Intervalo de confiança para a média populacional

Resumo

Um intervalo de confiança (IC) é um intervalo estimado de um parâmetro θ de interesse de uma população. Em vez de estimar o parâmetro por um único valor, é dado um intervalo de estimativas prováveis. O quanto estas estimativas são prováveis será determinado pelo nível de confiança γ, que reflete a probabilidade do parâmetro populacional estar contido dentro deste intervalo. Assim, podemos definir um IC como:

$$IC(\theta\ ,\ \gamma) =]t_1, t_2[\quad (1)$$

sendo t_1 o limite inferior e t_2 o limite superior do intervalo. Devido à sua natureza aleatória, é improvável que duas amostras de uma determinada população irá render intervalos de confiança idênticos. Mas, se você repetir sua amostra várias vezes, uma determinada porcentagem dos intervalos de confiança resultantes conteria o parâmetro populacional desconhecido. Observe a Figura 1:

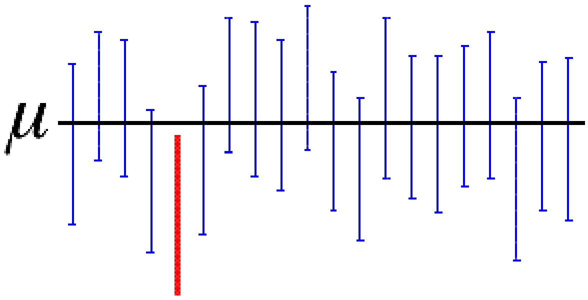

Figura 1: Diferentes amostras, com suas médias comparadas à média amostral μ

Aqui, a linha preta horizontal confiança representa o valor do parâmetro μ, que é a média desconhecida da população, . Os intervalos de confiança azuis verticais sobrepostos à linha horizontal contém o valor da média da população. O intervalo de confiança vermelho totalmente abaixo da linha horizontal não contém esse valor. Um intervalo de confiança de 95% indica que 19 em 20 amostras (95%) da mesma população produzem intervalos de confiança contendo o parâmetro da população. Vejamos um exemplo para a determinação de intervalos de confiança a seguir.

Exemplo: A duração da vida de uma peça de automóvel tem desvio padrão populacional igual a 5 horas. Com uma amostra aleatória de 100 peças, obteve-se a média amostral $\overline{x} = 500h$. Construa um IC com nível de confiança de 95% de confiança.

Solução:

Os limites superior e inferior do IC são definidos pela seguinte expressão:

$$IC = \overline{x} \ \pm \ Z_c \frac{\sigma}{\sqrt{n}} \quad (2)$$

Para o nosso exemplo, o desvio padrão amostral σ é 5 e o tamanho n da amostra é 100. Z_c é a variável que segue a distribuição normal padrão, cuja probabilidade associada é o nível de confiança γ. Retorne ao módulo relacionado à distribuição normal padrão, para relembrar a utilização da Tabela de distribuição normal padrão.

Em nosso caso, utiliza-se um nível de confiança de 95%, ou seja, a probabilidade da média populacional estar dentro do intervalo é de 0,95. Consultando a Tabela, vemos que o valor da variável Z_c correspondente a esse valor de probabilidade é $Z_c = 1,96$. Logo, os intervalos do IC são calculados usando (2):

$$t_1 = \overline{x} + 1,96.\frac{5}{\sqrt{100}} = 500 + 0,98 = 500,98 \quad (3)$$

$$t_2 = \overline{x} - 1,96.\frac{5}{\sqrt{100}} = 500 - 0,98 = 499,02 \quad (4)$$

$$IC = [499,02 \ ; \ 500,98] \quad (5)$$

Portanto, é possível afirmar com 95% de certeza que a média amostral está contida no intervalo obtido na equação (5)