CE085 - Estatística Inferencial
Teoremas Limites

Prof. Wagner Hugo Bonat

Laboratório de Estatística e Geoinformação - LEG
Curso de Bacharelado em Estatatística
Universidade Federal do Paraná - UFPR

31 de agosto de 2018

Conteúdo

Conteúdo

1. Motivação, Definições e Desigualdades.
2. Leis dos Grandes Números (LGN)
    - LGN de Chebyshev;
    - LGN de Kolmogorov;
    - LGN de Markov.
3. Teorema Central do Limite
    - Teorema de Lindeberg-Levy;
    - Teorema de Lindeberg-Levy multivariado;
    - Teorema de Lindeberg-Feller;
    - Teorema de Liapounov;
    - Teorema de Slutsky.

Motivação, Definições e Desigualdades

Motivação

- Sejam $Y_1, Y_2, \ldots, Y_n$ variáveis aleatórias (v.a) independentes e idênticamente distribuídas (iid) com $E(Y_i) = \mu$ e $V(Y_i) = \sigma^2$.
- Objetivo: Baseado em um conjunto de realizações de Y_i *estimar* o valor μ.
- Estratégia simples: Usar a versão empírica de μ, i.e.

$$\hat{\mu} = \sum_{i=1}^{n} \frac{Y_i}{n}.$$

- $\hat{\mu}$ é uma variável aleatória?

Motivação

- Sejam $Y_1, Y_2, \ldots, Y_n$ variáveis aleatórias (v.a) independentes e idênticamente distribuídas (iid) com $E(Y_i) = \mu$ e $V(Y_i) = \sigma^2$.
- Objetivo: Baseado em um conjunto de realizações de Y_i *estimar* o valor μ.
- Estratégia simples: Usar a versão empírica de μ, i.e.

$$\hat{\mu} = \sum_{i=1}^{n} \frac{Y_i}{n}.$$

- $\hat{\mu}$ é uma variável aleatória?
- Qual a esperança de $\hat{\mu}$?

Motivação

- Sejam $Y_1, Y_2, \ldots, Y_n$ variáveis aleatórias (v.a) independentes e idênticamente distribuídas (iid) com $E(Y_i) = \mu$ e $V(Y_i) = \sigma^2$.
- Objetivo: Baseado em um conjunto de realizações de Y_i *estimar* o valor μ.
- Estratégia simples: Usar a versão empírica de μ, i.e.

$$\hat{\mu} = \sum_{i=1}^{n} \frac{Y_i}{n}.$$

- $\hat{\mu}$ é uma variável aleatória?
- Qual a esperança de $\hat{\mu}$?
- Qual a variância de $\hat{\mu}$?

Motivação

- Sejam $Y_1, Y_2, \ldots, Y_n$ variáveis aleatórias (v.a) independentes e idênticamente distribuídas (iid) com $E(Y_i) = \mu$ e $V(Y_i) = \sigma^2$.
- Objetivo: Baseado em um conjunto de realizações de Y_i *estimar* o valor μ.
- Estratégia simples: Usar a versão empírica de μ, i.e.

$$\hat{\mu} = \sum_{i=1}^{n} \frac{Y_i}{n}.$$

- $\hat{\mu}$ é uma variável aleatória?
- Qual a esperança de $\hat{\mu}$?
- Qual a variância de $\hat{\mu}$?
- Qual é a distribuição de probabilidade de $\hat{\mu}$?

Sequência de v.a

- Definição 1 (Sequência de v.a's): Sejam $X_1, X_2, \ldots, X_n$ v.a's (v.a) com função de probabilidade (fp) ou função densidade de probabilidade (fdp) $f(X, \theta)$. Para uma função $g(\cdot)$ define-se uma sequência de v.a's como,

$$
\begin{aligned}
Y_1 &= g(X_1) \\
Y_2 &= g(X_1, X_2) \\
&\vdots \\
Y_n &= g(X1, \ldots, X_n).
\end{aligned}
$$

- Estatísticas amostrais são funções do tamanho da amostra e podem ser tratadas como sequência de v.a's.
- Exemplos: Média e variância amostral.

