ME414 - Estatística para Experimentalistas

Intervalos de Confiança

Prof. Carlos Trucíos
ctrucios@unicamp.br
ctruciosm.github.io

Instituto de Matemática, Estatística e Computação Científica,
Universidade Estadual de Campinas

Aula 12

Estimação por Intervalo: Introdução

Intervalo de confiança com variância conhecida

Intervalo de confiança com variância desconhecida

Intervalos de confiança para populações não normais

Tamanho da amostra

Estimação por Intervalo: Introdução

Estimação por Intervalo

- Até agora temos focado unicamente em estimadores pontuais (um único valor: média amostral, variância amostral, etc).

Estimação por Intervalo

- Até agora temos focado unicamente em estimadores pontuais (um único valor: média amostral, variância amostral, etc).
- Estimação pontual não permite medir a incerteza associada.

Estimação por Intervalo

- Até agora temos focado unicamente em estimadores pontuais (um único valor: média amostral, variância amostral, etc).
- Estimação pontual não permite medir a incerteza associada.
- Estimação pontual não permite saber qual é a possível magnitude de erro que estamos cometendo.

Estimação por Intervalo

- Até agora temos focado unicamente em estimadores pontuais (um único valor: média amostral, variância amostral, etc).
- Estimação pontual não permite medir a incerteza associada.
- Estimação pontual não permite saber qual é a possível magnitude de erro que estamos cometendo.

Estimação por Intervalo

- Até agora temos focado unicamente em estimadores pontuais (um único valor: média amostral, variância amostral, etc).
- Estimação pontual não permite medir a incerteza associada.
- Estimação pontual não permite saber qual é a possível magnitude de erro que estamos cometendo.

A forma geral de uma estimação por intervalo é:

$$\text{Estimação por ponto} \pm \text{margem de erro}$$

Estimação por Intervalo

Suponha que queremos estimar a média μ de uma distribuição qualquer com variância σ^2 (conhecida).

Estimação por Intervalo

Suponha que queremos estimar a média μ de uma distribuição qualquer com variância σ^2 (conhecida).

Pelo TCL temos:

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim_{approx} N(0,1)$$

Estimação por Intervalo

Suponha que queremos estimar a média μ de uma distribuição qualquer com variância σ^2 (conhecida).

Pelo TCL temos:

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim_{approx} N(0, 1)$$

Chamemos $e = |\bar{X} - \mu|$ e $\sigma_{\bar{X}} = \sigma/\sqrt{n}$.

Estimação por Intervalo

Suponha que queremos estimar a média μ de uma distribuição qualquer com variância σ^2 (conhecida).

Pelo TCL temos:

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim_{approx} N(0, 1)$$

Chamemos $e = |\bar{X} - \mu|$ e $\sigma_{\bar{X}} = \sigma/\sqrt{n}$.

Agora, podemos determinar a probabilidade de cometer erros de certa magnitude. Por exemplo, $P(e \leq 1.96\sigma_{\bar{X}})$

Estimação por Intervalo

$$P(e \leq 1.96\sigma_{\bar{X}}) = P(|\bar{X} - \mu| \leq 1.96\sigma_{\bar{X}}) = P(|Z| \leq 1.96)$$

$$P(-1.96 \leq Z \leq 1.96) = 0.95$$

```
pnorm(1.96)-pnorm(-1.96)
```

```
## [1] 0.9500042
```

Estimação por Intervalo

Por outro lado,

$$P(|\bar{X} - \mu| \leq 1.96\sigma_{\bar{X}}) = 0.95$$

Estimação por Intervalo

Por outro lado,

$$P(|\bar{X} - \mu| \leq 1.96\sigma_{\bar{X}}) = 0.95$$

$$P(-1.96\sigma_{\bar{X}} \leq \bar{X} - \mu \leq 1.96\sigma_{\bar{X}}) = 0.95$$

Estimação por Intervalo

Por outro lado,

$$P(|\bar{X} - \mu| \leq 1.96\sigma_{\bar{X}}) = 0.95$$

$$P(-1.96\sigma_{\bar{X}} \leq \bar{X} - \mu \leq 1.96\sigma_{\bar{X}}) = 0.95$$

$$P(\underbrace{\bar{X} - 1.96\sigma_{\bar{X}}}_{\text{limite inferior}} \leq \mu \leq \underbrace{\bar{X} + 1.96\sigma_{\bar{X}}}_{\text{limite superior}}) = 0.95$$

Estimação por Intervalo

Por outro lado,

$$P(|\bar{X} - \mu| \leq 1.96\sigma_{\bar{X}}) = 0.95$$

$$P(-1.96\sigma_{\bar{X}} \leq \bar{X} - \mu \leq 1.96\sigma_{\bar{X}}) = 0.95$$

$$P(\underbrace{\bar{X} - 1.96\sigma_{\bar{X}}}_{\text{limite inferior}} \leq \mu \leq \underbrace{\bar{X} + 1.96\sigma_{\bar{X}}}_{\text{limite superior}}) = 0.95$$

0.95 é denotado por δ ou $1 - \alpha$ e é chamado de coeficiente de confiança.

Estimação por Intervalo

O que significa esse intervalo?

Estimação por Intervalo

O que significa esse intervalo?

Para ilustrar,

Estimação por Intervalo

O que significa esse intervalo?

