Métodos de estimação (intervalar)

Prof. Caio Azevedo

Introdução

- Vimos, até o momento, como construir mecanismos apropriados, sob diferentes aspectos, com fins de inferência pontual (estimativa pontual), acerca dos parâmetros de interesse.
- Embora, à essas estimativas (estimadores) tenhamos associado algum tipo de (medida de) precisão (variância, erro-padrão, erro-quadrático médio, veja aqui), muitas vezes é de grande relevância prover algum tipo de estimativa intervalar.
- Nesta parte, discutiremos a construção, propriedades e interpretação desse tipo de processo inferencial.

Estimação intervalar

- Suponha que $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, σ^2 conhecido, e :
  - amostra 1: $\widetilde{\mu} = \overline{x} = 78,00$.
  - amostra 2: $\widetilde{\mu} = \overline{x} = 8,00$.
- Qual das duas estimativas é mais confiável (em algum sentido)?
- Como medir a magnitude de erros associados às estimativas?
- Sabemos, nesse caso, que $\overline{X} \sim N(\mu, \sigma^2/n)$.
- Note que, $E = \overline{X} - \mu \sim N(0, \sigma^2/n)$ e $\sqrt{n}\frac{(\overline{X}-\mu)}{\sigma} \sim N(0,1)$.

Estimação intervalar

- Gostaríamos de que:
  - $P(|\overline{X} - \mu| > \epsilon)$ (pequeno, por exemplo 0,05).
  - $P(|\overline{X} - \mu| < \epsilon)$ (grande, por exemplo 0,95).
- Suponha que: $P(|\overline{X} - \mu| \leq \epsilon) = 0,95$. Assim,

$$
\begin{aligned}
& P\left(\left|\sqrt{n}\frac{(\overline{X} - \mu)}{\sigma}\right| \leq \frac{\sqrt{n}\epsilon}{\sigma}\right) = 0,95 \\
\leftrightarrow \quad & P\left(\left|\sqrt{n}\frac{(\overline{X} - \mu)}{\sigma}\right| \leq 1,96\right) = 0,95 \\
\leftrightarrow \quad & P\left(\overline{X} - 1,96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + 1,96\frac{\sigma}{\sqrt{n}}\right) = 0,95
\end{aligned}
$$

Estimação intervalar

- Assim,
  - 95% dos intervalos da forma $\left[\overline{x} - 1,96\frac{\sigma}{\sqrt{n}}; \overline{x} + 1,96\frac{\sigma}{\sqrt{n}}\right]$ contém o parâmetro μ (interpretação clássica).
  - A probabilidade do intervalo $\left[\overline{X} - 1,96\frac{\sigma}{\sqrt{n}}; \overline{X} + 1,96\frac{\sigma}{\sqrt{n}}\right]$ conter o verdadeiro valor de μ é 95%
- Def: Sejam $T_1 = T_1(\boldsymbol{X})$ e $T_2 = T_2(\boldsymbol{X})$ duas estatísticas tais que, $T_1(\boldsymbol{x}) \leq T_2(\boldsymbol{x})$, $\boldsymbol{x} \subseteq \mathcal{X}$.
  - Dizemos que $[T_1, T_2]$ é um intervalo de confiança para θ se $P(T_1 \leq \theta \leq T_2)$ não depende θ.
  - Se $\gamma \in (0,1)$ é tal que $P(T_1 \leq \theta \leq T_2) = \gamma$, γ é chamado de coeficiente de confiança.
  - Podemos também utilizar a notação $\alpha = 1 - \gamma$.

Estimação intervalar

- Notação: $IC(\theta;\gamma) = [T_1, T_2]$: intervalo de confiança de $100\gamma\%$ para θ.
- Def: Dizemos que $Q(\boldsymbol{X};\theta)$ é uma quantidade pivotal se a distribuição de Q não depender de θ mas sua forma (funcional), sim.

Estimação intervalar

- Exemplo 1: $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, $\boldsymbol{\theta} = (\mu, \sigma^2)'$.
  - $Q(\boldsymbol{X}; \boldsymbol{\theta}) = \sqrt{n}\left(\frac{\overline{X}-\mu}{\sigma}\right) \sim N(0,1)$. (para $\boldsymbol{\theta}$)
  - $Q(\boldsymbol{X}; \boldsymbol{\theta}) = \left[\sqrt{n}\left(\frac{\overline{X}-\mu}{\sigma}\right)\right]^2 \sim \chi^2_1$. (para $\boldsymbol{\theta}$).
  - $Q(\boldsymbol{X}; \boldsymbol{\theta}) = (n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}$. (para σ^2).
  - $Q(\boldsymbol{X}; \sigma^2) = \sum_{i=1}^n \frac{(X_i-\mu)^2}{\sigma^2} \sim \chi^2_n$.(para σ^2, μ conhecido).
- Exemplo 2: $X_i \overset{iid}{\sim} \exp(\theta)$, $\mathcal{E}(X) = \theta$.
  - $Q(\boldsymbol{X}; \theta) = \frac{1}{\theta}\sum_{i=1}^n X_i \sim \text{gama}(n, 1)$.
  - $Q(\boldsymbol{X}; \theta) = \frac{2}{\theta}\sum_{i=1}^n X_i \sim \chi^2_{2n}$.

Estimação intervalar

- Lembrete: Se $X \in FL(\theta)$, então $Q(X;\theta) = X - \theta$ é uma quantidade pivotal.
- Lembrete: Se $X \in FES(\theta)$, então $Q(X;\theta) = \frac{X}{\theta}$ é uma quantidade pivotal.
- Lembrete: Se $X \in FLE(\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\mu, \sigma)'$, então $Q(X;\boldsymbol{\theta}) = \frac{X-\mu}{\sigma}$ é uma quantidade pivotal.
- Exemplo 3: $X_i \overset{iid}{\sim} U(0,\theta)$ (família de escala).
  - $Q(\boldsymbol{X};\theta) = \frac{X}{\theta} \sim U(0,1)$.
  - $Q(\boldsymbol{X};\theta) = \frac{\sum_{i=1}^{n} X_i}{\theta}$.
  - $Q(\boldsymbol{X};\theta) = \frac{Y_n}{\theta}$, $Y_n = \max(X_1, ..., X_n)$.