Convergência em probabilidade

- Definição 2 (Convergência em probabilidade):
- Seja $Y_1, \ldots, Y_n$ uma sequência de v.a's. Dizemos que Y_n converge em probabilidade para uma constante ou v.a Y, se $\forall \epsilon > 0$, tem-se

$$\lim_{n \to \infty} P(|Y_n - Y| > \epsilon) = 0.$$

- Notação: $Y_n \xrightarrow{P} Y$.
- Observações:
  1. $Y_n \xrightarrow{P} Y$ é equivalente a $(Y_n - Y) \xrightarrow{P} 0$.
  2. Para vetores aleatórios, tem-se

$$\boldsymbol{Y}_n = (Y_{n1}, \ldots, Y_{nk})^\top, \boldsymbol{Y}_n \xrightarrow{P} \boldsymbol{Y} \quad \text{se}$$

$$Y_{ni} \xrightarrow{P} \boldsymbol{Y}_i \quad \text{para} \quad i = 1, \ldots, n.$$

Desigualdades

- Markov: Seja Y uma v.a não negativa com $E(Y) = \mu$ e $\epsilon > 0$. Então,

$$P(Y \geq \epsilon) \leq \frac{E(Y)}{\epsilon}.$$

- Demonstração.
- Chebyshev: Seja Y qualquer v.a com $E(Y) = \mu$ e $Var(Y) = \sigma^2$ ambos finitos. Então, para todo $\epsilon > 0$

$$P(|Y - \mu| > \epsilon) \leq \frac{Var(Y)}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}.$$

- Demonstração.
- Corolário: Seja $\psi(\cdot)$ uma função monotônica, então

$$P(\psi(|Y|) > \psi(\epsilon)) \leq \frac{E(\psi(Y))}{\psi(\epsilon)}.$$

Leis dos Grandes Números

Lei dos Grandes Números: Chebyshev

- Teorema 1: Seja $Y_1, \ldots, Y_n$ v.a. iid com $E(Y_i) = \mu$ e $Var(Y_i) = \sigma^2$ ambos finitos. Então,

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{P} E(Y_i) = \mu.$$

- Demonstração.
- Ilustração computacional.

Lei dos Grandes Números: Kolmogorov

- Teorema 2: Seja $Y_1, \ldots, Y_n$ v.a. iid com $E(|Y_i|) < \infty$ e $E(Y_i) = \mu$. Então,
$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{P} E(Y_i) = \mu.$$
- Condição $E(|Y_i|) < \infty$ significa que $E(|Y_i|) = \int |y_i| f(y_i, \theta) dy_i < \infty$. Controla as caudas da distribuição.
- Caudas não podem ser muito pesadas, tal que $E(|Y_i|) = \infty$, mas ainda podem ser pesadas a ponto de $E(Y_i^2) = \infty$.
- Kolmogorov não requer que $Var(Y_i)$ exista.
- Kolmororov cobre distribuições com caudas pesadas como a t-Student.
- Demonstração é trabalhosa em sua forma geral.
- Ilustração computacional.

Lei dos Grandes Números: Markov

- Teorema 3: Seja $Y_1, \ldots, Y_n$ uma amostra não correlacionada com médias finitas $E(Y_i) = \mu_i$ e variâncias uniformemente limitadas $Var(Y_i) = \sigma_i^2 \leq M < \infty$ para $i = 1, \ldots, n$. Então,

$$\bar{Y} - \bar{\mu} = \frac{1}{n}\sum_{i=1}^{n} Y_i - \frac{1}{n}\sum_{i=1}^{n} \mu_i = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \mu_i) \xrightarrow{P} 0.$$

- LGN de Markov não requer amostras iid.
- Relaxar iid observações requer uma restrição mais forte na variância da v.a.
- Demonstração.
- Ilustração computacional.

