Capítulo 6 - Estimação por intervalo de confiança

Estimação por intervalo de confiança

Seja $(X_1, X_2, \ldots, X_n)$ uma amostra aleatória de uma população com função de distribuição $F(\theta)$ onde θ é desconhecido.

Uma **estimativa por intervalo de confiança** de um parâmetro θ desconhecido, é um intervalo da forma $l < \theta < u$, onde os extremos l e u dependem do valor que o estimador $\hat{\Theta}$ do parâmetro θ, assumir para uma dada concretização $x_1, \ldots, x_n$ da amostra aleatória.

Como diferentes concretizações produzirão estimativas pontuais $\hat{\theta}$ distintas e, consequentemente diferentes extremos l e u, estes extremos são observações de variáveis aleatórias L e U, respectivamente.

A partir da distribuição de amostragem de $\hat{\Theta}$, poderemos encontrar extremos L e U tais que

$$P(L < \theta < U) = 1 - \alpha$$

com $0 < \alpha < 1$.

Estimação por intervalo de confiança

- O intervalo resultante $l < \theta < u$, é designado por **intervalo de confiança** $(1-\alpha)$ para o parâmetro θ.

- As quantidades l e u são denominadas **limites de confiança inferior** e **superior**, respectivamente

- $(1-\alpha)$ é chamado **coeficiente de confiança** do intervalo
  Normalmente são usados coeficientes de confiança superiores a 90%.

Iremos determinar ICs para os nossos parâmetros populacionais de interesse, i.e., para μ**,** σ^2**,** σ **e** p (respectivamente, média, variância, desvio padrão e proporção populacionais)

Estimação por intervalo de confiança

Definição (Intervalo Aleatório de Confiança)

Seja $(X_1, X_2, \ldots, X_n)$ *uma amostra aleatória de uma população com função de distribuição* $F(\theta)$, θ *desconhecido. Considere as estatísticas*

$$T_1(X_1, X_2, \ldots, X_n) \qquad \text{e} \qquad T_2(X_1, X_2, \ldots, X_n),$$

tais que $P(T_1 < \theta < T_2) = 1 - \alpha$, *onde* $\alpha \in]0, 1[$ *não depende de* θ. *Então* $]T_1, T_2[$ *é um* **intervalo aleatório de confiança** *para* θ.

Estimação por intervalo de confiança

Definição (Intervalo de Confiança)

Seja $(x_1, x_2, \ldots, x_n)$ *uma realização da amostra aleatória e sejam*

$$t_1 = T_1(x_1, x_2, \ldots, x_n) \qquad \text{e} \qquad t_2 = T_2(x_1, x_2, \ldots, x_n),$$

os valores das estatísticas T_1 *e* T_2 *.*

Ao intervalo $]t_1, t_2[$ *chamamos* **intervalo de confiança** $(1-\alpha) \times 100\%$ *para* θ *onde o valor* $(1-\alpha)$ *representa o* **nível** *(ou coeficiente)* **de confiança** *do intervalo e* α *o* **nível de significância***.*

Escrevemos habitualmente

$$IC_{(1-\alpha)100\%}(\theta) =]t_1, t_2[.$$

Note que não podemos afirmar que o valor populacional θ pertence ao intervalo de confiança deduzido. O que se garante com esta construção é apenas que, em m concretizações da a.a., digamos

$$x_1^1, \ldots, x_n^1, \quad \ldots \quad , x_1^m, \ldots, x_n^m \qquad (m \text{ grande})$$

e m ICs $]t_1^1, t_2^1[, \ldots,]t_1^m, t_2^m[$ calculados com base em cada uma destas concretizações, então θ irá pertencer a $(1-\alpha)100\%$ desses desses m intervalos.

Considere-se a amostra aleatória $(X_1, \ldots, X_n)$ de uma população $X \sim N(\mu, \sigma^2)$ de variância σ^2 conhecida e média $\underline{\mu \text{ desconhecida}}$.