Estimação intervalar

- Exemplo: Seja $X_1, ..., X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, σ^2 conhecido. Temos que $Q(\boldsymbol{X}; \mu) = \frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \sim N(0,1)$.
- Seja $\gamma \in (0,1)$, e dada a relação $P(q_1 \leq Q \leq q_2) = \gamma$, obteremos os valores de q_1 e q_2 (possivelmente $q_2 = q_2(q_1)$).
- Note que

$$
\begin{aligned}
P\left(q_1 \leq \frac{\sqrt{n}(\overline{X} - \mu)}{\sigma} \leq q_2\right) &= \Phi(q_2) - \Phi(q_1) = \gamma \\
&\vdots \\
P\left(\overline{X} - q_2\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} - q_1\frac{\sigma}{\sqrt{n}}\right) &= \gamma
\end{aligned}
$$

Estimação intervalar

- Assim, $IC(\mu,\gamma) = \left[\overline{X} - q_2\frac{\sigma}{\sqrt{n}}; \overline{X} - q_1\frac{\sigma}{\sqrt{n}}\right]$ e L = comprimento do IC é dado por

$$\overline{X} - q_1\frac{\sigma}{\sqrt{n}} - \left(\overline{X} - q_2\frac{\sigma}{\sqrt{n}}\right) = (q_2 - q_1)\frac{\sigma}{\sqrt{n}}$$

- Dado que $\frac{\sigma}{\sqrt{n}}$ está "fixo", podemos trabalhar com $L = q_2 - q_1$, para determinar $q_i, i = 1, 2$.
- Com efeito, queremos minimizar (lembrando que $q_2 = g(q_1)$ e definindo $Z \sim N(0,1)$)

$$\begin{aligned} Q(q_1,\lambda) \equiv Q &= L + \lambda\left(\Phi(q_2) - \Phi(q_1) - \gamma\right) \\ &= q_2 - q_1 + \lambda\left(\Phi(q_2) - \Phi(q_1) - \gamma\right) \end{aligned}$$

Estimação intervalar

- Assim, temos que:

$$\begin{aligned}
\frac{\partial}{\partial q_1} Q &= \frac{dq_2}{dq_1} - 1 + \lambda \left(f_Z(q_2)\frac{dq_2}{dq_1} - f_Z(q_1) \right) \\
\frac{\partial}{\partial \lambda} Q &= \Phi(q_2) - \Phi(q_1) - \gamma = 0 \rightarrow f_Z(q_2)\frac{dq_2}{dq_1} - f_Z(q_1) = 0
\end{aligned}$$

- Portanto, $\frac{dq_2}{dq_1} = 1$ e, consequentemente, $f_Z(q_2) = f_Z(q_1)$, logo:

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} e^{-q_2^2/2} &= \frac{1}{\sqrt{2\pi}} e^{-q_1^2/2} \rightarrow q_2^2 = q_1^2 \\
&\rightarrow q_2 = q_1 \text{ ou } q_2 = -q_1
\end{aligned}$$

Estimação intervalar

- Logo, o IC de menor comprimento é dado por:

$$IC(\mu, \gamma) = \left[\overline{X} - q_2 \frac{\sigma}{\sqrt{n}}; \overline{X} + q_2 \frac{\sigma}{\sqrt{n}}\right].$$

  Se $\gamma = 0,95$, então $q_2 = 1,96$, por exemplo.

- OBS: repetir o exercício considerando σ^2 desconhecido e usando (provando): que $\frac{\sqrt{n}(\overline{X} - \mu)}{S} \sim t_{(n-1)}$, em que $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ e $S = \sqrt{S^2}$.

Exemplo

- Seja $X_1, ..., X_n \overset{iid}{\sim} U(0, \theta)$, encontre um IC ótimo para θ.
- Já tínhamos visto anteriormente que: $f_{Y_n}(y) = \frac{ny^{n-1}}{\theta} \mathbb{1}_{(0,\theta)}(y)$. Definindo $Q(\boldsymbol{X}; \theta) = W_n = \frac{Y_n}{\theta}$, podemos provar que (exercício) $f_{W_n}(w) = nw^{n-1} \mathbb{1}_{(0,1)}(w)$, ou seja $W \sim \text{beta}(n, 1)$
- Assim, para um dado n, temos que

$$
\begin{aligned}
P(q_1 \leq Q \leq q_2) &= \gamma \rightarrow P(Q \leq q_2) - P(Q \leq q_1) = \gamma \\
&\rightarrow F_{W_n}(q_2) - F_{W_n}(q_1) = q_2^n - q_1^n = \gamma \quad (1)
\end{aligned}
$$

do qual se obtem os valores de q_2 e q_1.

Exemplo

- Note, ainda, que:

$$P(q_1 \leq Q \leq q_2) = P\left(q_1 \leq \frac{Y_n}{\theta} \leq q_2\right) = P\left(\frac{Y_n}{q_2} \leq \theta \leq \frac{Y_n}{q_1}\right)$$

- Assim, $IC(\theta;\gamma) = \left[\frac{Y_n}{q_2}; \frac{Y_n}{q_1}\right]$, o que leva à $L = \frac{Y_n}{q_1} - \frac{Y_n}{q_2}$ ou $L \propto \frac{1}{q_1} - \frac{1}{q_2}$.
- Queremos minimizar $L = \frac{1}{q_1} - \frac{1}{q_2}$ $(q_2 = g(q_1))$ sujeito à (1). Assim, temos que

$$\frac{\partial L}{\partial q_1} = -\frac{1}{q_1^2} + \frac{1}{q_2^2}\frac{dq_2}{dq_1}(*)$$

Exemplo

- Cont.

$$nq_2^{n-1}\frac{dq_2}{dq_1} - nq_1^{n-1} = 0 \rightarrow \frac{dq_2}{dq_1} = \frac{q_1^{n-1}}{q_2^{n-1}}(**)$$

- Substituindo (**) em (*) vem que:

$$\frac{\partial L}{\partial q_1} = -\frac{1}{q_1^2} + \frac{1}{q_2^2}\frac{q_1^{n-1}}{q_2^{n-1}} = \frac{q_1^{n-1}}{q_2^{n+1}} - \frac{1}{q_1^2} = \frac{q_1^{n+1} - q_2^{n+1}}{q_1^2 q_2^{n+1}} < 0(q_1 < q_2)$$

Exemplo

- (Cont.) Isto é $L(q_1)$ é decrescente e, assim, o valor de $q_2 = g(q_1)$ que torna L mínimo é 1 ($q_2 \in (0,1)$). Assim:

$$\begin{aligned} P(q_1 \leq Q \leq q_2) &= q_2^n - q_1^n = \gamma \leftrightarrow 1 - \gamma = q_1^n \\ &\leftrightarrow q_1 = (1-\gamma)^{1/n} \end{aligned}$$

Portanto, $IC(\theta;\gamma) = \left[Y_n; \frac{Y_n}{(1-\gamma)^{1/n}} \right]$

Método da quantidade pivotal

- O método da quantidade pivotal (QP) é uma das técnicas mais utilizadas para obtenção de IC's.
- Dado $\gamma \in (0,1)$ e uma QP Q, tal que $P(q_1 \leq Q \leq q_2) = \gamma$, se existirem estatísticas $T_1 = T_1(\boldsymbol{X})$ e $T_2 = T_2(\boldsymbol{X})$, tais que

$$q_1 \leq Q(\boldsymbol{x};\theta) \leq q_2 \leftrightarrow T_1(\boldsymbol{x}) \leq \theta \leq T_2(\boldsymbol{x})$$

- Então $IC(\theta;\gamma) = [T_1(\boldsymbol{X}); T_2(\boldsymbol{X})]$
- Exemplo: Seja $X_1, ..., X_n$ uma aa de $X \sim N(\mu, \sigma^2)$. Vamos obter uma IC para σ^2 com a) μ conhecido e b) μ desconhecido.