Manipulação de limites e convergência em probabilidade

- Teorema 4: Teorema de Slutsky

  Sejam Y_n e Z_n sequências de v.a's e sejam b, c e d constantes.

  a) Se $Y_n \xrightarrow{P} c$ então $bY_n \xrightarrow{P} bc$.
  b) Se $Y_n \xrightarrow{P} c$ e $Z_n \xrightarrow{P} d$ então $Y_n + Z_n \xrightarrow{P} c + d$.
  c) Se $Y_n \xrightarrow{P} c$ e $Z_n \xrightarrow{P} d$ então $\frac{Y_n}{Z_n} \xrightarrow{P} \frac{c}{d}$ desde que $d \neq 0$ e $Y_n Z_n \xrightarrow{P} cd$.
  d) Se $Y_n \xrightarrow{P} c$ e $h(\cdot)$ é uma função contínua então $h(Y_n) \xrightarrow{P} h(c)$.

Teorema do Limite Central

Convergência em distribuição

- Seja $Y_1, \ldots, Y_n$ uma sequência de v.a. Dizemos que Y_n converge em distribuição para uma v.a Y e escrevemos $Y_n \xrightarrow{D} Y$ se

$$F_{Y_n}(y) = P(Y_n \leq y) \to F_Y(y) = P(Y \leq y), \quad n \to \infty.$$

- Na maioria das aplicações Y será normal ou qui-quadrado.
- Convergência em distribuição em geral é demonstrada via Teorema Central do Limite.
- Se n é grande usamos a distribuição de Y como uma aproximação para a distribuição de Y_n.
- Distribuição exata de Y_n em geral é dificil de se obter.

Função geradora de momentos

- A função geradora de momentos da v.a Y é $M_Y(t) = E(e^{tY})$, ou seja
$$M_Y(t) = \int_{-\infty}^{\infty} e^{ty} f(y) dy.$$
- Diferenciando $M_Y(t)$, temos
$$\begin{aligned} M_Y^{(r)}(t) &= \frac{d^r}{dt^r} \int_{-\infty}^{\infty} e^{ty} f(y) dy \\ &= \int_{-\infty}^{\infty} \left(\frac{d^r}{dt^r} e^{ty} f(y) \right) dy = \int_{-\infty}^{\infty} y^r e^{ty} f(y) dy. \end{aligned}$$
- Fazendo $t = 0$, temos
$$\int_{-\infty}^{\infty} y^r e^{ty} f(y) dy = E(Y^r).$$

Função geradora de momentos

- Proposição 1: Se X e Y são v.a independentes com fgm $M_X(t)$ e $M_Y(t)$. Seja $Z = X + Y$, então $M_Z(t) = M_X(t)M_Y(t)$.
- Proposição 2: Se X é uma v.a com fgm $M_X(t)$ e $Y = a + bX$, então $M_Y(t) = e^{at}M_X(bt)$.
- Teorema da Continuidade de Levy: Seja F_n uma sequência de funções de distribuições acumuladas com correspondentes funções geradoras de momentos $M_n(t)$. Seja F uma cdf com fgm $M(t)$. Se $M_n(t) \to M(t) \quad \forall t$ em um intervalo aberto contendo zero, então $F_n(y) \to F_Y(y) \quad \forall y$.
- Função geradora de momentos distribuição normal padrão $e^{t^2/2}$.

Teorema Central do Limite

- Teorema Lindeberg-Levy: Seja $Y_1, \ldots, Y_n$ uma amostra iid com $E(Y_i) = \mu$ e $V(Y_i) = \sigma^2 < \infty$. Então,

$$\sqrt{n}\left(\frac{\bar{Y} - \mu}{\sigma}\right) \xrightarrow{D} Z \sim N(0, 1), \quad \text{para} \quad n \rightarrow \infty.$$

- Isso significa que, para todo $y \in \Re$,

$$P(Y_n \leq y) \rightarrow \Phi(y) \quad \text{quando} \quad n \rightarrow \infty,$$

onde

$$\Phi(y) = \int_{-\infty}^{y} \phi(z)dz \quad \text{e} \quad \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right).$$

- Forma alternativa: $\bar{Y} \sim N(\mu, \sigma^2/n)$.