No sentido de construir um intervalo aleatório de confiança (IAC) $(1-\alpha) \times 100\%$ para o parâmetro desconhecido μ, consideramos o estimador pontual de μ, $\bar{X}$ (que sabemos ser centrado e consistente em média quadrática), e ainda a **variável pivot**

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1).$$

A ideia da construção do IAC para μ passa por considerar que

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$$

e deduzir a partir daí o IAC de interesse.

Note que no gráfico ao lado, se tem que z_a é o valor real tal que $P(Z > z_a) = a$.

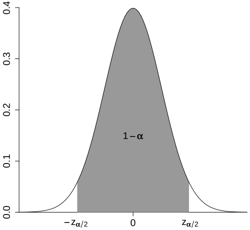


Figura: Distribuição da variável pivot Z.

O valor $z_{\alpha/2}$ é obtido através da resolução da equação:

$$P\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) = 1-\alpha \Leftrightarrow P(Z < z_{\alpha/2}) - P(Z \leq -z_{\alpha/2}) = 1-\alpha$$

$$\Leftrightarrow \Phi(z_{\alpha/2}) - \Phi(-z_{\alpha/2}) = 1-\alpha \Leftrightarrow \Phi(z_{\alpha/2}) = 1-\alpha/2$$

$$\Leftrightarrow z_{\alpha/2} = \Phi^{-1}\left(1-\alpha/2\right)$$

Determinação dos extremos do intervalo aleatório:

$$-z_{\alpha/2} < Z < z_{\alpha/2} \Leftrightarrow -z_{\alpha/2} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}$$

$$\Leftrightarrow -z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}} < \overline{X}-\mu < z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}}$$

$$\Leftrightarrow -z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}} - \overline{X} < -\mu < z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}} - \overline{X}$$

$$\Leftrightarrow \overline{X} - z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}}$$

Logo, o intervalo aleatório de confiança $(1-\alpha)\times 100\%$ para μ é:

$$IAC_{(1-\alpha)\times 100\%}(\mu) = \left]\bar{X} - z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}}\ ;\ \bar{X} + z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}}\right[$$

Assim, observando-se uma concretização $(x_1, x_2, \ldots, x_n)$ da amostra aleatória, o intervalo de confiança $(1-\alpha)\times 100\%$ para μ vem

$$IC_{(1-\alpha)\times 100\%}(\mu) = \left]\overline{x} - z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}}\ ;\ \overline{x} + z_{\alpha/2}\ \frac{\sigma}{\sqrt{n}}\right[.$$

Exemplo

Consideramos a população do peso das formigas Solenopsis, medido em décimas de grama, que sabemos ter distribuição Normal com média μ e variância $\sigma^2 = 2^2$, $X \sim N(\mu, 2^2)$. Desta população observámos a amostra de 4 pesos, $(8, 13, 9, 8.5)$, a qual usámos para obter uma estimativa de μ, $\bar{x} = 9.625dg$. Queremos agora determinar limites inferior e superior de um intervalo de confiança a 95% para μ.

Uma vez que a população segue uma **distribuição Normal** de valor médio μ desconhecido e **variância σ^2 conhecida**, o IC a $(1-\alpha)100\%$ será dado como indicado no slide acima.

Observamos do enunciado:

- informação populacional: $\sigma^2 = 4$
- informação amostral: $n = 4$; $\bar{x} = 9.625$
- nível de confiança para o IC: $1 - \alpha = 0.95 \Leftrightarrow \alpha = 0.05 \Leftrightarrow \frac{\alpha}{2} = 0.025$ pelo que $z_{\frac{\alpha}{2}} = z_{0.025} = q_{0.975} \underset{tabela}{\simeq} 1.96$ vindo portanto $-z_{\frac{\alpha}{2}} = -z_{0.025} \simeq -1.96$