Para ilustrar,

- vamos a selecionar 100 amostras aleatórias de tamanho 50 de uma $N(0,1)$

Estimação por Intervalo

O que significa esse intervalo?

Para ilustrar,

- vamos a selecionar 100 amostras aleatórias de tamanho 50 de uma $N(0, 1)$
- Para cada amostra, calcularemos os intervalos de confiança 80%, 90%, 95%.

Estimação por Intervalo

O que significa esse intervalo?

Para ilustrar,

- vamos a selecionar 100 amostras aleatórias de tamanho 50 de uma $N(0, 1)$
- Para cada amostra, calcularemos os intervalos de confiança 80%, 90%, 95%.
- Vamos a contar quantas vezes o intervalo cobre o parâmetro.

Estimação por Intervalo

O que significa esse intervalo?

Para ilustrar,

- vamos a selecionar 100 amostras aleatórias de tamanho 50 de uma $N(0, 1)$
- Para cada amostra, calcularemos os intervalos de confiança 80%, 90%, 95%.
- Vamos a contar quantas vezes o intervalo cobre o parâmetro.

Estimação por Intervalo

O que significa esse intervalo?

Para ilustrar,

- vamos a selecionar 100 amostras aleatórias de tamanho 50 de uma $N(0, 1)$
- Para cada amostra, calcularemos os intervalos de confiança 80%, 90%, 95%.
- Vamos a contar quantas vezes o intervalo cobre o parâmetro.

  Os pontos vermelhos representam a estimação pontual e as linhas verdes representam o intervalo de confiança.

Estimação por Intervalo

Estimação por Intervalo

Tabela 1: Porcentagem de IC abaixo o IC, dentro do IC e acima do IC

Int. Conf.	Abaixo	Dentro	Acima
IC80%	8	78	14
IC90%	7	86	7
IC95%	3	93	4

Estimação por Intervalo

Tabela 1: Porcentagem de IC abaixo o IC, dentro do IC e acima do IC

Int. Conf.	Abaixo	Dentro	Acima
IC80%	8	78	14
IC90%	7	86	7
IC95%	3	93	4

O que significa o intervalo de confiança?

Voltando ao exemplo onde

$$P(\bar{X} - 1.96\sigma_{\bar{X}} \leq \mu \leq \bar{X} + 1.96\sigma_{\bar{X}}) = 0.95$$

Se pudessemos escolher várias amostras aleatórias de tamaho n e construir intervalos da forma $\langle \bar{X} - 1.96\sigma_{\bar{X}}; \bar{X} + 1.96\sigma_{\bar{X}} \rangle$, 95% deles conteriam o parâmetro (desconhecido) μ.

Intervalo de confiança com variância conhecida

Intervalo de confiança com variância conhecida

Intervalo de Confiança para μ: Caso σ conhecido

Seja $X_1, \ldots, X_n \sim N(\mu, \sigma)$, então:

Intervalo de confiança com variância conhecida

Intervalo de Confiança para μ: Caso σ conhecido

Seja $X_1, \ldots, X_n \sim N(\mu, \sigma)$, então:

- Se quisermos o IC 95% $\langle \bar{X} - 1.96\frac{\sigma}{\sqrt{n}}; \bar{X} + 1.96\frac{\sigma}{\sqrt{n}} \rangle$

Intervalo de confiança com variância conhecida

Intervalo de Confiança para μ: Caso σ conhecido

Seja $X_1, \ldots, X_n \sim N(\mu, \sigma)$, então:

- Se quisermos o IC 95% $\langle \bar{X} - 1.96\frac{\sigma}{\sqrt{n}}; \bar{X} + 1.96\frac{\sigma}{\sqrt{n}} \rangle$
- Se quisermos o IC 99% $\langle \bar{X} - 2.57\frac{\sigma}{\sqrt{n}}; \bar{X} + 2.57\frac{\sigma}{\sqrt{n}} \rangle$

Intervalo de confiança com variância conhecida

Intervalo de Confiança para μ: Caso σ conhecido

Seja $X_1, \ldots, X_n \sim N(\mu, \sigma)$, então:

- Se quisermos o IC 95% $\langle \bar{X} - 1.96\frac{\sigma}{\sqrt{n}}; \bar{X} + 1.96\frac{\sigma}{\sqrt{n}} \rangle$
- Se quisermos o IC 99% $\langle \bar{X} - 2.57\frac{\sigma}{\sqrt{n}}; \bar{X} + 2.57\frac{\sigma}{\sqrt{n}} \rangle$
- Em geral, se quisermos IC com nível de confiança $100\delta\%$,

$$\langle \bar{X} - Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \rangle,$$

em que $\alpha = 1 - \delta$ e $Z_{1-\alpha/2}$ é o quantil $1 - \alpha/2$ da distribuição $N(0, 1)$

Intervalo de confiança com variância conhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma = 0.5)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Intervalo de confiança com variância conhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma = 0.5)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$

Intervalo de confiança com variância conhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma = 0.5)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- Utilizaremos $\alpha = 1 - \delta = 0.1$ e 0.05

Intervalo de confiança com variância conhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma = 0.5)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- Utilizaremos $\alpha = 1 - \delta = 0.1$ e 0.05
- Para $\alpha = 0.1$, $Z_{1-\alpha/2} = Z_{1-0.1/2} = Z_{0.95}$ (usaremos para o IC90%)

