# 1 Estimação de Parâmetros

Vários tipos de estudos tem o objetivo de obter conclusões (fazer inferências) a respeito de parâmetros de uma população. A impossibilidade de avaliar toda a população faz com que a partir de amostras possamos obter estimativas daqueles parâmetros. A generalização da amostra para a população é feita com o auxílio de um modelo estatístico para a situação em estudo, estas generalizações estão sempre associadas um grau de incerteza e, consequentemente, uma probabilidade de erro. A teoria da estimação preocupa-se com a obtenção do respectivo de um estimador para um determinado parâmetro, com intuito de descrever o seu comportamento com o menor erro possível.

Parâmetro: É uma constante (um número) que caracteriza uma população. Exemplo: média populacional $\mu$ , variância populacional $\sigma^2$ , etc. Em geral, os parâmetros são desconhecidos.

Estimador: É uma expressão algébrica utilizada para obter um valor aproximado de um parâmetro. Exemplo:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

.

Estimativa: É o valor numérico de um estimador. É determinada usando os dados amostrais.

Exemplo: Mediante uma pesquisa queremos conhecer o tamanho médio dos estudantes universitários do Brasil.

- População: Todas os estudantes universitários do Brasil;
- Amostra: por exemplo, 500 estudantes;
- Parâmetro: Média das alturas .
- Estimador:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

  .
- Estimativa: $\overline{X} = 1,7m$ (valor aproximado para $\mu$) .

# 2 Estimação

É um processo de indução, na qual usamos dados extraídos de uma amostra para produzir inferência sobre a população. Esta inferência só será válida se a amostra for significativa.

Tipos de Estimações de Parâmetros

1. Estimação Pontual;
2. Estimação Intervalar

## 2.1 Estimação Pontual

É usada quando a partir da amostra procura-se obter um único valor de certo parâmetro populacional, ou seja, obter estimativas a partir dos valores amostrais.

A estimativas são os valores amostrais obtidos para a média, variância, proporção, etc. Os valores de $\overline{X}$, $S^2$, $S$ estimam, respectivamente $\mu$, $\sigma^2$ e $\sigma$.

## 2.2 Estimação Intervalar

Uma outra maneira de se calcular um estimativa de um parâmetro desconhecido, é construir um intervalo de confiança $[a, b]$ para esse parâmetro com uma probabilidade de $1-\alpha$ (nível de confiança) de que o intervalo contenha o verdadeiro parâmetro, usando as distribuições de amostragem podemos obter expressões do tipo:

$$P(a \leq \mu \leq b) = 1-\alpha$$

Dessa maneira $\alpha$ será o nível de significância, isto é, o erro que se estará cometendo ao afirmar que o parâmetro está entre o limite inferior e o superior calculado.

### 2.2.1 Intervalo de Confiança para proporção $p$

Consideremos uma população cujos elementos podem ser classificados em dois tipos: Sucesso e Insucesso. Pretende-se estimar a proporção $p$ de sucessos na população.

Dada uma amostra de tamanho $n$, uma estimativa pontual de $p$ da proporção de sucessos é dada por

$$\hat{p} = \frac{x}{n}$$

.

Pelo teorema do limite cental, quando $n$ for suficientemente grande $\hat{p}$ tem distribuição aproximadamente normal, com média $\mu_{\hat{p}} = p$ e variância a $\sigma^2_{\hat{p}} = \frac{pq}{n}$, em que:

$$z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$$

Fixando uma probabilidade de confiança $(1-\alpha)$ , o intervalo de confiança para uma proporção pode ser obtido da seguinte forma:

$$P\left(\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} \leq p \leq \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right) = 1-\alpha$$

onde:$z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}$ é a margem de erro da proporção e $z_{\frac{\alpha}{2}}$ é o valor da curva normal padrão acima do qual encontramos uma área de $\frac{\alpha}{2}$.

Exemplo: Uma empresa de pesquisa de mercado faz contato com 30 pessoas para saber a satisfação a uma determinada marca de refrigerante, 12 delas respondem que gosta da referida marca. Obtenha o intervalo de confiança de 95% para proporção de pessoas que gostam da marca.