e portanto o IC a 95% para μ vem finalmente

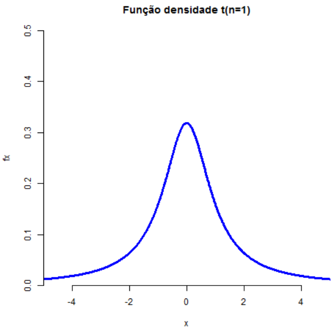
$$IC_{95\%}(\mu) \simeq \left]9.625 - 1.96\ \frac{2}{\sqrt{4}}\ ;\ 9.625 + 1.96\ \frac{2}{\sqrt{4}}\right[\simeq \left]7.665, 11.585\right[.$$

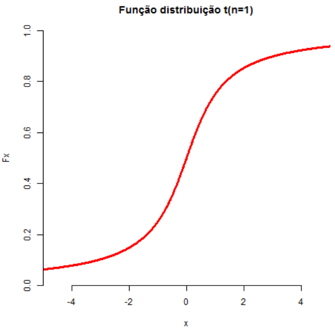
Distribuição t-Student

Definição (Distribuição t de Student)

Uma v.a. T diz-se ter distribuição t de Student com n graus de liberdade, e escreve-se $T \sim t_n$, se a sua função densidade probabilidade é dada por:

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{t^2}{n}\right)^{-\frac{(n+1)}{2}}, \qquad t \in \mathbb{R}, \quad \Gamma(a) = \int_0^{\infty} x^{a-1}e^{-x}dx,\ a > 0$$





Intervalos de Confiança para a média populacional, μ

População Normal com variância σ^2 desconhecida

Considere-se a amostra aleatória $(X_1, \ldots, X_n)$ de uma população $X \sim N(\mu, \sigma^2)$ de variância σ^2 e média μ desconhecidas.

No sentido de construir um intervalo aleatório de confiança (IAC) $(1-\alpha) \times 100\%$ para o parâmetro desconhecido μ, consideramos o estimador pontual de μ, $\bar{X}$ (que sabemos ser centrado e consistente em média quadrática), o estimador pontual de σ^2, S^2 (que sabemos ser centrado e consistente em média quadrática[1]) e ainda a **variável pivot**

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

A ideia da construção do IAC para μ passa por considerar que

$$P(-t_{n-1,\alpha/2} < T < t_{n-1,\alpha/2}) = 1 - \alpha$$

e deduzir a partir daí o IAC de interesse.

Note que no gráfico ao lado, se tem que $t_{n-1,a}$ é o valor real tal que

$$P(T > t_{n-1,a}) = a.$$

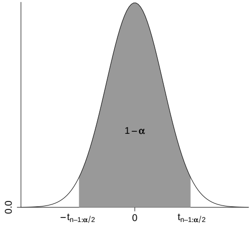


Figura: Distribuição da variável pivot T.

1

O valor $t_{n-1,\alpha/2}$ é obtido através da resolução da equação:

$$P\left(T > t_{n-1,\alpha/2}\right) = \alpha/2 \Leftrightarrow P(T \leq t_{n-1,\alpha/2}) = 1 - \alpha/2$$

$$\Leftrightarrow F_{t_{n-1}}(t_{n-1,\alpha/2}) = 1 - \alpha/2 \Leftrightarrow t_{n-1,\alpha/2} = F^{-1}_{t_{n-1}}\,(1 - \alpha/2)$$

Determinação dos extremos do intervalo aleatório:

$$-t_{n-1,\alpha/2} < T < t_{n-1,\alpha/2} \Leftrightarrow -t_{n-1,\alpha/2} < \frac{\overline{X} - \mu}{S/\sqrt{n}} < t_{n-1,\alpha/2}$$

$$\Leftrightarrow -t_{n-1,\alpha/2}\ \frac{S}{\sqrt{n}} < \overline{X} - \mu < t_{n-1,\alpha/2}\ \frac{S}{\sqrt{n}}$$

$$\Leftrightarrow -t_{n-1,\alpha/2}\ \frac{S}{\sqrt{n}} - \overline{X} < -\mu < t_{n-1,\alpha/2}\ \frac{S}{\sqrt{n}} - \overline{X}$$

$$\Leftrightarrow \overline{X} - t_{n-1,\alpha/2}\ \frac{S}{\sqrt{n}} < \mu < \overline{X} + t_{n-1,\alpha/2}\ \frac{S}{\sqrt{n}}$$