Intervalo de confiança com variância conhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma = 0.5)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- Utilizaremos $\alpha = 1 - \delta = 0.1$ e 0.05
- Para $\alpha = 0.1$, $Z_{1-\alpha/2} = Z_{1-0.1/2} = Z_{0.95}$ (usaremos para o IC90%)
- Para $\alpha = 0.05$, $Z_{1-\alpha/2} = Z_{1-0.05/2} = Z_{0.975}$ (usaremos para o IC95%)

Intervalo de confiança com variância conhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma = 0.5)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- Utilizaremos $\alpha = 1 - \delta = 0.1$ e 0.05
- Para $\alpha = 0.1$, $Z_{1-\alpha/2} = Z_{1-0.1/2} = Z_{0.95}$ (usaremos para o IC90%)
- Para $\alpha = 0.05$, $Z_{1-\alpha/2} = Z_{1-0.05/2} = Z_{0.975}$ (usaremos para o IC95%)

Intervalo de confiança com variância conhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma = 0.5)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- Utilizaremos $\alpha = 1 - \delta = 0.1$ e 0.05
- Para $\alpha = 0.1$, $Z_{1-\alpha/2} = Z_{1-0.1/2} = Z_{0.95}$ (usaremos para o IC90%)
- Para $\alpha = 0.05$, $Z_{1-\alpha/2} = Z_{1-0.05/2} = Z_{0.975}$ (usaremos para o IC95%)

```
c(qnorm(0.95), qnorm(0.975))
```

```
## [1] 1.644854 1.959964
```

Intervalo de confiança com variância conhecida

$n = 8$, $\bar{X} = 3.0625$, $\sigma = 0.5$, $Z_{0.95} = 1.64$, $Z_{0.975} = 1.96$.

$$\langle \bar{X} - Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \rangle$$

Intervalo de confiança com variância conhecida

$$n = 8,\ \bar{X} = 3.0625,\ \sigma = 0.5,\ Z_{0.95} = 1.64,\ Z_{0.975} = 1.96.$$

$$\langle \bar{X} - Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \rangle$$

- IC 90%

$$\langle 3.0625 - 1.64\frac{0.5}{\sqrt{8}}; 3.0625 + 1.64\frac{0.5}{\sqrt{8}} \rangle \quad \equiv \langle 2.772586; 3.352414 \rangle$$

Intervalo de confiança com variância conhecida

$$n = 8,\ \bar{X} = 3.0625,\ \sigma = 0.5,\ Z_{0.95} = 1.64,\ Z_{0.975} = 1.96.$$

$$\langle \bar{X} - Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \rangle$$

- IC 90%

$$\langle 3.0625 - 1.64\frac{0.5}{\sqrt{8}}; 3.0625 + 1.64\frac{0.5}{\sqrt{8}} \rangle \quad \equiv \langle 2.772586; 3.352414 \rangle$$

- IC 95%

$$\langle 3.0625 - 1.96\frac{0.5}{\sqrt{8}}; 3.0625 + 1.96\frac{0.5}{\sqrt{8}} \rangle \quad \equiv \langle 2.716018; 3.408982 \rangle$$

Intervalo de confiança com variância desconhecida

Intervalo de confiança com variância desconhecida

- Os IC apresentados anteriormente são utilizados quando conhecemos o verdadeiro valor do parâmetro σ.

Intervalo de confiança com variância desconhecida

- Os IC apresentados anteriormente são utilizados quando conhecemos o verdadeiro valor do parâmetro σ.
- Contudo, na prática dificilmente conhecemos σ e temos que estimar esse valor por $\hat{\sigma}$

Intervalo de confiança com variância desconhecida

- Os IC apresentados anteriormente são utilizados quando conhecemos o verdadeiro valor do parâmetro σ.
- Contudo, na prática dificilmente conhecemos σ e temos que estimar esse valor por $\hat{\sigma}$
- Quando utilizamos $\hat{\sigma}$ em lugar de σ para construir o IC, devemos substituir a distribuição Normal pela distribuição T com $n-1$ graus de liberdade.

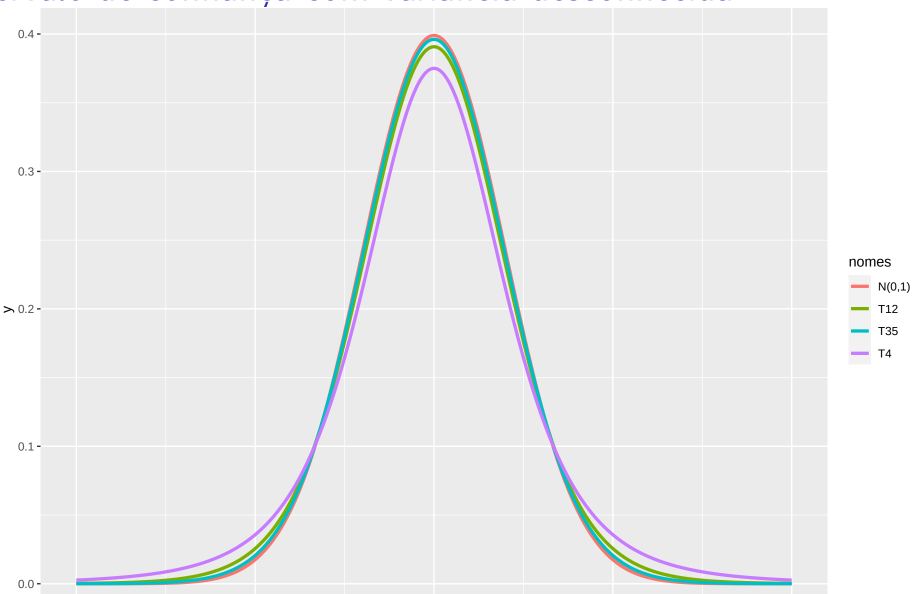
Intervalo de confiança com variância desconhecida

