Inferência Estatística II

Prof. Fernando de Souza Bastos
fernando.bastos@ufv.br

Departamento de Estatística
Programa de Pós-Graduação em Estatística Aplicada e Biometria
Universidade Federal de Viçosa
Campus UFV - Viçosa

Sumário

Teorema 1

Se $X_n \xrightarrow{P} X$, *então* $X_n \xrightarrow{D} X$.

Teorema 1

Se $X_n \xrightarrow{P} X$, *então* $X_n \xrightarrow{D} X$.

Demonstração do Teorema 1

Seja x um ponto de continuidade de $F_X(x)$, a função de distribuição acumulada (FDA) de X. Queremos mostrar que $F_{X_n}(x) \to F_X(x)$ à medida que $n \to \infty$, onde $F_{X_n}(x)$ é a FDA de X_n.

Demonstração do Teorema 1

Dividimos o evento $\{X_n \leq x\}$ em dois subconjuntos: um onde $|X_n - X| < \varepsilon$ e outro onde $|X_n - X| \geq \varepsilon$. Assim, podemos reescrever:

$$
\begin{aligned}
F_{X_n}(x) &= P(X_n \leq x) \\
&= P(\{X_n \leq x\} \cap \{|X_n - X| < \varepsilon\}) \\
&+ P(\{X_n \leq x\} \cap \{|X_n - X| \geq \varepsilon\}) \\
&\leq P(X \leq x + \varepsilon) + P(|X_n - X| \geq \varepsilon)
\end{aligned}
$$

Essa é uma decomposição da probabilidade em duas partes: uma onde X_n está "perto" de X (a diferença é menor que ε) e outra onde X_n está "longe" de X (a diferença é maior ou igual a ε).

Demonstração do Teorema 1

- Quando $|X_n - X| < \varepsilon$, sabemos que X_n está perto de X, então $X_n \leq x$ implica que $X \leq x + \varepsilon$.

Demonstração do Teorema 1

- Quando $|X_n - X| < \varepsilon$, sabemos que X_n está perto de X, então $X_n \leq x$ implica que $X \leq x + \varepsilon$.
- O segundo termo, $P(\{X_n \leq x\} \cap \{|X_n - X| \geq \varepsilon\})$, é menor ou igual a $P(|X_n - X| \geq \varepsilon)$, que é simplesmente a probabilidade de X_n estar longe de X.

Demonstração do Teorema 1

$P(|X_n - X| \geq \varepsilon) \to 0$ conforme $n \to \infty$, pois $X_n \to X$ em probabilidade. Portanto, podemos concluir:

$$\lim_{n\to\infty} F_{X_n}(x) \leq F_X(x+\varepsilon)$$

Isso nos dá a estimativa superior (upper bound) da função de distribuição acumulada de X_n.

Demonstração do Teorema 1

Agora, para obter a **estimativa inferior**, começamos reescrevendo $P(X_n \leq x)$ utilizando o complemento:

$$P(X_n \leq x) = 1 - P(X_n > x)$$

Dividimos a probabilidade $P(X_n > x)$ em dois pedaços:

$$\begin{aligned} P(X_n > x) = {} & P(\{X_n > x\} \cap \{|X_n - X| < \varepsilon\}) \\ & + P(\{X_n > x\} \cap \{|X_n - X| \geq \varepsilon\}) \end{aligned}$$

Demonstração do Teorema 1

Como $P(\{X_n > x\} \cap \{|X_n - X| < \varepsilon\})$ é menor que $P(X > x - \varepsilon)$, podemos usar a seguinte desigualdade:

$$P(X_n > x) \leq P(X \geq x - \varepsilon) + P(|X_n - X| \geq \varepsilon)$$

- O primeiro termo, $P(X \geq x - \varepsilon)$, é a probabilidade de X ser maior ou igual a $x - \varepsilon$. Isso é uma aproximação para lidar com o fato de que X_n está próximo de X.

