Lista 2 - MAE0315

01/06/2020

1)

Considere

$$N = 90000 \quad N_1 = 35000 \quad N_2 = 55000$$

Onde N representa o tamanho da população, N_1 representa o tamanho do estrato *casas* e N_2 representa o tamanhho do estrato *apartamentos*. Sabendo que o consumo médio de eletricidade para as casas(μ_1) é o dobro do que o consumo de energia elétrica para apartamentos(μ_2).Ou seja,

$$\mu_1 = 2\mu_2$$

como sabemos também que o desvio padrão é proporcional à média, podemos afirmar

$$\sigma_1 = 2\sigma_2$$

a)

Para alocar de modo ótimo, uma amostra estratificada (com reposição) de 900 domicílios, para estimar o consumo médio de energia para os domicílios da cidade é possível utilizar a alocação ótima de Neyman, obtendo assim o seguinte plano amostral:

$$n_1 = n\frac{W_1\sigma_1/\sqrt{c_1}}{W_1\sigma_1/\sqrt{c_1} + W_2\sigma_2/\sqrt{c_2}}$$

$$n_2 = n\frac{W_2\sigma_2/\sqrt{c_2}}{W_1\sigma_1/\sqrt{c_1} + W_2\sigma_2/\sqrt{c_2}}$$

Como não possuimos nenhuma informação para o custo em cada estrato, consiraremos que o custo é constante em todos os estratos. Desta forma obteremos,

$$n_1 = n\frac{W_1\sigma_1}{W_1\sigma_1 + W_2\sigma_2} = 900\frac{\frac{35000}{90000}\sigma_1}{\frac{35000}{90000}\sigma_1 + \frac{55000}{90000}\sigma_2}$$

$$n_2 = n\frac{W_2\sigma_2}{W_1\sigma_1 + W_2\sigma_2} = 900\frac{\frac{55000}{90000}\sigma_2}{\frac{35000}{90000}\sigma_1 + \frac{55000}{90000}\sigma_2}$$

Porém, como sabemos que $\sigma_1 = 2\sigma_2$,

$$n_1 = 900\frac{\frac{35000}{90000}\sigma_1}{\frac{35000}{90000}\sigma_1 + \frac{55000}{90000}\sigma_2} = 900\frac{\frac{35000}{90000}2\sigma_2}{\frac{35000}{90000}2\sigma_2 + \frac{55000}{90000}\sigma_2} = 900\frac{70000}{125000} = 504$$

$$n_2 = 900\frac{\frac{55000}{90000}\sigma_2}{\frac{35000}{90000}\sigma_1 + \frac{55000}{90000}\sigma_2} = 900\frac{55000}{125000} = 396$$

b)

Considere P_1 e P_2 as proporções de casas e apartamentos que seguem as práticas de economiia de energia repectivamente e sabemos previamente que

$$P_1 = 0.45 \quad e \quad P_2 = 0.25$$

e considere

$$Q_i = 1 - P_1, \quad i = 1, 2$$

Então para estimar a proporção de demicílios que seguem práticas de economia de energia podemos utilizar a alocação ótima de Neyman, para o caso onde o estimador é a proporção. Assim como no item **a)**, não temos nenhuma informação sobre o custo em cada estrato, então iremos considerá-lo constante em todos os estratos, obtendo as seguintes expressões:

$$n_1 = n \frac{N_1\sqrt{P_1Q_1}}{N_1\sqrt{P_1Q_1} + N_2\sqrt{P_2Q_2}}$$

$$n_2 = n \frac{N_2\sqrt{P_2Q_2}}{N_1\sqrt{P_1Q_1} + N_2\sqrt{P_2Q_2}}$$

Substituindo os valores correspondentes, obtemos

$$n_1 = 900 \frac{35000\sqrt{0.45 * 0.55}}{35000\sqrt{0.45 * 0.55} + 55000\sqrt{0.25 * 0.75}} \approx 380$$

$$n_2 = 900 \frac{55000\sqrt{0.25 * 0.75}}{35000\sqrt{0.45 * 0.55} + 55000\sqrt{0.25 * 0.75}} \approx 520$$

c)

Considerando o plano amostral descrito, temos que em cada estrato foi feito uma amostra aleatória simples sem reposição em cada estrato, seguindo uma alocação uniforme. Desta forma, considere a seguinte estatística:

$$\bar{y}_{es} = \sum_{i=1}^{2} W_i \bar{y}_i$$

que é um estimador não viesado da média populacional, e também um estimador para sua variância

$$\hat{Var}_{AASs}(\bar{y}_{es}) = \sum_{i=1}^{2} Wi^2(1 - f_h)\frac{s_h^2}{n_h}$$

Fazendo as contas:

$$\bar{y}_{es} = (\frac{35000}{90000}800 + \frac{55000}{90000}500) \approx 617 \quad (KWh)$$

$$\hat{Var}_{AASs}(\bar{y}_{es}) = (\frac{35}{90})^2(1 - \frac{45}{3500})\frac{200^2}{450} + (\frac{55}{90})^2(1 - \frac{45}{5500})\frac{120^2}{450} \approx 13.27 + 11.85 = 25.12 \quad (KWh)^2$$

E portanto o erro padrão é dado por:

$$SE = \sqrt{\hat{Var}_{AASs}(\bar{y}_{es})} = \sqrt{25.12} \approx 5 \quad (KWh)$$

Assumindo normalidade pelo Teorema Central do Limite(TLC), contruiremos um intervalo de confiança de 95% para o consumo médio populacional:

$$IC(\mu, 95\%) = [617 - 1.96(5), 617 + 1.96(5)] = [607.2, 626.8]$$

2)

a)

O estimador razão é o mais adequado, pois trata-se de um problema onde o valor esperado do ganho do peso dos frangos com peso inicial nulo, com o uso da ração nova, é nulo. Para realizar um teste de hipóteses adequado, primeiro irei ajustar um modelo de regressão linear simples para o ganho do peso após 30 dias, com esta motivação considere o seguinte modelo:

$$y_i = \alpha + \beta x_i$$

```
##
## Call:
## lm(formula = yi ~ xi, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.0600 -0.0475 -0.0020  0.0510  0.0560
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.5400     0.3778   1.429  0.19079
## xi            1.0400     0.2516   4.134  0.00328 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05626 on 8 degrees of freedom
## Multiple R-squared:  0.6811, Adjusted R-squared:  0.6412
## F-statistic: 17.09 on 1 and 8 DF,  p-value: 0.003282
```

Pelo código do R apresentado, é coerente a hipótese de que o ganho após 30 dias é nulo, para frangos com peso inicial nulo, pois observamos um valor-p consideravelmente alto(0.20) e por isso não há evidência para rejeitar a hipótese nula

b)

Uma estimativa para o peso médio dos frangos após 30 dias é

$$\hat{\mu}_y = 2.156$$

e uma estimativa para seu erro padrão,com quatro casas decimais, é

$$SE(\hat{\mu}_y) = 0.0188$$

c)

Desconsiderando a informação sobre o peso inicial dos frangos. Temos uma amostra aleatória simples com reposição, pois não temos nenhuma informação sobre o tamanho da população de frangos. então o peso médio estimado dos frangos após 30 dias é

$$\hat{\mu}_y = \frac{1}{10}\sum_{i=1}^{10} y_i = 2.1$$

e uma estimativa para seu erro padrão com quatro casas decimais é dada por

$$SE(\hat{\mu}_y) = 0.0297$$