- Os IC apresentados anteriormente são utilizados quando conhecemos o verdadeiro valor do parâmetro σ.
- Contudo, na prática dificilmente conhecemos σ e temos que estimar esse valor por $\hat{\sigma}$
- Quando utilizamos $\hat{\sigma}$ em lugar de σ para construir o IC, devemos substituir a distribuição Normal pela distribuição T com $n-1$ graus de liberdade.
- Em geral, se quisermos IC com nível de confiança $100\delta\%$,

$$\langle \bar{X} - t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X} + t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}} \rangle,$$

em que $\alpha = 1 - \delta$ e $t_{1-\alpha/2,n-1}$ é o quantil $1-\alpha/2$ da distribuição T com $n-1$ graus de liberdade

Intervalo de confiança com variância desconhecida



Intervalo de confiança com variância desconhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Intervalo de confiança com variância desconhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$

Intervalo de confiança com variância desconhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- $\hat{\sigma} = 0.5125$

Intervalo de confiança com variância desconhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- $\hat{\sigma} = 0.5125$
- $\alpha = 1 - \delta = 0.1$ e 0.05

Intervalo de confiança com variância desconhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- $\hat{\sigma} = 0.5125$
- $\alpha = 1 - \delta = 0.1$ e 0.05
- $t_{1-\alpha/2,n-1} = t_{1-0.1/2,7} = t_{0.95,7}$ (usaremos para o IC90%)

Intervalo de confiança com variância desconhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- $\hat{\sigma} = 0.5125$
- $\alpha = 1 - \delta = 0.1$ e 0.05
- $t_{1-\alpha/2,n-1} = t_{1-0.1/2,7} = t_{0.95,7}$ (usaremos para o IC90%)
- $t_{1-\alpha/2,n-1} = t_{1-0.05/2,7} = t_{0.975,7}$ (usaremos para o IC95%)

Intervalo de confiança com variância desconhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- $\hat{\sigma} = 0.5125$
- $\alpha = 1 - \delta = 0.1$ e 0.05
- $t_{1-\alpha/2,n-1} = t_{1-0.1/2,7} = t_{0.95,7}$ (usaremos para o IC90%)
- $t_{1-\alpha/2,n-1} = t_{1-0.05/2,7} = t_{0.975,7}$ (usaremos para o IC95%)

Intervalo de confiança com variância desconhecida

Exemplo

Seja 3.1, 3.5, 2.6, 3.4, 3.8, 3, 2.9 e 2.2 uma amostra aleatória de tamaho $n = 8$ proveniente de uma distribuição $N(\mu, \sigma)$. Calcule:

- IC para μ com nível de confiança $\delta = 0.90$
- IC para μ com nível de confiança $\delta = 0.95$

Solução

- $\bar{X} = 3.0625$
- $\hat{\sigma} = 0.5125$
- $\alpha = 1 - \delta = 0.1$ e 0.05
- $t_{1-\alpha/2,n-1} = t_{1-0.1/2,7} = t_{0.95,7}$ (usaremos para o IC90%)
- $t_{1-\alpha/2,n-1} = t_{1-0.05/2,7} = t_{0.975,7}$ (usaremos para o IC95%)

```
c(qt(0.95,7), qt(0.975,7))
```

```
## [1] 1.894579 2.364624
```

Intervalo de confiança com variância desconhecida

$n = 8$, $\bar{X} = 3.0625$, $\hat{\sigma} = 0.5125$, $t_{0.95,7} = 1.89$, $t_{0.975,7} = 2.36$.

$$\langle \bar{X} - t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X} + t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

Intervalo de confiança com variância desconhecida

$$n = 8, \ \bar{X} = 3.0625, \ \hat{\sigma} = 0.5125, \ t_{0.95,7} = 1.89, \ t_{0.975,7} = 2.36.$$

$$\langle \bar{X} - t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X} + t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

- IC 90%

$$\langle 3.0625 - 1.89\frac{0.5125}{\sqrt{8}}; 3.0625 + 1.89\frac{0.5125}{\sqrt{8}} \rangle \quad \equiv \langle 2.720039; 3.404961 \rangle$$

Intervalo de confiança com variância desconhecida

$$n = 8,\ \bar{X} = 3.0625,\ \hat{\sigma} = 0.5125,\ t_{0.95,7} = 1.89,\ t_{0.975,7} = 2.36.$$

$$\langle \bar{X} - t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X} + t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

- IC 90%

$$\langle 3.0625 - 1.89\frac{0.5125}{\sqrt{8}}; 3.0625 + 1.89\frac{0.5125}{\sqrt{8}} \rangle \quad \equiv \langle 2.720039; 3.404961 \rangle$$

