]$

<u>Decisão</u>:

$t_0 \notin R_{0.05} \Rightarrow$ Não se rejeita H_0 ao nível de significância de 5%

<u>Conclusão</u>:

Os dados recolhidos não fornecem evidência para afirmar que o valor médio da taxa de colesterol dos indivíduos doentes seja inferior a 200 cg/l.

Alternativamente:

<u>Valor-p</u>:

$P = P(t_{(49)} \leq -0.47) \approx P(\mathcal{N}(0, 1) \leq -0.47) = \Phi(-0.47) = 1 - \Phi(0.47) = 1 - 0.68082 = 0.31918$

<u>Decisão</u>:

Rejeita-se H_0 para níveis de significância $\alpha \geq 0.31918 \Rightarrow$ Não se rejeita H_0 ao nível de significância de 5%

b) -----