Demonstração do Teorema 1

Como $P(\{X_n > x\} \cap \{|X_n - X| < \varepsilon\})$ é menor que $P(X > x - \varepsilon)$, podemos usar a seguinte desigualdade:

$$P(X_n > x) \leq P(X \geq x - \varepsilon) + P(|X_n - X| \geq \varepsilon)$$

- O primeiro termo, $P(X \geq x - \varepsilon)$, é a probabilidade de X ser maior ou igual a $x - \varepsilon$. Isso é uma aproximação para lidar com o fato de que X_n está próximo de X.
- O segundo termo, $P(|X_n - X| \geq \varepsilon)$, representa a probabilidade de X_n estar distante de X (mais de ε).

Demonstração do Teorema 1

Assim, podemos expressar $P(X_n \leq x)$ como:

$$P(X_n \leq x) = 1 - P(X_n > x)$$

Substituímos o limite que encontramos para $P(X_n > x)$:

$$P(X_n \leq x) \geq 1 - P(X \geq x - \varepsilon) - P(|X_n - X| \geq \varepsilon)$$

Ou, de forma mais compacta:

$$F_{X_n}(x) \geq F_X(x - \varepsilon) - P(|X_n - X| \geq \varepsilon)$$

Demonstração do Teorema 1

Sabemos que, como $X_n \to X$ em probabilidade, temos $P(|X_n - X| \geq \varepsilon) \to 0$ conforme $n \to \infty$. Assim, no limite:

$$\lim_{n\to\infty} F_{X_n}(x) \geq F_X(x - \varepsilon)$$

Agora, combinamos as duas estimativas (superior e inferior) que obtivemos:

$$F_X(x - \varepsilon) \leq \lim_{n\to\infty} F_{X_n}(x) \leq F_X(x + \varepsilon)$$

Finalmente, fazendo $\varepsilon \to 0$, chegamos à conclusão desejada:

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

Outros Teoremas Importantes:

Teorema 2

Se $X_n \xrightarrow{D} a,$ *então* $X_n \xrightarrow{P} a,$ *a constante.*

Outros Teoremas Importantes:

Teorema 2

Se $X_n \xrightarrow{D} a,$ *então* $X_n \xrightarrow{P} a,$ *a constante.*

Teorema 3

Se $X_n \xrightarrow{D} X$ *e* $Y_n \xrightarrow{P} 0$ *então* $X_n + Y_n \xrightarrow{D} X.$

Outros Teoremas Importantes:

Teorema 2

Se $X_n \xrightarrow{D} a$, *então* $X_n \xrightarrow{P} a$, *a constante.*

Teorema 3

Se $X_n \xrightarrow{D} X$ *e* $Y_n \xrightarrow{P} 0$ *então* $X_n + Y_n \xrightarrow{D} X$.

Teorema 4

Se $X_n \xrightarrow{D} X$ *e* g *é uma função contínua no suporte de* X, *então*

$$g(X_n) \xrightarrow{D} g(X).$$

Teorema de Slutsky

Teorema 5

Sejam $X_n, \ A_n$ *e* B_n, *variáveis aleatórias com* $X_n \xrightarrow{D} X, \ A_n \xrightarrow{P} a$ *e* $B_n \xrightarrow{P} b, \ a, b$ *constantes reais. Então,*

$$A_nX_n + B_n \xrightarrow{D} aX + b.$$

Para 🏠

Exercícios 5.2.2, 5.2.3, 5.2.6, 5.2.12, 5.2.15, 5.2.17, 5.2.19 e 5.2.20

Momentos de Ordem k

Definição 1

Para $k = 1, 2, \ldots,$ *o momento de ordem* k *da variável* X *é definido por* $E(X^k)$*, desde que essa quantidade exista. Se* $E(X) = \mu < \infty$*, definimos o momento central de ordem* k *por* $E[(X - \mu)^k]$*, sempre que essa quantidade exista. De modo similar, o momento absoluto de ordem* k *da variável aleatória* X *é definido por* $E(|X|^k)$*.*

Exemplo 1

Considerando $X \sim \Gamma(\alpha, \beta)$, calcule seus momentos.