Método da quantidade pivotal

- a) Temos que $Q = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 \sim \chi^2_n$ é uma quantidade pivotal, assim

$$
\begin{aligned}
& P\left(q_1 \leq \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 \leq q_2\right) = \gamma \\
\leftrightarrow \quad & P\left(\frac{1}{q_2}\sum_{i=1}^{n}(X_i - \mu)^2 \leq \sigma^2 \leq \frac{1}{q_1}\sum_{i=1}^{n}(X_i - \mu)^2\right) = \gamma
\end{aligned}
$$

- Assim $IC(\sigma^2, \gamma) = \left[\frac{1}{q_2}\sum_{i=1}^{n}(X_i - \mu)^2; \frac{1}{q_1}\sum_{i=1}^{n}(X_i - \mu)^2\right]$ é um IC de $100\gamma\%$ para σ^2, em que q_1 e q_2 são quantis apropriados oriundos de uma distribuição χ^2_n, por exemplo

Método da quantidade pivotal

- Por exemplo, $P(X \leq q_1) = \frac{1-\gamma}{2}; P(X \geq q_2) = \frac{1-\gamma}{2}, X \sim \chi^2_n$
- O comprimento desse IC é $CIC = \sum_{i=1}^{n}(X_i - \mu)^2 \left(\frac{1}{q_1} - \frac{1}{q_2}\right)$ e seu respectivo comprimento esperado é $ECIC = \mathcal{E}(CIC) = n\sigma^2 \left(\frac{1}{q_1} - \frac{1}{q_2}\right)$.
- Exercício: minimizar o CIC/ECIC.

Método da quantidade pivotal

- b) Temos que $Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$ é uma quantidade pivotal $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$, assim

$$
\begin{aligned}
& P\left(q_1 \leq \frac{(n-1)S^2}{\sigma^2} \leq q_2\right) = \gamma \\
\leftrightarrow \quad & P\left(\frac{(n-1)S^2}{q_2} \leq \sigma^2 \leq \frac{(n-1)S^2}{q_1}\right) = \gamma
\end{aligned}
$$

- Assim $IC(\sigma^2, \gamma) = \left[\frac{(n-1)S^2}{q_2}; \frac{(n-1)S^2}{q_1}\right]$ é um IC de $100\gamma\%$ para σ^2, em que q_1 e q_2 são quantis apropriados oriundos de uma distribuição $\chi^2_{(n-1)}$, por exemplo

Método da quantidade pivotal

- $P(X \leq q_1) = \frac{1-\gamma}{2}; P(X \geq q_2) = \frac{1-\gamma}{2}, X \sim \chi^2_{(n-1)}$
- O comprimento desse IC é $CIC = (n-1)S^2\left(\frac{1}{q_1} - \frac{1}{q_2}\right)$ e seu respectivo comprimento esperado é $ECIC = \mathcal{E}(CIC) = (n-1)\sigma^2\left(\frac{1}{q_1} - \frac{1}{q_2}\right)$.
- Exercício: minimizar o CIC/ECIC.

Resultados

- OBS: Seja $X_1, ..., X_n$ uma aa de uma vac $X \sim F_X(.;\theta)$ (sua fda). Então

$$Q = -\sum_{i=1}^{n} \ln F_X(X_i;\theta) \sim \text{gama}(n,1) \tag{2}$$

- Analogamente,

$$Q = -2\sum_{i=1}^{n} \ln F_X(X_i;\theta) \sim \chi^2_{2n} \tag{3}$$

- Seja $[T_1, T_2]$ um IC para θ com cc (coeficiente de confiança) γ. Se $\tau(\theta)$ é uma função estritamente monótona então, pode-se obter um IC para $\tau(\theta)$ com cc γ, a partir de uma IC para θ.

Resultados

- Por exemplo, se $\tau(\theta)$ é estritamente crescente, então

$$P(\tau(T_1) \leq \tau(\theta) \leq \tau(T_2)) = P(T_1 \leq \theta \leq T_2) = \gamma$$

- Portanto, $[\tau(T_1), \tau(T_2)]$ é um $IC[\tau(\theta); \gamma]$.
- Exemplo: Seja $X_1, ..., X_n$ uma aa de $X \sim \exp(\theta)$, $\mathcal{E}(X) = \theta$, obtenha um IC para o primeiro quartil. Temos que

$$P(X \leq q) = 1/4 \rightarrow 1 - e^{-q/\theta} = 1/4 \rightarrow q = -\theta \ln(3/4) = \tau(\theta) \quad (4)$$

Resultados

- Por outro lado, sabemos que $Q = \frac{1}{\theta}\sum_{i=1}^{n} X_i \sim \text{gama}(n, 1)$. Assim

$$
\begin{aligned}
& P\left(q_1 \leq \frac{1}{\theta}\sum_{i=1}^{n} X_i \leq q_2\right) = \gamma \\
\leftrightarrow \quad & P\left(\frac{\sum_{i=1}^{n} X_i}{q_2} \leq \theta \leq \frac{\sum_{i=1}^{n} X_i}{q_1}\right) = \gamma
\end{aligned}
$$

em que q_1 e q_2 são quantis obtidos a partir de uma distribuição gama$(n, 1)$.

Resultados

- Assim, um IC com cc γ para θ é dado por

$$\left[\frac{\sum_{i=1}^{n} X_i}{q_2}; \frac{\sum_{i=1}^{n} X_i}{q_1}\right]$$

e, para $\tau(\theta)$ dado por

$$\left[-\ln(3/4)\frac{1}{q_2}\sum_{i=1}^{n} X_i; -\ln(3/4)\frac{1}{q_1}\sum_{i=1}^{n} X_i\right]$$

Resultados

- Exemplo: Seja $X_i \overset{iid}{\sim} U(0,\theta), i = 1, 2..., n$. Já vimos que um IC com cc $100\gamma\%$ é dado por $IC(\theta;\gamma) = \left[\frac{Y_n}{q_2}; \frac{Y_n}{q_1}\right]$, com $CIC = Y_n\left(\frac{1}{q_1} - \frac{1}{q_2}\right)$ e $ECIC = \frac{n\theta}{n+1}\left(\frac{1}{q_1} - \frac{1}{q_2}\right)$.
- Exercício, utilizar a QP: $-\sum_{i=1}^{n} \ln F_X(X_i;\theta)$ para obter um $IC(\theta;\gamma)$, em que F é a fda de X e comparar com o resultado acima.