Logo, o intervalo aleatório de confiança $(1-\alpha)\times 100\%$ para μ é:

$$IAC_{(1-\alpha)\times 100\%}(\mu) = \left]\bar{X} - t_{n-1,\alpha/2}\ \frac{S}{\sqrt{n}}\ ;\ \bar{X} + t_{n-1,\alpha/2}\ \frac{S}{\sqrt{n}}\right[$$

Assim, observando-se uma concretização $(x_1, x_2, \ldots, x_n)$ da amostra aleatória, o intervalo de confiança $(1-\alpha)\times 100\%$ para μ vem

$$IC_{(1-\alpha)\times 100\%}(\mu) = \left]\overline{x} - t_{n-1,\alpha/2}\ \frac{s}{\sqrt{n}}\ ;\ \overline{x} + t_{n-1,\alpha/2}\ \frac{s}{\sqrt{n}}\right[.$$

Exemplo

Consideramos a população X do peso das formigas Solenopsis, medido em décimas de grama, que sabemos ter distribuição $N(\mu, \sigma^2)$ com média μ e variância σ^2 desconhecidas. Desta população observámos a amostra de 4 pesos, $(8, 13, 9, 8.5)$, a qual usámos para obter uma estimativa de μ, $\bar{x} = 9.625dg$. Queremos determinar um IC a 95% para μ.

Uma vez que a população segue uma **distribuição Normal** de valor médio μ e variância σ^2 **desconhecidas**, o IC a $(1-\alpha)100\%$ será dado como indicado no slide acima. Observamos:

- <u>informação populacional</u>: $X \sim N(\mu, \sigma^2)$ com μ, σ^2 desconhecidos
- <u>informação amostral</u>: $n = 4$; $\bar{x} = 9.625$; $s^2 = \dfrac{1}{3}\displaystyle\sum_{i=1}^{4}(x_i - \bar{x})^2 \simeq 5.229$
- <u>nível de confiança para o IC</u>: $1 - \alpha = 0.95 \Leftrightarrow \alpha = 0.05 \Leftrightarrow \frac{\alpha}{2} = 0.025$ pelo que $t_{3,\frac{\alpha}{2}} = t_{3,0.025} = \underset{\substack{tabela \\ t-student}}{\simeq} 3.182$ vindo portanto $-t_{3,\frac{\alpha}{2}} = -t_{3,0.025} \simeq -3.182$

e portanto o IC a 95% para μ vem finalmente

$$IC_{95\%}(\mu) \simeq \left]9.625 - 3.182\,\frac{\sqrt{5.229}}{\sqrt{4}}\ ;\ 9.625 + 3.182\,\frac{\sqrt{5.229}}{\sqrt{4}}\right[\simeq \left]5.987, 13.263\right[.$$

Intervalos de Confiança para a média populacional, μ

A tabela abaixo faz um sumário das variáveis pivot a usar na construção de intervalos de confiança para μ no caso em que:

- a distribuição da população é Normal
- a distribuição da população é desconhecida
- a distribuição é conhecida mas não é Normal.

distinguindo os casos em que a variância σ^2 é e não é conhecida.