$$\begin{aligned} E(X^k) &= \int_0^\infty x^k \cdot \frac{\beta^\alpha x^{\alpha-1} \cdot e^{-\beta x}}{\Gamma(\alpha)}\, dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+k-1} \cdot e^{-\beta x}\, dx \end{aligned}$$

Sabe-se que

$$\int_0^\infty x^{p-1} \cdot e^{-qx}\, dx = \Gamma(p) \cdot q^{-p}$$

Exemplo 1

Aplicando essa propriedade à integral, obtemos:

$$\int_0^\infty x^{\alpha+k-1} \cdot e^{-\beta x}\, dx = \Gamma(\alpha + k) \cdot (\beta)^{-(\alpha+k)}$$

Substituindo na expressão original:

$$E(X^k) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \Gamma(\alpha + k) \cdot (\beta)^{-(\alpha+k)}$$

$$E(X^k) = \frac{\Gamma(\alpha + k)}{\Gamma(\alpha)} \cdot \beta^{-k} = \frac{\alpha(\alpha + 1) \cdots (\alpha + k - 1)}{\beta^k}$$

Esse é o valor esperado $E(X^k)$ em termos das funções Gama e do parâmetro β.

Exemplo 1

Observe que se $\alpha = 1$, X tem distribuição exponencial de parâmetro $\beta > 0$ e, assim,

$$E(X^k) = \frac{k!}{\beta^k}$$

Se fizermos $k = 1$, obtemos a média desta variável que é $\dfrac{1}{\beta}$.

Exemplo 2

Sabendo que $X \overset{\text{iid}}{\sim} B(n, p)$. Encontre o momento central de ordem 1 e 2 desta variável.

$$
\begin{aligned}
E(X^2) &= \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=0}^{n} [i(i-1) + i] \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=2}^{n} i(i-1) \binom{n}{i} p^i (1-p)^{n-i} + E(X) \\
&= n(n-1) \sum_{i=2}^{n} \binom{n-2}{i-2} p^i (1-p)^{n-i} + E(X)
\end{aligned}
$$

Exemplo 2

$$
\begin{aligned}
E(X^2) &= n(n-1)p^2 \sum_{j=0}^{n-2} \binom{n-2}{j} p^j (1-p)^{(n-2)-j} + E(X) \\
&= n(n-1)p^2 [p + (1-p)]^{(n-2)} + np \\
&= n^2 p^2 + np(1-p)
\end{aligned}
$$

Exemplo 2

$$\begin{aligned} E(X^2) &= n(n-1)p^2 \sum_{j=0}^{n-2} \binom{n-2}{j} p^j (1-p)^{(n-2)-j} + E(X) \\ &= n(n-1)p^2 [p + (1-p)]^{(n-2)} + np \\ &= n^2 p^2 + np(1-p) \end{aligned}$$

Assim, $Var(X) = E(X - E(X))^2 = E(X^2) - (E(X))^2 = np(1-p)$

Função Geradora de Momentos

Definição 2

A função geradora de momentos de uma variável aleatória X é definida por $M_X(t) = E(e^{tX}), \ t \in \mathbb{R}$

Função Geradora de Momentos

Definição 2

A função geradora de momentos de uma variável aleatória X é definida por $M_X(t) = E(e^{tX}), \ t \in \mathbb{R}$

Observação Importante:

O momento de ordem k de uma variável aleatória X pode ser encontrado utilizando a função geradora de momentos. Para encontrar o momento de ordem k, derivamos k vezes em relação a t a função geradora de momentos $M_X(t)$ e então avaliamos em $t = 0$:

$$\mathbb{E}[X^k] = \frac{d^k}{dt^k} M_X(t)\bigg|_{t=0}$$

Teorema 6

Seja $\{X_n\}_{n\geq 1}$ *uma sequência de variáveis aleatórias com fgm* $M_{X_n}(t)$ *que existe para* $|t| < h$ *para todo* n. *Seja* X *uma variável aleatória com fgm* $M_X(t)$, *que existe para* $|t| \leq h_1 \leq h$. *Se* $\lim_{n\rightarrow\infty} M_{X_n}(t) = M_X(t)$ *para* $|t| \leq h_1$, *então* $X_n \xrightarrow{D} X$.