Alguns resultados de IC para populações normais

- Sejam $X_1, ..., X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. Queremos construir uma região de confiança para $\boldsymbol{\theta} = (\mu, \sigma^2)$, ou seja (gráfico slide seguinte):
- Em outras palavras:

$$P(T_1 \leq \mu \leq T_2, S_1 \leq \sigma^2 \leq S_2) = \gamma$$

- Sejam $Q_1(\boldsymbol{X}; \boldsymbol{\theta}) = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{(n-1)}$ e $Q_2(\boldsymbol{X}; \theta) = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$. Assim

$$P\left(\overline{X} - q_1\frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + q_2\frac{S}{\sqrt{n}}; \frac{(n-1)S^2}{q_3} \leq \sigma^2 \leq \frac{(n-1)S^2}{q_4}\right) = \gamma$$

Alguns resultados de IC para populações normais

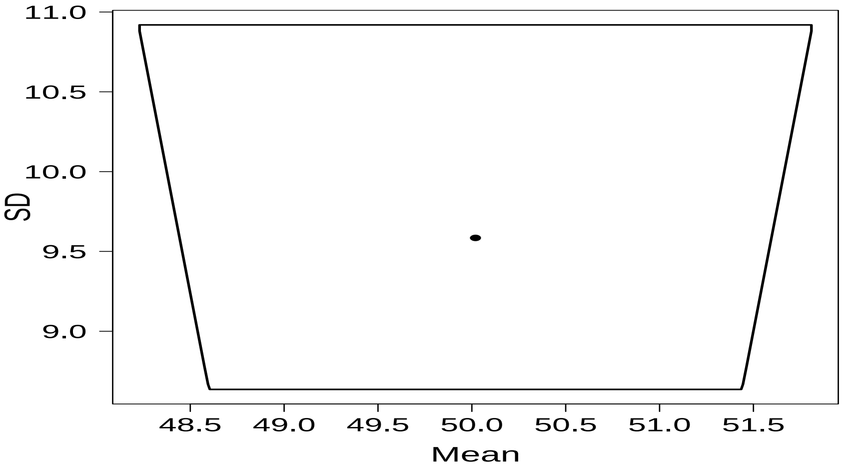
- Para contornar o problema acima (dependência entre Q_1 e Q_2, podemos considerar outras quantidades pivotais, por exemplo:

$$Q_1^*(\boldsymbol{X};\boldsymbol{\theta}) = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1);\, Q_2^*(\boldsymbol{X};\boldsymbol{\theta}) = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

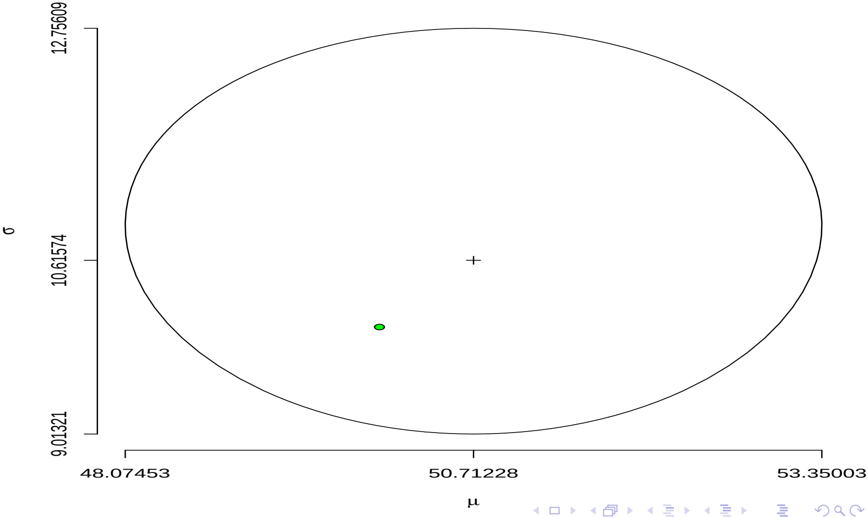
- Assim:

$$\begin{aligned}
& P\left(-q_1 \leq Q_1^* \leq q_1;\, q_2 \leq Q_2^* \leq q_2^*\right) \\
= \quad & P\left(\left|\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right| \leq q_1;\, \frac{(n-1)S^2}{q_2^*} \leq \sigma^2 \leq \frac{(n-1)S^2}{q_2}\right) \\
= \quad & P\left(\left|\frac{\overline{X}-\mu}{q_1/\sqrt{n}}\right| \leq \sigma;\, \frac{(n-1)S^2}{q_2^*} \leq \sigma^2 \leq \frac{(n-1)S^2}{q_2}\right) \\
= \quad & P\left(\left|\frac{\overline{X}-\mu}{q_1/\sqrt{n}}\right| \leq \sigma\right) P\left(\frac{(n-1)S^2}{q_2^*} \leq \sigma^2 \leq \frac{(n-1)S^2}{q_2}\right)
\end{aligned}$$

Exemplo de região de confiança $N(\mu, \sigma^2)$



Exemplo de região de confiança $N(\mu, \sigma^2)$



Alguns resultados de IC para populações normais

- Portanto, podemos determinar um IC para σ^2, ou seja

$$IC(\sigma^2;\gamma) = \left[\frac{(n-1)S^2}{q_2^*};\frac{(n-1)S^2}{q_2}\right]$$

- E, para cada valor para σ^2, constante no respectivo IC, construir um $IC(\mu,\gamma)$, ou seja

$$IC(\mu;\gamma) = \left[\overline{X} - q_1\frac{\sigma}{\sqrt{n}};\overline{X} + q_1\frac{\sigma}{\sqrt{n}}\right].$$

Observações

- O IC aleatório $IC(\theta, \gamma) = [T_1, T_2]$, $T_i = T_i(\boldsymbol{X}), i = 1, 2$, tem a seguinte interpretação: a probabilidade do intervalo conter o verdadeiro valor do parâmetro é γ, lembrando que

$$P(T_1 \leq \theta \leq T_2) = \gamma.$$

- De posse de uma amostra aletória, e definidos quantis $(q_1, q_2)'$ (de forma apropriada), teremos um intervalo numérico, $IC(\theta, \gamma) = [t_1, t_2]$, $t_i = t_i(\boldsymbol{x})$, $i = 1, 2$, para o qual a interpretação é: retirando-se m amostras e calculando-se o intervalo de confiança numérico para cada uma delas, espera-se que $100\gamma\%$ deles contenham o verdadeiro valor do parâmetro. Exemplos (próximo slide):

Observações e exemplos

- (Cont. observações) Para que um IC esteja devidamente definido, devemos sempre apresentar como obter os valores (quantis) (q_1, q_2) associados á distribuição utilizada para a sua construção.
- $N(\mu, \sigma^2)$, σ^2 desconhecido, n$= 20$, $\overline{x} = 103,72$, $\widetilde{\sigma}^2 = 256,94$, $\gamma = 0,95$, $q_2 = 2,09$; $q_1 = -2,09$ (distribuição $t_{20-1=19}$ graus de liberdade, $IC(\mu, 0,95) = [96,21; 111,22]$.
- $N(\mu, \sigma^2)$, μ desconhecido, n$= 35$, $\overline{x} = 203,41$, $\widetilde{\sigma}^2 = 331,44$, $\gamma = 0,99$, $q_2 = 58,96$; $q_1 = 16,50$ (distribuição $\chi^2_{(35-1=34)}$ graus de liberdade, $IC(\mu, 0,99) = [190,94; 682,31]$.
- $U(0, \theta)$, n $= 30$, $\gamma = 0,90$, $\widetilde{\theta} = y_n = 93,09$ $IC(\theta, \gamma) = [93,09; 100,52]$ (de menor comprimento, com base no máximo amostral).