Nesse caso o sucesso é o gosto pela marca de refrigerante

$$\hat{p} = \frac{x}{n} = \frac{12}{30} = 0,40$$

Como $\hat{p} = 0,40$, temos que $\hat{q} = 1 - \hat{p} = 1 - 0,40 = 0,60$
Como queremos o intervalo de confiança a 95%, temos que:

$$1 - \alpha = 0,95 \Rightarrow \alpha = 1 - 0,95 = 0,05 \Rightarrow \frac{\alpha}{2} = \frac{0,05}{2} = 0,025$$

Assim, temos que o valor tabelado de $z_{\frac{\alpha}{2}} = 1,96$

$$P\left(\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}} \leq p \leq \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right) = 0,95$$
$$P\left(0,40 - 1,96\sqrt{\frac{0,40 \times 0,60}{30}} \leq p \leq 0,40 + 1,96\sqrt{\frac{0,40 \times 0,60}{30}}\right) = 0,95$$
$$P\left(0,40 - 0,08 \leq p \leq 0,40 + 0,80\right) = 0,95$$
$$P\left(0,32 \leq p \leq 0,48\right) = 0,95$$

Assim,

$$IC_{95\%}(\mu) = [0,32;\ 0,48]$$

### 2.2.2 Intervalo de Confiança para média $\mu$ com variância $\sigma^2$ conhecida

Como já vimos anteriormente, $\overline{X}$ (média amostral) tem distribuição normal de média $\mu$ e variância $\frac{\sigma^2}{n}$, assim um intervalo de $(1 - \alpha)$ de confiança para $\mu$ será dado por:

$$P\left(\overline{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Exemplo: Um pesquisador obteve a partir de uma amostra uma média $\overline{X} = 180cm$ para altura de uma determinado grupo de pessoas utilizando uma amostra n=40, sabe-se que a variância populacional da altura é de $\sigma^2 = 100cm^2$. Qual o intervalo de confiança a 90% e 95% para a média populacional.

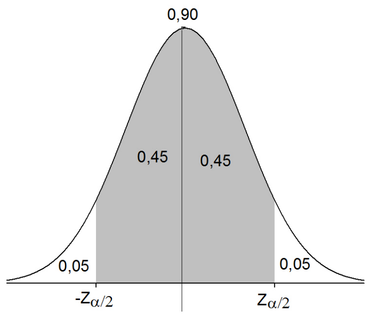
Primeiramente temos que obter o valor tabelado de Z, como queremos o intervalo de confiança a 90%, temos que:

$$1 - \alpha = 0,90 \Rightarrow \alpha = 1 - 0,90 = 0,10 \Rightarrow \frac{\alpha}{2} = \frac{0,10}{2} = 0,05$$

Assim, temos que procurar na tabela qual o valor de Z que deixa $0,05$ de probabilidade acima dele.

Olhando na tabela o valor em que $P(0 < Z < z) = 0,45$, temos que $z = 1,65$, logo o valor $Z_{\frac{\alpha}{2}} = 1,65$

$$P\left(\overline{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$
$$P\left(180 - 1,65\frac{\sqrt{200}}{\sqrt{40}} \leq \mu \leq 180 + 1,65\frac{\sqrt{200}}{\sqrt{40}}\right) = 0,90$$
$$P\left(176,31 \leq \mu \leq 183,69\right) = 0,90$$




ou seja, o intervalo de confiança a $90\%$ para a média é

$$IC_{90\%}(\mu) = [176,31;\ 183,69]$$

Fazendo o mesmo processo temos que a$95\%$:

$$1-\alpha = 0,95 \Rightarrow \alpha = 1-0,95 = 0,05 \Rightarrow \frac{\alpha}{2} = \frac{0,05}{2} = 0,025$$

Então $Z_{0,025} = 1,96$, assim

$$P\left(\overline{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$
$$P\left(180 - 1,96\frac{\sqrt{200}}{\sqrt{40}} \leq \mu \leq 180 + 1,96\frac{\sqrt{200}}{\sqrt{40}}\right) = 0,95$$
$$P\left(178,61 \leq \mu \leq 187,38\right) = 0,95$$

Assim,

$$IC_{95\%}(\mu) = [178,61;\ 187,38]$$

Observa-se que aumentando o nível de confiança, também temos o aumento do intervalo de confiança.