- IC 95%

$$\langle 3.0625 - 2.36\frac{0.5125}{\sqrt{8}}; 3.0625 + 2.36\frac{0.5125}{\sqrt{8}} \rangle \quad \equiv \langle 2.634877; 3.490123 \rangle$$

Intervalos de confiança para populações não normais

Intervalos de confiança para populações não normais

- Até agora temos visto como calcular intervalos de confiança quando nossa amostra aleatória $X_1, \ldots, X_n \sim N(\mu, \sigma)$

Intervalos de confiança para populações não normais

- Até agora temos visto como calcular intervalos de confiança quando nossa amostra aleatória $X_1, \ldots, X_n \sim N(\mu, \sigma)$
- Vimos o caso quendo σ é conhecido

$$\langle \bar{X} - Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \rangle$$

Intervalos de confiança para populações não normais

- Até agora temos visto como calcular intervalos de confiança quando nossa amostra aleatória $X_1, \ldots, X_n \sim N(\mu, \sigma)$
- Vimos o caso quendo σ é conhecido

$$\langle \bar{X} - Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \rangle$$

- Também vimos o caso quendo σ é desconhecido

$$\langle \bar{X} - t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X} + t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

Intervalos de confiança para populações não normais

- Até agora temos visto como calcular intervalos de confiança quando nossa amostra aleatória $X_1, \ldots, X_n \sim N(\mu, \sigma)$
- Vimos o caso quendo σ é conhecido

$$\langle \bar{X} - Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \rangle$$

- Também vimos o caso quendo σ é desconhecido

$$\langle \bar{X} - t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X} + t_{1-\alpha/2,n-1}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

- Mas o que acontece se $X_1, \ldots, X_n$ não seguirem uma distribuição Normal?

Intervalos de confiança para populações não normais

- Quando $X_1, \ldots, X_n$ não seguirem uma distribuição Normal podemos utilizar o TCL (desde que n seja grande).

Intervalos de confiança para populações não normais

- Quando $X_1, \ldots, X_n$ não seguirem uma distribuição Normal podemos utilizar o TCL (desde que n seja grande).
- Teremos que, independente da distribuição de $X_1, \ldots, X_n$, um intervalo aproximado, quando n for grande, será

$$\langle \bar{X} - Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

Intervalos de confiança para populações não normais

- Quando $X_1, \ldots, X_n$ não seguirem uma distribuição Normal podemos utilizar o TCL (desde que n seja grande).
- Teremos que, independente da distribuição de $X_1, \ldots, X_n$, um intervalo aproximado, quando n for grande, será

$$\langle \bar{X} - Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

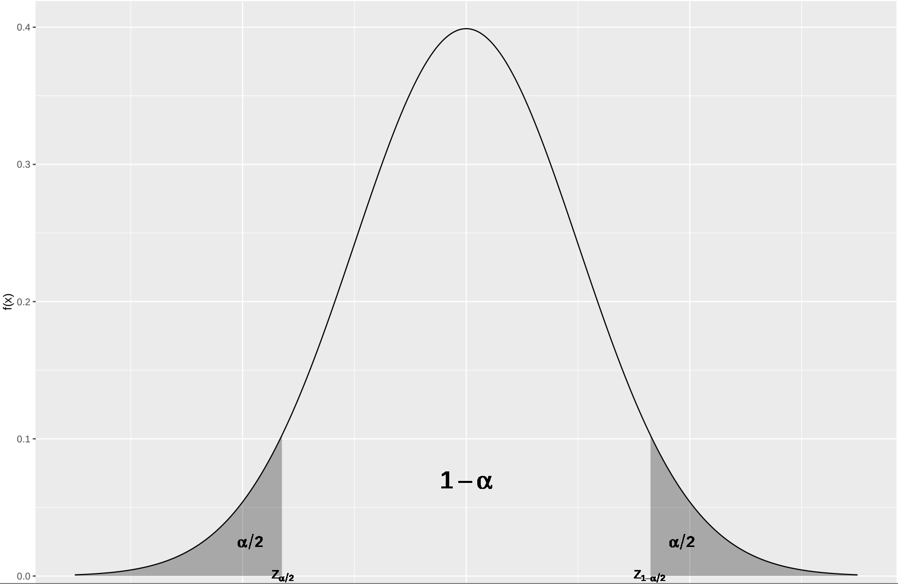
Intervalos de confiança para populações não normais

- Quando $X_1, \ldots, X_n$ não seguirem uma distribuição Normal podemos utilizar o TCL (desde que n seja grande).
- Teremos que, independente da distribuição de $X_1, \ldots, X_n$, um intervalo aproximado, quando n for grande, será

$$\langle \bar{X} - Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

O que são esses valores $Z_{1-\alpha/2}$?

Intervalos de confiança



Intervalos de confiança

Uma pesquisa baseada em uma amostra de tamanho 600, descubriu que as famílias pretendem gastar em média 649 reais nas festas de final do ano. Se o desvio padrão da amostra foi de 175 reais:

Intervalos de confiança

Uma pesquisa baseada em uma amostra de tamanho 600, descubriu que as famílias pretendem gastar em média 649 reais nas festas de final do ano. Se o desvio padrão da amostra foi de 175 reais:

- Com 95% de confiança, qual é a margem de erro?

