Ex.3.26

6 de dezembro de 2020 21:04

3.26. Num estudo sobre o número de folhas por planta de tabaco, obtiveram-se os seguintes dados:

n º de folhas	17	18	19	20	21	22	23	24
n º de plantas	3	22	44	42	22	10	6	1

a) Indique a unidade estatística e a variável em estudo.

b) Determine a média e a mediana da amostra. Compare-as e comente.

c) Com base nesta amostra poder-se-á afirmar que 90% das plantas de tabaco têm 21 ou menos folhas? Justifique.

a) A VARIÁVEL EM ESTUDO É O NÚMERO DE FOLHAS POR PLANTA.
A UNIDADE ESTATÍSTICA É A PLANTA.

b) CONSIDERANDO A TABELA DE FREQUÊNCIAS :

	xi	ni	xni	Fi
1	17	3	51	0.0200000
2	18	22	396	0.1666667
3	19	44	836	0.4600000
4	20	42	840	0.7400000
5	21	22	462	0.8866667
6	22	10	220	0.9533333
7	23	6	138	0.9933333
8	24	1	24	1.0000000
9	NA	150	2967	NA

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i n_i}{n} = \frac{2967}{150} = 19.78 \text{ folhas}$$

$$\tilde{x} = \left(x_{(75)} + x_{(76)}\right)/2 = 20$$

$\bar{x} < \tilde{x}$ POIS HÁ UMA ASSIMETRIA DA DIST. PARA A ESQUERDA

c) SEJA p A PROPORÇÃO DE PLANTAS DE TABACO COM $\underbrace{\text{NÚMERO DE FOLHAS INFERIOR OU IGUAL A 21}}_{\text{"SUCESSO"}}$.

SEJA X A V.A. Nº DE PLANTAS COM NÚMERO DE FOLHAS IGUAL OU INFERIOR A 21 EM $\underbrace{n=150 \text{ PLANTAS}}_{\text{PROVAS}}$

SE p É CONSTANTE E AS PROVAS SÃO INDEPENDENTES ENTÃO $X \cap B(n=150, p)$

A QUESTÃO QUE SE COLOCA É SE SE PODE ADMITIR QUE $p = 0.9$.

HÁ DUAS FORMAS DE RESPONDER À QUESTÃO:
A – CONSTRUIR UM I.C. PARA p
ou
B – TESTAR A HIPÓTESE $H_0 : p = 0.9$

A:

p	provas repetidas independência	$\hat{p} - z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n} < p < \hat{p} + z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n}$

A:

p	provas repetidas independência n 'grande'	$\hat{p} - z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n} < p < \hat{p} + z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n}$

NESTE CASO $\hat{p}$ É A PROPORÇÃO DE PLANTAS COM Nº FOLHAS IGUAL OU INFERIOR À 21 NA AMOSTRA

$$\hat{p} = \frac{3+22+44+42+22}{150}$$

$$= \frac{133}{150} = 0.886$$

$$\hat{q} = 1 - \hat{p} = 0.114$$

$$n = 150$$

	xi	ni	xni	Fi
1	17	3	51	0.0200000
2	18	22	396	0.1666667
3	19	44	836	0.4600000
4	20	42	840	0.7400000
5	21	22	462	0.8866667
6	22	10	220	0.9533333
7	23	6	138	0.9933333
8	24	1	24	1.0000000
9	NA	150	2967	NA

· ESCOLHENDO $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

$z_\epsilon : P[Z > z_\epsilon] = \epsilon$

ϵ	0.100	0.050	0.025	0.010	0.005
z_ϵ	1.282	1.645	1.960	2.326	2.576

O I.C. A 95% PARA p É

$$\left]0.886 - 1.96\sqrt{\frac{0.886 \times 0.114}{150}}, 0.886 + 1.96\sqrt{\frac{0.886 \times 0.114}{150}}\right[$$

.886-1.96*sqrt(0.886*0.114/150)=0.835139605034959

.886+1.96*sqrt(0.886*0.114/150)=0.936860394965041

ISTO É, COM 95% DE CONFIANÇA, PODE AFIRMAR-SE QUE $0.835 < p < 0.936$

PORTANTO PODE ADMITIR-SE QUE $p = 0.9$.

B :

p	provas repetidas independência n 'grande'	$p = p_0$ $p \leq p_0$ $p \geq p_0$	$p \neq p_0$ $p > p_0$ $p < p_0$	$Z = \frac{X - n\,p_0}{\sqrt{n\,p_0\,(1-p_0)}}$	$Z < -z_{\alpha/2}$ ou $Z > z_{\alpha/2}$ $Z > z_\alpha$ $Z < -z_\alpha$

NESTE CASO, PRETENDE-SE AVERIGUAR SE $p = 0.9$

Passos a seguir na construção de um **teste estatístico**

1. Identificar o(s) parâmetro(s); especificar $\mathbf{H}_0$ **e** $\mathbf{H}_1$ e o nível de significância α.
2. Escolher uma variável aleatória – **Estatística de Teste**, que sob $\mathbf{H}_0$ terá distribuição conhecida (pelo menos aproximadamente).
3. Definir a **região de rejeição** ou **região crítica** – **RC** (conjunto de valores da estatística que são menos "plausíveis" caso $\mathbf{H}_0$ seja verdadeira, portanto levam a rejeitar $\mathbf{H}_0$).
4. Calcular **o valor** da estatística de teste, para a amostra observada.
5. Se o valor calculado $\in$ RC $\longrightarrow$ rejeita-se H_0
   o valor calculado $\notin$ RC $\longrightarrow$ não se rejeita H_0

1. O PARÂMETRO É p, A PROB. DE SUCESSO PARA $X \cap B(n, p)$

$$H_0: p = 0.9 \quad \text{vs} \quad H_1: p \neq 0.9$$

2. A ESTATÍSTICA DE TESTE É A V.A.

$$Z = \frac{X - n\,p_0}{\sqrt{n\,p_0\,(1-p_0)}}$$

SOB H_0, $Z \cap N(0,1)$

3. A REGIÃO CRÍTICA É DADA POR $Z < -z_{\alpha/2}$ ou $Z > z_{\alpha/2}$
ESCOLHENDO $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$
A REGIÃO CRÍTICA É:

$$]-\infty, -1.96[\cup]1.96, +\infty[$$

4. O VALOR OBSERVADO DA ESTATÍSTICA DE TESTE É

$$z_{CALC} = \frac{133 - 150 \times 0.9}{\sqrt{150 \times 0.9 \times 0.1}} = -0.544$$

(133-150*0.90)/sqrt(150*0.1*0.9)=-0.544331053951817

5. COMO z_{CALC} NÃO PERTENCE À REGIÃO CRÍTICA, ENTÃO NÃO SE REJEITA H_0 E ADMITE-SE QUE que 90% das plantas de tabaco têm 21 ou menos FOLHAS.