LEI FRACA DOS GRANDES NÚMEROS

Vanderlei da Costa Bueno

Instituto de Matemática e Estatística

Universidade de São Paulo, SP. Brasil

Setembro de 2020

Lei Fraca dos Grandes números

Consideremos a variável aleatória X que representa o valor numérico de um experimento aleatório e seja $X_1, X_2, ..., X_n$ uma amostra aleatória de X, isto é, cópias i.i.d. de X, de tamanho n, grande.

A **Lei dos Grandes Números** afirma que a média aritmética dos n valores observados é aproximadamente igual à média de $X, \mu = E[X]$, quando n é grande, isto é

$$\overline{X_n}(\omega) = \frac{\sum_{i=1}^{n} X_i(\omega)}{n} \rightarrow_{n\rightarrow\infty} \mu$$

onde $\omega = (\omega_1, ..., \omega_n), X_n(\omega) = X(\omega_n)$, ω_n são os ensaios sucessivos e w é o experimento composto.

Lei Fraca dos Grandes números

Exemplo 1

Considere que nas repetições independentes de um experimento estamos interessados nas ocorrências de um evento A .
Definimos a sequência de variáveis aleatórias $(X_n)_{n\geq 1}$ $X_n = 1$, se A ocorreu na n-ésima realização;

$X_n = 0 \;\; c.c.$

Desta maneira $P(X_n = 1) = P(A)$. Na n-ésima repetição do experimento calculamos a frequência relativa do evento A, isto é

$$F_n = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$\lim_{n\to\infty} F_n = E[X_1] = P(A).$$

O exemplo justifica o conceito frequêntista de probabilidade.

Lei Fraca dos Grandes números

Se a convergência é quase certa dizemos que a Lei é Forte, se a convergência é em probabilidade dizemos que a Lei é Fraca. Claramente, a Lei Forte implica a Lei Fraca dos Grandes Números.

Para provarmos a Lei Fraca dos Grandes Números usaremos a Desigualdade de Markov:

Lema 1

Seja X uma variável aleatória. Para todo ε e r, números reais positivos com $E[|X|^r] < \infty$, temos

$$P(|X| \geq \varepsilon) < \frac{E[|X|^r]}{\varepsilon^r}.$$

Prova

$$E[|X|^r] = \int_{-\partial}^{\infty} |x|^r dF(x) = \int_{\{|X|>\varepsilon\}} |x|^r dF(x) + \int_{\{|X|\leq\varepsilon\}} |x|^r dF(x) \geq$$

$$\int_{\{|X|>\varepsilon\}} |x|^r dF(x) \geq |\varepsilon|^r \int_{\{|X|>\varepsilon\}} dF(x) = |\varepsilon|^r P(|X| \geq \varepsilon).$$

Lei Fraca dos Grandes números

Corolário - Desigualdade de Chebyshev
Se X uma variável aleatória com média $E[X] = \mu$ e variância $\sigma^2 < \infty$, $\forall \varepsilon > 0$ temos

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

Prova Segue da desigualdade de Markov

Lei Fraca dos Grandes números

Teorema 1 Lei Fraca dos Grandes Números de Chebyshev
Se $(X_n)_{n\geq 1}$ é uma sequência de variáveis aleatórias independentes com médias μ e variâncias $\sigma^2 < \infty$, comuns, então

$$\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n} \rightarrow^{P} \mu.$$

Prova Segue da desigualdade de Chebyshev, desde que $E[\frac{\sum_{i=1}^{n} X_i}{n}] = \mu$ e $Var(\frac{\sum_{i=1}^{n} X_i}{n}) = \frac{\sigma^2}{n}$.

Teorema 2 Lei Fraca dos Grandes Números para variáveis aleatórias não correlacionadas.
Se $(X_n)_{n\geq 1}$ é uma sequência de variáveis aleatórias não correlacionadas, isto é, $cov(X_i, X_j) = 0, \forall i, j$, com médias μ e variâncias $\sigma^2 < \infty$, comuns, então

$$\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n} \rightarrow^{P} \mu.$$

Lei Fraca dos Grandes números

Prova

Sabemos que $E[\frac{\sum_{i=1}^{n} X_i}{n}] = \mu$.

$$Var(\overline{X_n}) = E[(\overline{X_n} - \mu)^2] = E[(\frac{\sum_{i=1}^{n} X_i}{n} - \mu)^2] = E[(\frac{\sum_{i=1}^{n} X_i - n\mu}{n})^2] =$$

$$\frac{1}{n^2} E[(\sum_{i=1}^{n} X_i - n\mu)^2] = \frac{1}{n^2} E\{[\sum_{i=1}^{n} (X_i - \mu)].[\sum_{j=1}^{n} (X_j - \mu)]\} =$$

Lei Fraca dos Grandes números

$$\frac{1}{n^2}E[\sum_{i=1}^{n}(X_i-\mu)^2+\sum_{i<j}^{n}(X_i-\mu).(X_j-\mu)]=\frac{1}{n^2}\{\sum_{i=1}^{n}E[(X_i-\mu)^2]+$$

$$\sum_{i<j}^{n}E[(X_i-\mu).(X_j-\mu)]\}=\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2+0=\frac{\sigma^2}{n}.$$

Assim, pela Desigualdade de Chebyshev

$$P(|\overline{X_n}-\mu|>\varepsilon)=P((\overline{X_n}-\mu)^2>\varepsilon^2)\leq\frac{E[(\overline{X_n}-\mu)^2]}{\varepsilon^2}=\frac{\sigma^2}{n\varepsilon^2}\rightarrow 0$$

quando $n\rightarrow\infty$.