Intervalos de confiança

Uma pesquisa baseada em uma amostra de tamanho 600, descubriu que as famílias pretendem gastar em média 649 reais nas festas de final do ano. Se o desvio padrão da amostra foi de 175 reais:

- Com 95% de confiança, qual é a margem de erro?
- Qual é o intervalo de confiança 95% para a média populacional μ?

Intervalos de confiança

Uma pesquisa baseada em uma amostra de tamanho 600, descubriu que as famílias pretendem gastar em média 649 reais nas festas de final do ano. Se o desvio padrão da amostra foi de 175 reais:

- Com 95% de confiança, qual é a margem de erro?
- Qual é o intervalo de confiança 95% para a média populacional μ?
- Qual é o intervalo de confiança 99% para a média populacional μ?

Intervalos de confiança

Uma pesquisa baseada em uma amostra de tamanho 600, descubriu que as famílias pretendem gastar em média 649 reais nas festas de final do ano. Se o desvio padrão da amostra foi de 175 reais:

- Com 95% de confiança, qual é a margem de erro?
- Qual é o intervalo de confiança 95% para a média populacional μ?
- Qual é o intervalo de confiança 99% para a média populacional μ?

Intervalos de confiança

Uma pesquisa baseada em uma amostra de tamanho 600, descubriu que as famílias pretendem gastar em média 649 reais nas festas de final do ano. Se o desvio padrão da amostra foi de 175 reais:

- Com 95% de confiança, qual é a margem de erro?
- Qual é o intervalo de confiança 95% para a média populacional μ?
- Qual é o intervalo de confiança 99% para a média populacional μ?

Solução

- $n = 600$ (grande), utilizamoso TCL

Intervalos de confiança

Uma pesquisa baseada em uma amostra de tamanho 600, descubriu que as famílias pretendem gastar em média 649 reais nas festas de final do ano. Se o desvio padrão da amostra foi de 175 reais:

- Com 95% de confiança, qual é a margem de erro?
- Qual é o intervalo de confiança 95% para a média populacional μ?
- Qual é o intervalo de confiança 99% para a média populacional μ?

Solução

- $n = 600$ (grande), utilizamoso TCL
- IC serão calculados utilizando:

$$\langle \bar{X} - \underbrace{Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}}_{\text{Margem de erro}} ; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

Intervalos de confiança

$$\langle \bar{X} - \underbrace{Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}}_{\text{Margem de erro}} ; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

Com 95% de confiança, qual é a margem de erro?

Intervalos de confiança

$$\langle \bar{X} - \underbrace{Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}}_{\text{Margem de erro}} ; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

Com 95% de confiança, qual é a margem de erro?

- $\delta = 0.95$, então $\alpha = 0.05$. Assim, $Z_{1-\alpha/2} = Z_{0.975} = 1.96$

Intervalos de confiança

$$\langle \bar{X} - \underbrace{Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}}_{\text{Margem de erro}} ; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

Com 95% de confiança, qual é a margem de erro?

- $\delta = 0.95$, então $\alpha = 0.05$. Assim, $Z_{1-\alpha/2} = Z_{0.975} = 1.96$
- Margem de erro: $Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} = 1.96\frac{175}{\sqrt{600}} = 14.00292 \approx 14$

Intervalos de confiança

$$\langle \bar{X} - \underbrace{Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}}_{\text{Margem de erro}} ; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

Com 95% de confiança, qual é a margem de erro?

- $\delta = 0.95$, então $\alpha = 0.05$. Assim, $Z_{1-\alpha/2} = Z_{0.975} = 1.96$
- Margem de erro: $Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} = 1.96\frac{175}{\sqrt{600}} = 14.00292 \approx 14$

Intervalos de confiança

$$\langle \bar{X} - \underbrace{Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}}_{\text{Margem de erro}} ; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

Com 95% de confiança, qual é a margem de erro?

- $\delta = 0.95$, então $\alpha = 0.05$. Assim, $Z_{1-\alpha/2} = Z_{0.975} = 1.96$
- Margem de erro: $Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} = 1.96\frac{175}{\sqrt{600}} = 14.00292 \approx 14$

Qual é o intervalo de confiança 95% para a média populacional μ?

Intervalos de confiança

$$\langle \bar{X} - \underbrace{Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}}_{\text{Margem de erro}} ; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle$$

Com 95% de confiança, qual é a margem de erro?

- $\delta = 0.95$, então $\alpha = 0.05$. Assim, $Z_{1-\alpha/2} = Z_{0.975} = 1.96$
- Margem de erro: $Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} = 1.96\frac{175}{\sqrt{600}} = 14.00292 \approx 14$

Qual é o intervalo de confiança 95% para a média populacional μ?

$$\langle \bar{X} - \underbrace{Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}}_{\text{Margem de erro}} ; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \rangle \approx \langle 649 - 14; 649 + 14 \rangle \approx \langle 635; 663 \rangle$$

Intervalos de confiança

Qual é o intervalo de confiança 99% para a média populacional μ?

Intervalos de confiança

Qual é o intervalo de confiança 99% para a média populacional μ?

- $\delta = 0.99$, então $\alpha = 0.01$. Assim, $Z_{1-\alpha/2} = Z_{0.995} = 2.57$

Intervalos de confiança

Qual é o intervalo de confiança 99% para a média populacional μ?