Alguns resultados de IC para populações normais

- Considere agora: $X_1, ..., X_n$ aa de $X \sim N(\mu_1, \sigma_1^2)$ e $Y_1, ..., Y_m$ aa de $Y \sim N(\mu_2, \sigma_2^2)$, $X_i \perp Y_j, \forall i, j$. Desejamos construir um $IC(\mu_1 - \mu_2; \gamma)$, sob $\sigma_1^2 = \sigma_2^2 = \sigma^2$ em que $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma^2)'$.
- Temos que (prova-exercício) :

$$Q(\boldsymbol{X}, \boldsymbol{Y}) = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}} \sim t_{(n+m-2)},$$

em que $S_X^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$, $S_Y^2 = \frac{1}{m-1} \sum_{i=1}^{m} (Y_i - \overline{Y})^2$, $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ e $\overline{Y} = \frac{1}{m} \sum_{i=1}^{m} Y_i$.

Alguns resultados de IC para populações normais

- Logo:

$$IC(\mu_1 - \mu_2; \gamma) = \left[(\overline{X} - \overline{Y}) \pm t_{\alpha/2} \sqrt{\left(\frac{1}{n} + \frac{1}{m} \right) \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}} \right]$$

em que $P(X \geq t_{\alpha/2}) = \frac{\alpha}{2}$, $X \sim t_{(n+m-2)}$.

- Exercício: Pesquisar o caso em que $\sigma_1^2 \neq \sigma_2^2$ são desconhecidos e coinstruir um IC quando σ_1^2 e σ_2^2 forem conhecidos.

Alguns resultados de IC para populações normais

- Suponha agora que: $(X_i, Y_i) \overset{iid}{\sim} N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\mu} = (\mu_1, \mu_2)'$ e $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$. Temos que $(\theta = (\mu_1, \mu_2, \sigma_1^2\sigma_1^2, \rho)')$:

$$Q(\boldsymbol{X}, \boldsymbol{Y}; \boldsymbol{\theta}) = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\frac{S_D}{\sqrt{n}}} \sim t_{(n-1)},$$

lembrando que $Z_i = X_i - Y_i \sim N(\mu_1 - \mu_2, \sigma_D^2)$, em que $\sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho$ e $S_D^2 = \frac{1}{n}\sum_{i=1}^n (Z_i - \overline{Z})^2$, $\overline{Z} = \frac{1}{n-1}\sum_{i=1}^n Z_i$ e $S_D = \sqrt{S_D^2}$. Assim:

$$IC(\mu_1 - \mu_2; \gamma) = \left[(\overline{X} - \overline{Y}) \pm t_{\alpha/2}\frac{S_D}{\sqrt{n}}\right]$$

em que $P(X \geq t_{\alpha/2}) = \frac{\alpha}{2}$, $X \sim t_{(n-1)}$.

Alguns resultados de IC para populações normais

- Exemplo: Construir um $IC\left[\frac{\sigma_1^2}{\sigma_2^2};\gamma\right]$, nas duas seguintes situações:
  a) μ_1 e μ_2 conhecidos ($\boldsymbol{\theta} = (\sigma_1^2, \sigma_2^2)'$). Temos que

$$Q(\boldsymbol{X}, \boldsymbol{Y}; \boldsymbol{\theta}) = \frac{\sigma_1^2}{\sigma_2^2}\frac{\sum_{i=1}^{m}(Y_i - \mu_2)^2/m}{\sum_{i=1}^{n}(X_i - \mu_1)^2/n} \sim F_{(m,n)}$$

- Assim

$$IC\left(\frac{\sigma_1^2}{\sigma_2^2};\gamma\right) = \left[\frac{1}{q_2}\frac{\sum_{i=1}^{n}(X_i - \mu_1)^2/n}{\sum_{i=1}^{m}(Y_i - \mu_2)^2/m}; \frac{1}{q_1}\frac{\sum_{i=1}^{n}(X_i - \mu_1)^2/n}{\sum_{i=1}^{n}(Y_i - \mu_2)^2/m}\right],$$

em que $P(X \leq q_1) = \alpha/2$ e $P(X \geq q_2) = \alpha/2$, $X \sim F_{(m,n)}$

Alguns resultados de IC para populações normais

- Exemplo: Construir $IC\left[\frac{\sigma_1^2}{\sigma_2^2};\gamma\right]$, nas duas seguintes situações:

  b) μ_1 e μ_2 desconhecidos ($\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)'$),
  $S_X^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2/(n-1)$ e $S_Y^2 = \sum_{i=1}^{m}(Y_i - \overline{Y})^2/(m-1)$.

$$Q(\boldsymbol{X}, \boldsymbol{Y}; \boldsymbol{\theta}) = \frac{\sigma_1^2}{\sigma_2^2}\frac{S_Y^2}{S_X^2} \sim F_{(m-1,n-1)}$$

- Assim

$$IC\left(\frac{\sigma_1^2}{\sigma_2^2};\gamma\right) = \left[\frac{1}{q_2}\frac{S_Y^2}{S_X^2};\frac{1}{q_1}\frac{S_Y^2}{S_X^2}\right],$$

em que $P(X \leq q_1) = \alpha/2$ e $P(X \geq q_2) = \alpha/2$, $X \sim F_{(m-1,n-1)}$

Outras formas de obtenção de IC's

- Seja T uma estatística com fda $F_T(.;\theta)$, em que F_T é monótona em θ. Isto é, $F_T(.;\theta)$ é crescente como função de θ, se $\forall \theta_1 \leq \theta_2$, $F_T(t;\theta_1) \leq F_T(t;\theta_2)$, $\forall t$ fixo.
- Teorema: Seja T uma estatística com fda $F_T(.;\theta)$ contínua. Dado $\alpha \in (0,1)$ e supondo que $\exists$ funções $\theta_L(t)$ e $\theta_U(t)$, tais que
  a) Se $F_T(.;\theta)$ é decrescente em θ, para cada $t \in B$ (suporte de T), definir $\theta_L(t)$ e $\theta_U(t)$ por:

  $$F_T(t;\theta_U(t)) = \alpha/2 \text{ e } F_T(t;\theta_L(t)) = 1 - \alpha/2$$

  ou (próximo slide)

Outras formas de obtenção de IC's

- Cont.
  b) Se $F_T(.;\theta)$ é crescente em θ e para cada $t \in B$ defina $\theta_L(t)$ e $\theta_U(t)$ por:

$$F_T(t;\theta_U(t)) = 1 - \alpha/2 \;\text{ e }\; F_T(t;\theta_L(t)) = \alpha/2$$