Teorema Central do Limite (versão multivariada)

- Teorema Lindeberg-Levy multivariado: Sejam $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ vetores aleatórios $p-$dimensionais com $E(\boldsymbol{Y}_i) = \boldsymbol{\mu}$ e $V(\boldsymbol{Y}_i) = [(\boldsymbol{Y}_i - \boldsymbol{\mu})(\boldsymbol{Y}_i - \boldsymbol{\mu})^\top] = \Sigma$, onde Σ é não-singular. Denote $\Sigma^{-1} = \Sigma^{-\frac{1}{2}}\Sigma^{-\frac{1}{2}}$. Então,

$$\sqrt{n}\Sigma^{-\frac{1}{2}}(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}) \xrightarrow{D} \boldsymbol{Z} \sim N(\boldsymbol{0}, \boldsymbol{I}),$$

  onde $N(\boldsymbol{0}, \boldsymbol{I})$ denota uma distribuição normal multivariada com média zero e matriz de covariância identidade,

$$f(\boldsymbol{z}) = (2\pi)^{-p/2} \exp\left(-\frac{1}{2}\boldsymbol{z}^\top \boldsymbol{z}\right).$$

- Formas alternativas

$$\sqrt{n}(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}) \sim N(\boldsymbol{0}, \Sigma)$$

$$\bar{\boldsymbol{Y}} \sim N(\boldsymbol{\mu}, n^{-1}\Sigma).$$

Teorema Central do Limite

- Teorema Lindeberg-Feller: Seja $Y_1, \ldots, Y_n$ v.a independentes $E(Y_i) = \mu_i$ e $V(Y_i) = \sigma_i^2 < \infty$. Defina, $\bar{\mu}_n = n^{-1} \sum_{i=1}^{n} \mu_i$ e $\bar{\sigma}_n^2 = n^{-1} \sum_{i=1}^{n} \sigma_i^2$. Suponha que

$$\lim_{n \to \infty} \max_i \frac{\sigma_i^2}{n \bar{\sigma}_n^2} = 0$$

$$\lim_{n \to \infty} \bar{\sigma}_n^2 = \bar{\sigma}^2 < \infty.$$

- Então

$$\sqrt{n} \left(\frac{\bar{Y} - \bar{\mu}_n}{\bar{\sigma}_n^2} \right) \xrightarrow{D} Z \sim N(0, 1).$$

- De forma equivalente,

$$\sqrt{n}(\bar{Y} - \bar{\mu}_n) \xrightarrow{D} N(0, \bar{\sigma}_n^2).$$

Teorema Central do Limite

- Teorema Liapounov: Sejam $Y_1, \ldots, Y_n$ v.a independentes $E(Y_i) = \mu_i$ e $V(Y_i) = \sigma_i^2 < \infty$. Suponha que

$$E[|Y_i - \mu_i|^{2+\delta}] \leq M < \infty$$

para algum $\delta > 0$. Se $\bar{\sigma}_n^2$ é positiva e finita para todo n suficientemente grande, então

$$\sqrt{n}\left(\frac{\bar{Y} - \bar{\mu}_n}{\bar{\sigma}_n}\right) \xrightarrow{D} Z \sim N(0, 1).$$

- Equivalentemente,

$$\sqrt{n}(\bar{Y} - \bar{\mu}_n) \xrightarrow{D} N(0, \bar{\sigma}^2).$$

- Liapounov é equivalente a Lindeberg-Feller só que mais simples de entender e verificar na prática.

Resultados para manipular TCL's

- Teorema Slutsky: Sejam Y_n e Z_n sequência de v.a tais que $Y_n \xrightarrow{D} Y$ e $Z_n \xrightarrow{P} c$, onde Y é uma v.a e c é uma constante. Então, os seguintes resultados valem quando $n \to \infty$:
  1. $Z_n Y_n \xrightarrow{D} cY$.
  2. $\frac{Y_n}{Z_n} \xrightarrow{D} \frac{Y}{c}$ desde que $c \neq 0$.
  3. $Y_n + Z_n \xrightarrow{D} Y + c$.
- Exercício: Seja $Y_1, \ldots, Y_n$ v.a iid com $E(Y_i) = \mu$ e $V(Y_i) = \sigma^2 < \infty$.
  a) Mostre que $\bar{\sigma}^2 \xrightarrow{P} \sigma^2$ onde $\bar{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \mu)^2$.
  b) Obtenha a distribuição aproximada de $\bar{\sigma}^2$.