- $\delta = 0.99$, então $\alpha = 0.01$. Assim, $Z_{1-\alpha/2} = Z_{0.995} = 2.57$
- $\bar{X} = 649$, $\hat{\sigma} = 175$, $n = 600$

Intervalos de confiança

Qual é o intervalo de confiança 99% para a média populacional μ?

- $\delta = 0.99$, então $\alpha = 0.01$. Assim, $Z_{1-\alpha/2} = Z_{0.995} = 2.57$
- $\bar{X} = 649$, $\hat{\sigma} = 175$, $n = 600$

Intervalos de confiança

Qual é o intervalo de confiança 99% para a média populacional μ?

- $\delta = 0.99$, então $\alpha = 0.01$. Assim, $Z_{1-\alpha/2} = Z_{0.995} = 2.57$
- $\bar{X} = 649$, $\hat{\sigma} = 175$, $n = 600$

$$\left\langle \bar{X} - Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}; \bar{X} + Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \right\rangle =$$

$$\left\langle 649 - 2.57\frac{175}{\sqrt{600}}; 649 + 2.57\frac{175}{\sqrt{600}} \right\rangle =$$

$$\left\langle 630.639; 667.361 \right\rangle$$

Tamanho da amostra

Como determinar o tamanho da amostra

- A margem de erro $Z_{1-\alpha/2}\dfrac{\hat{\sigma}}{\sqrt{n}}$ depende to tamanho da amostra n

Como determinar o tamanho da amostra

- A margem de erro $Z_{1-\alpha/2}\dfrac{\hat{\sigma}}{\sqrt{n}}$ depende to tamanho da amostra n
- Qual deve ser o tamaho de n para produzir uma margem de erro desejada?

Como determinar o tamanho da amostra

- A margem de erro $Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$ depende to tamanho da amostra n
- Qual deve ser o tamaho de n para produzir uma margem de erro desejada?
- Suponha que conhecemos o valor σ, se denotarmor por E à margem de erro, temos:

$$E = Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Como determinar o tamanho da amostra

- A margem de erro $Z_{1-\alpha/2}\dfrac{\hat{\sigma}}{\sqrt{n}}$ depende to tamanho da amostra n
- Qual deve ser o tamaho de n para produzir uma margem de erro desejada?
- Suponha que conhecemos o valor σ, se denotarmor por E à margem de erro, temos:

$$E = Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

- Resolvendo:

$$\sqrt{n} = Z_{1-\alpha/2}\frac{\sigma}{E} \longrightarrow n = Z^2_{1-\alpha/2}\frac{\sigma^2}{E^2}$$

Como determinar o tamanho da amostra

- A margem de erro $Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$ depende to tamanho da amostra n
- Qual deve ser o tamaho de n para produzir uma margem de erro desejada?
- Suponha que conhecemos o valor σ, se denotarmor por E à margem de erro, temos:

$$E = Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

- Resolvendo:

$$\sqrt{n} = Z_{1-\alpha/2}\frac{\sigma}{E} \longrightarrow n = Z^2_{1-\alpha/2}\frac{\sigma^2}{E^2}$$

Como determinar o tamanho da amostra

- A margem de erro $Z_{1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$ depende to tamanho da amostra n
- Qual deve ser o tamaho de n para produzir uma margem de erro desejada?
- Suponha que conhecemos o valor σ, se denotarmor por E à margem de erro, temos:

$$E = Z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

- Resolvendo:

$$\sqrt{n} = Z_{1-\alpha/2}\frac{\sigma}{E} \longrightarrow n = Z^2_{1-\alpha/2}\frac{\sigma^2}{E^2}$$

Esse tamanho de amostra fornece a margem de erro desejada, ao nível de confiança escolhido.

Como determinar o tamanho da amostra

$$n = Z^2_{1-\alpha/2}\frac{\sigma^2}{E^2}$$

Como determinar o tamanho da amostra

$$n = Z^2_{1-\alpha/2}\frac{\sigma^2}{E^2}$$

- A formula para determinar o tamanho da amostra depende do parâmetro deconhecido σ, o que fazer na prática?

Como determinar o tamanho da amostra

$$n = Z^2_{1-\alpha/2}\frac{\sigma^2}{E^2}$$

- A formula para determinar o tamanho da amostra depende do parâmetro deconhecido σ, o que fazer na prática?

[1] Estudo piloto é um estudo preliminar com o objetivo de validar o instrumento que utilizaremos para coletar a informaçñao, bem como para termos uma estimativa de σ

Como determinar o tamanho da amostra

$$n = Z^2_{1-\alpha/2}\frac{\sigma^2}{E^2}$$

- A formula para determinar o tamanho da amostra depende do parâmetro deconhecido σ, o que fazer na prática?
- Use a estimativa do desvio padrão calculada a partir de dados de estudos anteriores.
- Use um estudo piloto[1] para selecionar uma amostra preliminar e estimar o valor de σ.

[1] Estudo piloto é um estudo preliminar com o objetivo de validar o instrumento que utilizaremos para coletar a informaçñao, bem como para termos uma estimativa de σ

Como determinar o tamanho da amostra

Uma organização deseja realizar um estudo para estimar a média populacional do custo diário de aluguel de carros de tamaho médio em Campinas. O diretor do projeto especifica que a média populacional deve ser estimada com uma margem de erro de 3 reais e um grau de confiança de 95%. De um estudo piloto, sabe-se que o desvio padrão do custo diário de aluguel é de 15 reais. Determine o tamanho amostral que satisfaça as exigências do diretor.