### 2.2.3 Intervalo de Confiança para média $\mu$ com variância $\sigma^2$ desconhecida

Na prática quando não se conhece a média $\overline{X}$ também não se conhece a variância, nesse caso utilizamos o intervalo de confiança:

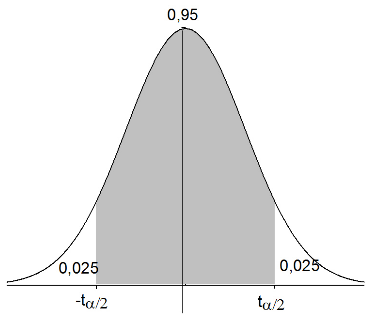
$$P\left(\overline{X} - t_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

Exemplo: Em uma determinada industria para verificar a qualidade dos rolamentos esféricos produzidos foi tomado uma amostra ao acaso um lote de 15 peças, fornecendo um diâmetro médio de 240cm com desvio padrão de 15cm . Encontre um intervalo de confiança de 95% para o diâmetro.

Primeiramente temos que obter o valor tabelado de t, como queremos o intervalo de confiança a $95\%$, temos que:

$$1-\alpha = 0,95 \Rightarrow \alpha = 1-0,95 = 0,05 \Rightarrow \frac{\alpha}{2} = \frac{0,05}{2} = 0,025$$

Olhando na tabela o valor que deixa 0,025 de área acima com $\nu = 15-1 = 14$, temos $t_{\frac{\alpha}{2}} = 2,145$




$$P\left(\overline{X} - t_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$
$$P\left(240 - 2,145\frac{15}{\sqrt{15}} \leq \mu \leq 180 + 2,145\frac{15}{\sqrt{15}}\right) = 0,95$$
$$P\left(231,69 \leq \mu \leq 248,31\right) = 0,95$$

Assim,

$$IC_{95\%}(\mu) = [231,69;\ 248,31]$$

### 2.2.4 Intervalo de Confiança para variância $\sigma^2$ e para o desvio padrão $\sigma$

Quando a população da qual foi amostra foi coletada for Normal, pode-se obter um intervalo de confiança para a variância $\sigma^2$ dada por:

$$P\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{\left(1-\frac{\alpha}{2}\right)}}\right) = 1 - \alpha$$

e IC para o desvio padrão é dado por

$$P\left(\sqrt{\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}}} \leq \sigma^2 \leq \sqrt{\frac{(n-1)S^2}{\chi_{\left(1-\frac{\alpha}{2}\right)}}} = 1 - \alpha\right)$$

Exemplo: No exemplo dos 15 peças de rolamentos esféricos, obter o intervalo de confiança de 95% para a variância e para o desvio padrão do maior eixo.

Temos que $\frac{0,05}{2} = 0,025$, nesse caso precisamos obter na tabela Qui-Quadrado o valores $\chi_{0,025}$ e $\chi_{1-0,025} = \chi_{0,975}$, com $\nu = 14$ graus de liberdade, então

$$\chi_{0,025} = 26,119 \quad \chi_{0,975} = 5,629$$

Nesse exemplo foi fornecido a variância amostral é $S^2 = 144$.

$$
\begin{aligned}
&P\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{\left(1-\frac{\alpha}{2}\right)}}\right) = 0,95\\
&P\left(\frac{14 \times 144}{26,119} \leq \sigma^2 \leq \frac{14 \times 144}{5,629}\right) = 0,95\\
&P\left(77,18 \leq \sigma^2 \leq 358,14\right) = 0,95
\end{aligned}
$$

A partir do intervalo da variância obtemo o IC do desvio padrão

$$
\begin{aligned}
&P\left(\sqrt{77,18} \leq \sigma \leq \sqrt{358,14}\right) = 0,95\\
&P\left(8,78 \leq \sigma \leq 18,92\right) = 0,95
\end{aligned}
$$

Assim,

$$IC_{95\%}(\sigma^2) = [77,18;\ 358,14] \quad IC_{95\%}(\sigma) = [8,78;\ 18,92]$$