- Então $IC(\mu;\gamma) = [\theta_L(T);\theta_U(T)]$

Outras formas de obtenção de IC's

- Exemplo: Seja $X_1 ..., X_n$ uma aa de X tal que $f_X(x;\mu) = e^{-(x-\mu)}\mathbb{1}_{(\mu,\infty)}(x)$ e defina $T = Y_1$ (mínimo), então $f_T(t;\mu) = ne^{-n(t-u)}$ e $F_T(t;\mu) = 1 - e^{-n(t-\mu)}$ (decrescente em μ). Assim

$$\begin{aligned} 1 - e^{-n(t-\mu_U(t))} &= \alpha/2 \rightarrow e^{-n(t-\mu_U(t))} = 1 - \alpha/2 \\ &= \mu_U(t) = t + \frac{1}{n}\ln(1-\alpha/2) \end{aligned}$$

- Analogamente $\mu_L(t) = t + \frac{1}{n}\ln(\alpha/2)$, em que $\alpha = 1 - \gamma$. Portanto

$$IC(\mu;\gamma) = \left[T + \frac{1}{n}\ln\left(\frac{\alpha}{2}\right); T + \frac{1}{n}\ln\left(1 - \frac{\alpha}{2}\right)\right].$$

Outras formas de obtenção de IC's

- Note que as definições anteriores implicam que: $P(\theta_L(T) \leq \theta \leq \theta_U(T)) = P(\theta_U(T) \leq \theta) - P(\theta_L(T) \leq \theta) = 1 - \alpha = \gamma$ (veja Casella & Berger (2006)).
- Exercício: Para o problema anterior, determine um IC para μ utilizando a QP: $Q(\boldsymbol{X}; \mu) = Y_1 - \mu \sim \exp(1/n)$.
- Exercício: Seja $X_1, ..., X_n$ uma aa de $X \sim U(0, \theta)$, tal que $F_{Y_n}(t; \theta) = \frac{t^n}{\theta^n}$. Encontre uma IC para θ, utilizando a metodologia descrita acima.

Teorema: Caso discreto

- Seja T uma estatística com distribuição discreta tal que $F_T(t;\theta) = P(T \leq t)$. Dado α e supondo que as funções $\theta_L(t)$ e $\theta_U(t)$ são tais que:
  a) Se $F_T(t;\theta)$ é decrescente em θ, para cada $t \in B$ defina $\theta_L(t)$ e $\theta_U(t)$ por
  $$P(T \leq t; \theta_U(t)) = \alpha/2 \ \text{ r } \ P(T \geq t; \theta_L(t)) = \alpha/2$$
  ou
  a) Se $F_T(t;\theta)$ é crescente em θ, para cada $t \in B$ defina $\theta_L(t)$ e $\theta_U(t)$ por
  $$P(T \geq t; \theta_U(t)) = \alpha/2 \ \text{ r } \ P(T \leq t; \theta_L(t)) = \alpha/2$$
- Então $[\theta_L(T); \theta_U(T)]$ é um IC para θ com cc $\gamma = 1 - \alpha$.

Teorema: Caso discreto

- Exemplo: Sejam $X_1, ..., X_n$ uma aa de $X \sim \text{Poisson}(\theta)$. Vamos considerar a estatística $T = \sum_{i=1}^{n} X_i$. Assim

$$f_T(t;\theta) = \frac{e^{-n\theta}(n\theta)^t}{t!} \mathbb{1}_{\{0,1,2,\ldots\}}(t)$$

$$F_T(t;\theta) = \sum_{x \leq t} \frac{e^{-n\theta}(n\theta)^x}{x!} = \sum_{x \leq t} \frac{(n\theta)^x}{e^{n\theta} x!}$$

- Que é decrescente em θ se $x \leq n\theta$. Portanto,

$$\begin{aligned} P(T \leq t; \theta_U(t)) &= \alpha/2 \ ; \ P(T \geq t; \theta_L(t)) = \alpha/2 \\ &\leftrightarrow \sum_{x=0}^{t} \frac{(n\theta_U(t))^x}{e^{n\theta_U(t)} x!} = \frac{\alpha}{2} \ ; \ \sum_{x=t}^{\infty} \frac{(n\theta_L(t))^x}{e^{n\theta_L(t)} x!} = \frac{\alpha}{2} \end{aligned}$$

Teorema: Caso discreto

- Obs: gama-Poisson. Se $X \sim \text{gama}(\alpha, \beta), \alpha \in \mathcal{N}$, então:

$$
\begin{aligned}
P(X \leq x) &= P(Y \geq \alpha) \\
P(X > x) &= P(Y < \alpha)
\end{aligned}
$$

em que $Y \sim \text{Poisson}(\beta)$. Assim, se

$$
\begin{aligned}
\frac{\alpha}{2} &= P(T \leq t, \theta_U(t)), T \sim \text{Poisson}(n\theta_U(t)) \\
&= P(T < t+1, \theta_U(t)) = P(X > n), X \sim \text{gama}(t+1, \theta_U(t)) \\
&= P(2\theta_U(t)X > 2n\theta_U(t)) = P(Z > 2n\theta_U(t)), Z \sim \chi^2_{2(t+1)}
\end{aligned}
$$

- Então $2n\theta_U(t) = q_2, P(Z > q_2) = 1 - \alpha/2, \rightarrow \theta_U(t) = \frac{1}{2n} q_2$.

Teorema: Caso discreto

- Analogamente, $\theta_L(t) = \frac{1}{2n} q_1, P(Z < q_1) = \alpha/2$.
- Portanto

$$IC(\theta, \gamma) = \left[\frac{1}{2n} q_1 ; \frac{1}{2n} q_2 \right]$$

- Obs: De uma forma geral, um IC bilateral para θ com cc γ, pode ser definido como

$$[T_1(\boldsymbol{X}); T_2(\boldsymbol{X})], P(T_1(\boldsymbol{X}) \leq \theta \leq T_2(\boldsymbol{X})) \geq \gamma.$$

Esta definição se aplica especialmente para as distribuições discretas.

Intervalos unilaterais

- Obs: Também podemos definir IC's, com cc γ, unilaterais à esquerda e à direita, respectivamente por:
  - $P(T_1 \leq \theta) = \gamma(\geq \gamma)$.
  - $P(T_2 \geq \theta) = \gamma(\geq \gamma)$
- Os intervalos gerados são da forma $[T_1; \theta_1]$ e $[\theta_2; T_2]$ em que θ_1 e θ_2 são, respectivamente, os limites superior e inferior do espaço paramétrico.