População	Variância	Variável Pivot
Pop. Normal de média μ	σ^2, conhecida	$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
	σ^2, desconhecida	$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$
Pop. desconhecida de média μ $(n \geq 30)$	σ^2, conhecida	$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$
	σ^2, desconhecida	$\frac{\overline{X}-\mu}{S/\sqrt{n}} \overset{a}{\sim} N(0,1)$

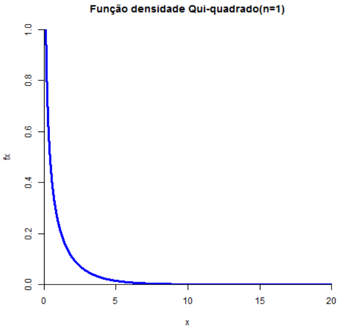
Observamos que para os dois últimos casos, a distribuição das variáveis pivot é assimptótica resultando da aplicação do TLC.

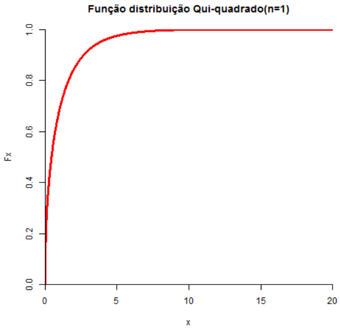
Distribuição Qui-quadrado

Definição (Distribuição do Qui Quadrado)

Uma variável aleatória X diz-se seguir uma distribuição Qui-quadrado com n graus de liberdade, e escrevemos $X \sim \chi^2_n$, se a sua função densidade probabilidade for dada por:

$$f(x) = \begin{cases} \frac{1}{\Gamma(n/2)2^{n/2}} e^{-x/2} x^{n/2-1}, & x > 0 \\ 0 \ , & x \leq 0 \end{cases} , \quad \Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx, \ a > 0$$





Intervalo de confiança para a variância populacional, σ^2

População Normal com média μ conhecida ou desconhecida

Considere-se a amostra aleatória $(X_1, \ldots, X_n)$ de uma população $X \sim N(\mu, \sigma^2)$ com <u>variância σ^2 desconhecida</u>.

No sentido de construir um intervalo aleatório de confiança (IAC) $(1-\alpha) \times 100\%$ para o parâmetro desconhecido σ^2, consideramos o estimador pontual de σ^2, S^2 (que sabemos ser centrado e consistente em média quadrática para σ^2), e ainda a **variável pivot**

$$X^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

A ideia da construção do IAC para σ^2 passa por considerar que

$$P(\chi^2_{n-1,1-\frac{\alpha}{2}} < X^2 < \chi^2_{n-1,\frac{\alpha}{2}}) = 1 - \alpha$$

e deduzir a partir daí o IAC de interesse.

Note que no gráfico ao lado, se tem que $\chi^2_{n-1,a}$ é o valor real tal que

$$P(X^2 > \chi^2_{n-1,a}) = a.$$

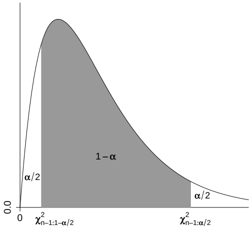


Figura: Distribuição da variável pivot X^2.

Os valores dos quantis são obtidos através das resoluções das equações:

$$P\left(X^2 > \chi^2_{n-1,\alpha/2}\right) = \alpha/2 \Leftrightarrow \chi^2_{n-1,\alpha/2} = F^{-1}_{\chi^2_{n-1}}\left(1-\alpha/2\right)$$

$$P\left(X^2 > \chi^2_{n-1,1-\alpha/2}\right) = 1-\alpha/2 \Leftrightarrow \chi^2_{n-1,1-\alpha/2} = F^{-1}_{\chi^2_{n-1}}\left(\alpha/2\right)$$

Determinação dos extremos do intervalo aleatório:

$$\chi^2_{n-1,1-\alpha/2} < X^2 < \chi^2_{n-1,\alpha/2} \Leftrightarrow \frac{\chi^2_{n-1,1-\alpha/2}}{(n-1)S^2} < \frac{1}{\sigma^2} < \frac{\chi^2_{n-1,\alpha/2}}{(n-1)S^2}$$

$$\Leftrightarrow \frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}}$$

Assim,

$$IC_{(1-\alpha)\times 100\%}(\sigma^2) = \left] \frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}} \; ; \; \frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}} \right[$$