Teorema 6

Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias com fgm $M_{X_n}(t)$ que existe para $|t| < h$ para todo n. Seja X uma variável aleatória com fgm $M_X(t)$, que existe para $|t| \leq h_1 \leq h$. Se $\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$ para $|t| \leq h_1$, então $X_n \xrightarrow{D} X$.

Observação importante na resolução de exercícios:

Se $\lim\limits_{n\to\infty}\left(1+\frac{b}{n}+\frac{\psi(n)}{n}\right)^{cn}$, em que b e c não dependem de n e, em que, $\lim\limits_{n\to\infty}\psi(n)=0$. Então,

$$\lim_{n\to\infty}\left(1+\frac{b}{n}+\frac{\psi(n)}{n}\right)^{cn}=\lim_{n\to\infty}\left(1+\frac{b}{n}\right)^{cn}=e^{bc}.$$

Exemplo 1

$$\lim_{n\to\infty}\left(1-\frac{t^2}{n}+\frac{t^2}{n^{3/2}}\right)^{-n/2}=\lim_{n\to\infty}\left(1-\frac{t^2}{n}+\frac{t^2/\sqrt{n}}{n}\right)^{-n/2}$$

Exemplo 1

$$\lim_{n\to\infty}\left(1-\frac{t^2}{n}+\frac{t^2}{n^{3/2}}\right)^{-n/2} = \lim_{n\to\infty}\left(1-\frac{t^2}{n}+\frac{t^2/\sqrt{n}}{n}\right)^{-n/2}$$
$$= \lim_{n\to\infty}\left(1-\frac{t^2}{n}\right)^{-n/2}$$
$$= e^{t^2/2}$$

Aqui, $b=-t^2$, $c=-\frac{1}{2}$ e $\psi(n)=\frac{t^2}{\sqrt{n}}$.

Exemplo 2

Considere $X_n \sim Binomial(n, p_n)$ e suponha $\lim_{n\to\infty} np_n = \lambda > 0$ (por exemplo, $p_n = \dfrac{1}{n+1}$, $\lim_{n\to\infty} np_n = 1$). Então, $X_n \xrightarrow{D} X$, em que $X \sim$Poisson(λ).

Exemplo 2

Considere $X_n \sim Binomial(n, p_n)$ e suponha $\lim_{n\to\infty} np_n = \lambda > 0$ (por exemplo, $p_n = \frac{1}{n+1}$, $\lim_{n\to\infty} np_n = 1$). Então, $X_n \xrightarrow{D} X$, em que $X \sim$Poisson(λ).

Demonstração

Temos que,

$$\begin{aligned}
M_{X_n}(t) = E(e^{tX_n}) &= \sum_{k=0}^{n} e^{tk}\binom{n}{k}p_n^k(1-p_n)^{n-k} \\
&= \left(1 - p_n + p_n e^t\right)^n = \left(1 + \frac{np_n}{n}(e^t - 1)\right)^n \\
\text{(para n grande)} &= \left(1 + \frac{\lambda}{n}(e^t - 1)\right)^n \xrightarrow[n\to\infty]{} \exp\left\{\lambda(e^t - 1)\right\}
\end{aligned}$$

Logo, $X_n \xrightarrow{D} X \sim$Poisson($\lambda$).

Quando a quantidade np_n se estabiliza em um valor $\lambda > 0$, estamos essencialmente controlando a média da binomial. À medida que $n \to \infty$ e p_n diminui de forma controlada, mantemos np_n constante, aproximando o comportamento da binomial ao de uma distribuição Poisson com parâmetro λ. A essência é que estamos explorando o comportamento assintótico da binomial, com p_n diminuindo à medida que n cresce, mas de modo que np_n permaneça fixo e igual a λ. Isso faz com que a média e variância da binomial "convirjam" para os parâmetros de uma Poisson.

Referências I

Hogg, RV, J McKean e AT Craig (2019). *Introduction to Mathematical Statistics*.