Intervalos unilaterais

- Exemplo: Seja $X_1, ..., X_n$ uma aa de $X \sim \exp(\theta)$. Temos que $S_X(x) = e^{-x/\theta} I_{(0,\infty)}(x)$;
- Assim $Q(\boldsymbol{X};\theta) = -\sum_{i=1}^{n} \ln S_X(X_i) \sim \text{gama}(n,1)$ e $Q(\boldsymbol{X};\theta)^* = 2Q(\boldsymbol{X};\theta) \sim \chi^2_{2n}$, logo:

$$Q^*(\boldsymbol{X};\theta) = -2\sum_{i=1}^{n}\left(\frac{-X_i}{\theta}\right) = \frac{2n\overline{X}}{\theta}$$

Intervalos unilaterais

- Portanto,

$$
\begin{aligned}
P\left(Q^*(\boldsymbol{X};\theta) \leq q_1\right) &= \gamma \leftrightarrow P\left(\frac{2n\overline{X}}{\theta} \geq q_1\right) = \gamma \\
&\rightarrow P\left(\theta \leq \frac{2n\overline{X}}{q_1}\right)
\end{aligned}
$$

- Assim, um IC unilateral à direita, com cc γ é dado por $IC(\theta;\gamma) = \left[0, \frac{2n\overline{X}}{q_1}\right]$. Analogamente, um IC unilateral à esquerda, com cc γ é dado por $IC(\theta;\gamma) = \left[\frac{2n\overline{X}}{q_2}, \infty\right]$

Intervalos de confiança unilaterais uniformemente mais acurados

- À semelhança do que ocorre na estimação pontual, no que ocorre com testes de hipóteses (TH) e pela relação entre IC e TH (como veremos mais à frente), é esperado que exista uma certa otimalidade de IC's construídos a partir das regiões de aceitação de testes uniformemente mais poderosos.
- De fato, em algumas situações, tal relação se mostra verdadeira.
- Um IC unilateral à esquerda, onde $\theta_L(\boldsymbol{X})$ é o seu limite inferior, é dito ser uniformemente (ICUMA) mais acurado se,

$$P(\theta_L(\boldsymbol{X}) \leq \theta) \leq P(\theta_L(\boldsymbol{X})^* \leq \theta), \forall \theta \in \Theta \quad (5)$$

e $\forall \theta_L^*(\boldsymbol{X})$ um outro limite inferior.

Intervalos de confiança unilaterais uniformemente mais acurados

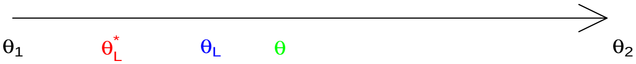
- Analogamente, um IC unilateral à direita, digamos $\theta_U(\boldsymbol{X})$ é dito ser uniformemente mais acurado se

$$P(\theta_U(\boldsymbol{X}) \geq \theta) \leq P(\theta_U(\boldsymbol{X})^* \geq \theta), \forall \theta \in \Theta \tag{6}$$

  e $\forall \theta_U^*(\boldsymbol{X})$ um outro limite superior.

- As condições (5) e (6) significam, respectivamente que $\theta_L(\boldsymbol{X})$ é estocasticamente maior do que $\theta_L^*(\boldsymbol{X})$, e que $\theta_U(\boldsymbol{X})$ é estocasticamente menor do que $\theta_U(\boldsymbol{X})$.
- Em outras palavras, a probabilidade de termos intervalos com menores comprimento é maior usando-se $\theta_L(\boldsymbol{X})$ ou $\theta_U(\boldsymbol{X})$.

ICUEUMA



ICUDUMA

Intervalos de confiança unilaterais uniformemente mais acurados

- Teorema: Seja $\boldsymbol{X} = (X_1, ..., X_n)'$ uma va com fdp $f_{\boldsymbol{X}}(.;\theta)$. Para cada $\theta_0 \in \Theta$ seja $A(\theta_0)$ a RA de um teste UMP de nível α para testar $H_0 : \theta = \theta_0$ vs $H_1 : \theta > \theta_0$ ($\theta < \theta_0$).
- Seja $C(\boldsymbol{X})$ o IC com cc $1-\alpha$ formado a partir de $A(\theta_0)$. Então $C(\boldsymbol{X})$ é o IC uniformemente mais acurado para θ com coeficiente de confiança $\gamma = 1-\alpha$.
- Voltaremos a esse assunto quando da apresentação da parte de teste de hipóteses.

Intervalos de confiança assintóticos

- Muitas vezes não dispomos de uma quantidade pivotal e/ou não é fácil obter sua distribuição exata.
- Ainda, utilizar os resultado (2) ou (3) pode não ser útil, devido à eventual complexidade funcional da fda, o que vale também para o Teorema relativo à fda de uma dada estatística
- No caso discreto, em particular, tais dificuldades se potencializam.
- Nesses casos, podemos recorrer a construção de IC's assintóticos, isto é, IC's baseados na distribuição assintótica de alguma estatística (estimador) ou quantidade pivotal.
- Vimos, por exemplo, resultados assintóticos relativos aos EMV e EMM, o que também existe para os estimadores de mínimos quadrados (ou aqueles que minimizam alguma função perda).

Intervalos de confiança assintóticos

- Seja $X_1, ..., X_n$ uma aa de X e suponha que $f_{\boldsymbol{X}}(.;\theta)$ satisfaz as CR. Se $\widehat{\theta}$ é o EMV de $\theta \subseteq \Theta \in \mathcal{R}$, então

$$\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow[n \to \infty]{D} N(0, I_1^{-1}(\theta))$$

ou $(\widehat{\theta} - \theta) \approx N(0, I^{-1}(\theta))$, para n suficientemente grande.

- Para obter um IC (assintótico) pafra θ, podemos usar a seguinte "quantidade pivotal" (assintótica):

$$Q(\boldsymbol{X};\theta) = \frac{\widehat{\theta} - \theta}{\sqrt{I^{-1}(\theta)}} \approx N(0,1)$$

para n suficientemente grande.

Intervalos de confiança assintóticos

- Dado $\gamma \in (0,1)$, usa-se a relação:

$$P\left(-q_{\alpha/2} \leq \frac{\widehat{\theta} - \theta}{\sqrt{I^{-1}(\theta)}} \leq q_{\alpha/2}\right) = \gamma \qquad (7)$$

- Duas opções:
    a) Usar a relação (7) e verificar se $\exists$ $T_1(\boldsymbol{X})$ e $T_2(\boldsymbol{X})$, $P\left(T_1(\boldsymbol{X}) \leq \theta \leq T_2(\boldsymbol{X})\right) = 1 - \alpha = \gamma$. Então $IC_A(\theta, \gamma) = [T_1(\boldsymbol{X}); T_2(\boldsymbol{X})]$, em que $P(Z \geq q_{\alpha/2}) = \alpha/2$, $Z \sim N(0,1)$.

Intervalos de confiança assintóticos

b) Como já vimos, $Q(\boldsymbol{X};\theta) = \frac{\widehat{\theta}-\theta}{\sqrt{I^{-1}(\widehat{\theta})}} \approx N(0,1)$, para n suficientemente grande. Então:

$$P\left(-q_{\alpha/2} \leq \frac{\widehat{\theta}-\theta}{\sqrt{I^{-1}(\widehat{\theta})}} \leq q_{\alpha/2}\right) = \gamma$$

Assim, $IC_A(\theta,\gamma) = \left[\widehat{\theta} - q_{\alpha/2}\sqrt{I^{-1}(\widehat{\theta})}, \widehat{\theta} + q_{\alpha/2}\sqrt{I^{-1}(\widehat{\theta})}\right]$, em que $P(Z \geq q_{\alpha/2}) = \alpha/2, Z \sim N(0,1)$.