Do resultado atrás obtido, muito simplesmente se constrói o intervalo de confiança para o **desvio padrão** populacional σ, resultando em:

$$IC_{(1-\alpha)\times 100\%}(\sigma) = \left] \sqrt{\frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}}} \; ; \; \sqrt{\frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}}} \right[$$

Intervalo de confiança para a variância populacional, σ^2

População Normal com média μ conhecida ou desconhecida

Exemplo

Consideramos a população X do peso das formigas Solenopsis, medido em décimas de grama, que sabemos ter distribuição $N(\mu, \sigma^2)$ com média μ e variância σ^2 desconhecidas. Desta população observámos a amostra de 4 pesos, $(8, 13, 9, 8.5)$, a qual usámos para obter uma estimativa de μ, $\bar{x} = 9.625dg$. Queremos determinar um IC a 95% para σ^2.

Uma vez que a população segue uma **distribuição Normal** de valor médio μ e variância σ^2 **desconhecidas**, o IC a $(1-\alpha)100\%$ para σ^2 será dado como indicado no slide acima. Tem-se

- informação populacional: $X \sim N(\mu, \sigma^2)$ com μ, σ^2 desconhecidos
- informação amostral: $n = 4$; $\bar{x} = 9.625$; $s^2 = \frac{1}{3}\sum_{i=1}^{4}(x_i - \bar{x})^2 \simeq 5.229$
- nível de confiança para o IC: $1 - \alpha = 0.95 \Leftrightarrow \alpha = 0.05 \Leftrightarrow \frac{\alpha}{2} = 0.025$ pelo que

  $\chi^2_{3,\frac{\alpha}{2}} = \chi^2_{3,0.025} \underset{tabela\ \chi^2}{\simeq} 9.348$ e $\chi^2_{3,1-\frac{\alpha}{2}} = \chi^2_{3,0.975} \underset{tabela\ \chi^2}{\simeq} 0.216$

e portanto o IC a 95% para σ^2 vem finalmente

$$IC_{95\%}(\sigma^2) \simeq \left]\frac{3 \times 5.229}{9.348}, \frac{3 \times 5.229}{0.216}\right[\simeq \left]1.678, 72.625\right[.$$

Exercícios

Intervalo de confiança para a proporção populacional, p

Assuma-se que os elementos de determinada população possuem uma dada característica, com uma certa probabilidade p desconhecida, independentemente uns dos outros.

Suponhamos que se selecciona uma amostra aleatória de n elementos desta população. Se X denotar o número desses elementos que possuem a referida característica, sabemos que $X \sim Bin(n, p)$.

Adicionalmente, se o tamanho amostral n for suficientemente grande, o Teorema Limite Central justifica que

$$Z = \frac{X - E(X)}{\sqrt{V(X)}} = \frac{X - np}{\sqrt{np(1-p)}} \overset{a}{\sim} N(0,1)$$

Notamos ainda que

- o estimador $\hat{P} = \frac{X}{n}$, dito **proporção amostral**, é um estimator centrado e consistente em média quadrática de p
- a variável pivot
$$Z = \frac{\hat{P} - p}{\sqrt{p(1-p)/n}} \overset{a}{\sim} N(0,1)$$

(deduz-se do resultado imediatamente acima - **TPC**)

Intervalo de confiança para a proporção populacional, p

No sentido de construir um intervalo aleatório de confiança (IAC) $(1-\alpha)\times 100\%$ para o parâmetro desconhecido p, consideramos o estimador pontual de p, $\hat{P}$ (que sabemos ser centrado e consistente em média quadrática para p), e ainda a **variável pivot**

$$Z = \frac{\hat{P} - p}{\sqrt{p(1-p)/n}} \overset{a}{\sim} N(0,1).$$

A ideia da construção do IAC para p passa então por considerar que $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$ e deduzir a partir daí o IAC de interesse.