Como determinar o tamanho da amostra

Uma organização deseja realizar um estudo para estimar a média populacional do custo diário de aluguel de carros de tamaho médio em Campinas. O diretor do projeto especifica que a média populacional deve ser estimada com uma margem de erro de 3 reais e um grau de confiança de 95%. De um estudo piloto, sabe-se que o desvio padrão do custo diário de aluguel é de 15 reais. Determine o tamanho amostral que satisfaça as exigências do diretor.

- $E = 3$, $\hat{\sigma} = 15$

Como determinar o tamanho da amostra

Uma organização deseja realizar um estudo para estimar a média populacional do custo diário de aluguel de carros de tamaho médio em Campinas. O diretor do projeto especifica que a média populacional deve ser estimada com uma margem de erro de 3 reais e um grau de confiança de 95%. De um estudo piloto, sabe-se que o desvio padrão do custo diário de aluguel é de 15 reais. Determine o tamanho amostral que satisfaça as exigências do diretor.

- $E = 3$, $\hat{\sigma} = 15$
- Nível de confiança 95%, ou seja $\delta = 0.95$ e $\alpha = 0.05 = 1 - \delta$

Como determinar o tamanho da amostra

Uma organização deseja realizar um estudo para estimar a média populacional do custo diário de aluguel de carros de tamaho médio em Campinas. O diretor do projeto especifica que a média populacional deve ser estimada com uma margem de erro de 3 reais e um grau de confiança de 95%. De um estudo piloto, sabe-se que o desvio padrão do custo diário de aluguel é de 15 reais. Determine o tamanho amostral que satisfaça as exigências do diretor.

- $E = 3$, $\hat{\sigma} = 15$
- Nível de confiança 95%, ou seja $\delta = 0.95$ e $\alpha = 0.05 = 1 - \delta$
- $Z_{1-\alpha/2} = Z_{0.975} = 1.96$

Como determinar o tamanho da amostra

Uma organização deseja realizar um estudo para estimar a média populacional do custo diário de aluguel de carros de tamaho médio em Campinas. O diretor do projeto especifica que a média populacional deve ser estimada com uma margem de erro de 3 reais e um grau de confiança de 95%. De um estudo piloto, sabe-se que o desvio padrão do custo diário de aluguel é de 15 reais. Determine o tamanho amostral que satisfaça as exigências do diretor.

- $E = 3$, $\hat{\sigma} = 15$
- Nível de confiança 95%, ou seja $\delta = 0.95$ e $\alpha = 0.05 = 1 - \delta$
- $Z_{1-\alpha/2} = Z_{0.975} = 1.96$
- Utilizando a formula:

$$n = Z^2_{1-\alpha/2}\frac{\sigma^2}{E^2} = (1.96)^2\frac{15^2}{3^2} = 216.09 \approx 217$$

Como determinar o tamanho da amostra

Uma organização deseja realizar um estudo para estimar a média populacional do custo diário de aluguel de carros de tamaho médio em Campinas. O diretor do projeto especifica que a média populacional deve ser estimada com uma margem de erro de 3 reais e um grau de confiança de 95%. De um estudo piloto, sabe-se que o desvio padrão do custo diário de aluguel é de 15 reais. Determine o tamanho amostral que satisfaça as exigências do diretor.

- $E = 3$, $\hat{\sigma} = 15$
- Nível de confiança 95%, ou seja $\delta = 0.95$ e $\alpha = 0.05 = 1 - \delta$
- $Z_{1-\alpha/2} = Z_{0.975} = 1.96$
- Utilizando a formula:

$$n = Z^2_{1-\alpha/2}\frac{\sigma^2}{E^2} = (1.96)^2\frac{15^2}{3^2} = 216.09 \approx 217$$

Como determinar o tamanho da amostra

Uma organização deseja realizar um estudo para estimar a média populacional do custo diário de aluguel de carros de tamaho médio em Campinas. O diretor do projeto especifica que a média populacional deve ser estimada com uma margem de erro de 3 reais e um grau de confiança de 95%. De um estudo piloto, sabe-se que o desvio padrão do custo diário de aluguel é de 15 reais. Determine o tamanho amostral que satisfaça as exigências do diretor.

- $E = 3$, $\hat{\sigma} = 15$
- Nível de confiança 95%, ou seja $\delta = 0.95$ e $\alpha = 0.05 = 1 - \delta$
- $Z_{1-\alpha/2} = Z_{0.975} = 1.96$
- Utilizando a formula:

$$n = Z^2_{1-\alpha/2}\frac{\sigma^2}{E^2} = (1.96)^2\frac{15^2}{3^2} = 216.09 \approx 217$$

(arredondamos pra cima).

Leituras recomendadas

- Anderson, D. R; Sweeney, D. J.; e Williams, T. A. (2008). *Estatística Aplicada à Administração e Economia*. 2ed. Cengage Learning. **Cap 8**
- Morettin, P.A; e Bussab, W. de O. (2004). *Estatística Básica*. 5ed, Saraiva. **Cap 11**