Intervalos de confiança assintóticos

- Exemplo: Seja $X_1, ..., X_n$ uma aa de $X \sim \exp(\theta)$. Sabemos que $\widehat{\theta} - \theta \approx N(0, I^{-1}(\theta))$, em que $I(\theta) = \frac{n}{\theta^2}$. Seja $Q(\boldsymbol{X};\theta) = \frac{\sqrt{n}(\widehat{\theta}-\theta)}{\theta}$, em que $\widehat{\theta} = \overline{X}$. Então:

$$
\begin{aligned}
& P\left(-q_{\alpha/2} \leq \frac{\sqrt{n}(\widehat{\theta}-\theta)}{\theta} \leq q_{\alpha/2}\right) = \gamma \\
\rightarrow \quad & P\left(-\frac{q_{\alpha/2}}{\sqrt{n}} \leq \frac{\widehat{\theta}}{\theta} - 1 \leq \frac{q_{\alpha/2}}{\sqrt{n}}\right) = P\left(-\frac{q_{\alpha/2}}{\sqrt{n}} \leq \frac{\widehat{\theta}}{\theta} - 1 \leq \frac{q_{\alpha/2}}{\sqrt{n}}\right) = \gamma \\
\rightarrow \quad & P\left(\frac{\widehat{\theta}}{1 + \frac{q_{\alpha/2}}{\sqrt{n}}} \leq \theta \leq \frac{\widehat{\theta}}{1 - \frac{q_{\alpha/2}}{\sqrt{n}}}\right) = \gamma
\end{aligned}
$$

Assim, $IC_A(\theta,\gamma) = \left[\frac{\widehat{\theta}}{1+\frac{q_{\alpha/2}}{\sqrt{n}}}; \frac{\widehat{\theta}}{1-\frac{q_{\alpha/2}}{\sqrt{n}}}\right]$, em que $P(Z \geq q_{\alpha/2}) = \alpha/2$, $Z \sim N(0,1)$.

Intervalos de confiança assintóticos

- Exemplo: Seja $X_1, ..., X_n$ uma aa de $X \sim$ Bernoulli(θ). Sabemos que $\frac{\widehat{\theta}-\theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \approx N(0,1)$, para n suficientemente grande.
- Contudo, este resultado não é útil para a construção de IC's pois não conseguiremos isolar θ como no caso anterior.
- Por outro lado, também temos que $\frac{\widehat{\theta}-\theta}{\sqrt{\frac{\widehat{\theta}(1-\widehat{\theta})}{n}}} \approx N(0,1)$, para n suficientemente grande.
- Assim, temos que

$$IC_A(\theta;\gamma) = \left[\widehat{\theta} - q_{\alpha/2}\sqrt{\frac{\widehat{\theta}(1-\widehat{\theta})}{n}}; \widehat{\theta} - q_{\alpha/2}\sqrt{\frac{\widehat{\theta}(1-\widehat{\theta})}{n}}\right],$$

em que $P(Z \geq q_{\alpha/2}) = \alpha/2, Z \sim N(0,1)$.

Intervalos de confiança assintóticos

- Seja $X_1, ..., X_n$ uma aa de $X \sim \text{beta}(\theta, 1)$, então o EMM de θ é dado por $\widehat{\theta} = \frac{\overline{X}}{1-\overline{X}}$ (veja os resultados assintóticos propostos para os EMM).
- Nesse caso, temos que $\widehat{\theta} - \theta \approx N\left(0, \frac{\theta(1+\theta)^2}{n(\theta+2)}\right)$, para n suficientemente grande. Então

$$Q(\boldsymbol{X};\theta) = \frac{\widehat{\theta} - \theta}{\sqrt{\frac{\theta(1+\theta)^2}{n(\theta+2)}}} \approx N(0,1)$$

para n suficientemente grande. Assim como para o caso da Bernoulli, seria difícil isolar θ para se obter um IC.

Intervalos de confiança assintóticos

- Contudo, podemos usar o seguinte resultado

$$Q(\boldsymbol{X};\theta) = \frac{\widehat{\theta} - \theta}{\sqrt{\frac{\widehat{\theta}(1+\widehat{\theta})^2}{n(\widehat{\theta}+2)}}} \approx N(0,1)$$

pois, nesse caso, $\widehat{\theta}$ é uma estimador consistente de θ. Assim, temos que

$$IC_A(\theta;\gamma) = \left[\widehat{\theta} - q_{\alpha/2}\sqrt{\frac{\widehat{\theta}(1+\widehat{\theta})^2}{n(\widehat{\theta}+2)}}; \widehat{\theta} + q_{\alpha/2}\sqrt{\frac{\widehat{\theta}(1+\widehat{\theta})^2}{n(\widehat{\theta}+2)}}\right],$$

em que $P(Z \geq q_{\alpha/2}) = \alpha/2, Z \sim N(0,1)$.

Intervalos de confiança assintóticos

- Por outro lado, sabemos que $\widehat{\theta} = -\frac{n}{\sum_{i=1}^{n} \ln X_i}$ é o EMV de θ e $I(\theta) = \frac{\theta^2}{n}$ (e, também, que satisfaz as CR).
- Assim, $Q(\boldsymbol{X};\theta) = \frac{\sqrt{n}(\widehat{\theta}-\theta)}{\theta} \approx N(0,1)$, para n suficientemente grande. Assim

$$P\left(-q_{\alpha/2} \leq \frac{\sqrt{n}(\widehat{\theta}-\theta)}{\theta} \leq q_{\alpha/2}\right) = \gamma \tag{8}$$

$$\rightarrow \quad P\left(\frac{\widehat{\theta}}{1+\frac{1}{\sqrt{n}}q_{\alpha/2}} \leq \theta \leq \frac{\widehat{\theta}}{1-\frac{1}{\sqrt{n}}q_{\alpha/2}}\right) = \gamma$$

- Portanto $IC(\theta,\gamma) = \left[\frac{\widehat{\theta}}{1+\frac{1}{\sqrt{n}}q_{\alpha/2}}; \frac{\widehat{\theta}}{1-\frac{1}{\sqrt{n}}q_{\alpha/2}}\right]$, em que $P(Z \geq q_{\alpha/2}) = \alpha/2$, $Z \sim N(0,1)$.

Intervalos de confiança assintóticos

- Exercício: Seja $X_1, ..., X_n$ uma aa de X em que

$$f_X(x;\theta) = \frac{e^{-(x-\theta)}}{[1+e^{-(x-\theta)}]^2}\mathbb{1}_{\mathcal{R}}(x),$$

a qual satisfaz as CR. Seja $\widehat{\theta}$ o EMV de θ. Encontre sua distribuição assintótica e, a partir dela, obtenha um IC_A com cc γ para θ.