

Intervalo de confiança para média

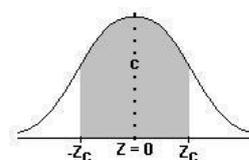
Professor Fiore

Com estatística é possível estimar os parâmetros populacionais a partir de amostras e para compreender bem isto é necessário distinguir os valores populacionais, como a média populacional (μ) e o desvio padrão populacional (σ), dos valores amostrais como a média amostral (\bar{x}) e desvio padrão amostral (s); onde cada qual tem seu próprio símbolo. Usando uma boa estratégia de amostragem, os valores amostrais tendem aos populacionais, mas sempre haverá certa diferença entre eles, pois um é um estimador pontual calculado com base nos dados da amostra e o outro é o valor 'real' da população.

Uma alternativa para melhorar os estimadores \bar{x} e s , são os intervalos de confiança, onde em vez de usar um **valor pontual** para tal, é usado uma **estimativa intervalar**, de acordo com uma confiança.

Intervalo de confiança para média – amostra grande ($n \geq 30$)

O procedimento para construção do intervalo de confiança para média tem base no **teorema do limite central**, que diz que, para uma amostra de tamanho grande ($n \geq 30$) a distribuição amostral das médias se comporta como uma distribuição normal, mesmo para dados não normalmente distribuídos.



Neste caso, uma 'margem de erro' pode ser calculada considerando como confiança a área sob a curva da distribuição normal e entre $-z_c$ e z_c . A tabela abaixo indica os principais valores para a confiança e quando nada for indicado considere por padrão uma confiança de $c = 95\%$, onde $z_c = 1,96$.

Intervalo de confiança para média ($n \geq 30$)						
Confiança (c)	80%	85%	90%	95%	98%	99%
z_c	1,28	1,44	1,65	1,96	2,33	2,58

Como o desvio padrão das médias das amostras é igual a razão entre o desvio padrão populacional e a raiz quadrada do tamanho da amostra, a margem de erro para a média, para uma amostra de tamanho n e considerando uma confiança c , pode ser calculada com a fórmula a seguir.

$$E = z_c \frac{\sigma}{\sqrt{n}}$$

E o intervalo de confiança para média pode ser escrito de uma das duas formas abaixo.

$$\bar{x} - E < \mu < \bar{x} + E$$

$$\bar{x} \pm E$$

Apesar da fórmula para a margem de erro indicar o uso do valor do desvio padrão populacional, para 'amostras grandes' ($n \geq 30$), o valor do desvio padrão amostral é tão próximo do parâmetro populacional que o substitui, $s \approx \sigma$.

É possível reduzir a margem de erro, sem alterar a confiança e sabendo previamente o valor do desvio padrão, pois de acordo com a **lei dos grandes números**, quanto maior o tamanho da amostra, mais aprimorado será o valor da média amostral em relação ao valor da média populacional. Assim para atingir uma margem de erro pré-determinada basta usar a fórmula ao lado.

$$n = \left(\frac{z_c \sigma}{E} \right)^2$$

Cálculo de intervalo de confiança para média – amostra pequena ($n < 30$)

Quando a amostra for pequena ($n < 30$) mas os dados forem normalmente distribuídos, há duas opções. Caso o desvio padrão amostral seja conhecido, o procedimento será o mesmo usado para amostras grandes ($n < 30$). Se o desvio padrão amostral não for conhecido, é necessário usar a distribuição t de Student no lugar da tabela da distribuição normal, mas o cálculo segue praticamente o mesmo, conforme o diagrama abaixo.

Para dados de amostras pequenas ($n < 30$) onde a população não for normalmente distribuída é sugerido aumentar o tamanho da amostra para 30 ou mais, caso contrário, não é possível construir o intervalo.

Quando usar a distribuição normal ou a distribuição t?