Isto é feito como na dedução dos intervalos de confiança anteriores mas com a dificuldade adicional de que neste caso o parâmetro desconhecido p aparece agora também no denominador!

No sentido de simplificar o problema, considera-se habitualmente o denominador aproximado por $\sqrt{\hat{P}(1-\hat{P})/n}$, i.e, tomamos a variável pivot

$$Z \simeq \frac{\hat{P} - p}{\sqrt{\hat{P}(1-\hat{P})/n}} \overset{a}{\sim} N(0,1).$$

vindo então o IAC dado por **TPC**

$$IAC_{(1-\alpha)\times 100\%}(p) \underset{a}{\simeq} \left] \hat{P} - z_{\alpha/2}\sqrt{\hat{P}(1-\hat{P})/n}\ ;\ \hat{P} + z_{\alpha/2}\sqrt{\hat{P}(1-\hat{P})/n} \right[$$

Intervalo de confiança para p, Exemplo

Exemplo

De 200 casos de pessoas com cancro do cólon, aleatoriamente detectadas, 12 morreram após 5 anos da detecção.

1. *Estime pontualmente a probabilidade de uma pessoa que contraia o cancro do cólon morrer após 5 anos da sua detecção.*
2. *Estime* p *por um intervalo de confiança a* 90% *e determine a sua amplitude.*
3. *Quanto deveria aumentar ao tamanho da sua amostra aleatória de forma a que a largura do intervalo de confiança a 90% para a probabilidade considerada na alínea anterior fosse inferior a 0.01?*

(1) Observamos então:

- informação populacional:
  $X =$ nº pessoas que morreram de cancro do cólon em n casos,
  sendo p a probabilidade de morrer (sucesso) $\sim Bin(n, p)$
- informação amostral: $n = 200$; $x = 12$;

pelo que uma estimativa pontual de p é dada por $\hat{p} = \frac{12}{200} = 0.06$.

Intervalo de confiança para p, Exemplo

(2) Relativamente ao IC pretendido tem-se

- nível de confiança para o IC: $1-\alpha=0.9 \Leftrightarrow \alpha=0.1 \Leftrightarrow \frac{\alpha}{2}=0.05$ pelo que $z_{\frac{\alpha}{2}}=z_{0.05} \underset{tabela}{=\simeq} 1.645$ e $-z_{\frac{\alpha}{2}}=-z_{0.05} \underset{tabela}{=\simeq} -1.645$

e portanto o IC a 90% para p vem finalmente

$$IC_{90\%}(p) \underset{a}{\simeq} \left]0.06-1.645\sqrt{0.06\times 0.94/200}, 0.06+1.645\sqrt{0.06\times 0.94/200}\right[\simeq \left]0.032, 0.088\right[$$

sendo a sua amplitude dada por $\mathcal{A} \simeq 0.088-0.032=0.056$.

(3) Seguidamente observamos que a amplitude do $IAC_{(1-\alpha)100\%(p)}$ é dada por

$$\mathcal{A}=2\times z_{\frac{\alpha}{2}}\sqrt{\hat{P}(1-\hat{P})/n}$$

pelo que temos então que resolver a equação seguinte em ordem a n

$$2\times 1.645\sqrt{0.06\times 0.94/n}<0.01 \Leftrightarrow \sqrt{n}>\frac{2\times 1.645\sqrt{0.06\times 0.94}}{0.01} \Leftrightarrow \sqrt{n}>78.13317$$

$$\Rightarrow n>78.13317^2 \underset{n\in\mathbb{N}}{\Rightarrow} n\geq 6105$$

pelo que para se ter a amplitude pretendida para o mesmo grau de confiança teríamos que aumentar a amostra de 200 para pelo menos 6105.

Exercícios