Lei Fraca dos Grandes números

Na realidade, somente é necessário que os X_i's sejam não correlacionados assintóticamente no sentido: $\lim_{|i-j|\to\infty} cov(X_i, X_j) = 0$.

Teorema 3 Lei Fraca dos Grandes Números para variáveis aleatórias assintóticamente não correlacionadas.
Seja $(X_n)_{n\geq 1}$ uma sequência de variáveis aleatórias assintóticamente não correlacionadas, com média comum μ e variâncias $Var(X_i) = \sigma_i^2 < c$, uniformemente limitadas por uma constante c. Então

$$\overline{X_n} = \frac{\sum_{i=1}^n X_i}{n} \to^P \mu.$$

Lei Fraca dos Grandes números

Prova
Pela desigualdade de Chebyshev temos

$$P(|\overline{X_n} - \mu| > \varepsilon) \leq \frac{Var(\overline{X_n})}{\varepsilon^2}, \forall \varepsilon > o.$$

Portanto é suficiente provar que $\lim_{n\to\infty} Var(\overline{X_n}) = 0$.
Pela desigualdade de Schwarz temos que $\forall i, j$,

$$|cov(X_i, X_j)| = |E[(X_i - \mu).(X_j - \mu)]| \leq$$

$$\sqrt{E[(X_i - \mu)^2].E[(X_j - \mu)^2]} = \sigma_i \sigma_j \leq c.$$

Como $\lim_{|i-j|\to\infty} cov(X_i, X_j) = 0$ temos que

$$\forall \delta > 0, \quad \exists m \in \mathbb{N} \ \mid \ se \ \ |i-j| > m \rightarrow |cov(X_i, X_j)| < \frac{\delta}{2}.$$

Lei Fraca dos Grandes números

Sabemos que

$$Var(\overline{X_n}) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} cov(X_i, X_j) \leq$$

$$\frac{1}{n^2}[\sum_{|i-j|\leq m} nc + \sum_{|i-j|>m} |cov(X_i, X_j)|] \leq$$

$$\frac{n.m.c}{n^2} + \frac{n(n-m)\delta}{2n^2} \leq \delta, \quad para \quad n \geq \frac{2mc}{\delta}.$$

Lei Fraca dos Grandes números

Teorema 4 Lei Fraca dos Grandes Números de Khintchine, para variáveis aleatórias independentes e identicamente distribuidas com média finita.
Seja $(X_n)_{n\geq 1}$ uma sequência de variáveis aleatórias i.i.d. com $E[X_1] = \mu < \infty$.Então

$$\overline{X_n} = \frac{\sum_{i=1}^{n} X_i}{n} \to^{P} \mu.$$

Prova

Devemos provar que $\forall \varepsilon > 0$ e $\forall \delta > 0$, $P(|\overline{X_n} - \mu| > \varepsilon) \leq \delta$.
Sejam $\gamma = E[|X|]$ e $\alpha = \frac{\delta.\varepsilon^2}{12.\gamma}$. Defina a sequência de variáveis aleatórias $(Y_n)_{n \geq 1}$

$$Y_n = X_n \quad se \quad -\alpha.n \leq X_n \leq \alpha.n \quad e \quad 0 \quad c.c.$$

Então os Y_n's são independentes com médias $\mu_n = \int_{-\alpha.n}^{\alpha.n} ydF(y)$ e variância $\sigma_n^2 = \int_{-\alpha.n}^{\alpha.n} y^2 dF(y) - \mu_n^2$.
Claramente, se $n \to \infty$, $\alpha.n \to \infty$ e $\mu_n \to \mu$,isto é

$$\exists n_0 \;\mid\; \forall n \geq n_0, \to |\mu_n - \mu| \leq \frac{\varepsilon}{2}.$$

Lei Fraca dos Grandes números

Portanto

$$P(|\overline{X_n} - \mu| > \varepsilon) = P(|\overline{X_n} - \mu_n + \mu_n - \mu| > \varepsilon) \leq P(|\mu_n - \mu| > \frac{\varepsilon}{2}) +$$

$$P(|\overline{X_n} - \mu_n| > \frac{\varepsilon}{2}) \leq P(|\overline{X_n} - \mu_n| > \frac{\varepsilon}{2}) \leq$$

$$P(|\overline{X_n} - \mu_n| > \frac{\varepsilon}{2}, \overline{X_n} = \overline{Y_n}) + P(|\overline{X_n} - \mu_n| > \frac{\varepsilon}{2}, \overline{X_n} \neq \overline{Y_n}) \leq$$

$$P(|\overline{Y_n} - \mu_n| > \frac{\varepsilon}{2}) + P(\overline{X_n} \neq \overline{Y_n}).$$

Note que

$$\sigma_n^2 \leq \int_{-\alpha.n}^{\alpha.n} y^2 dF(y) \leq \alpha.n \int_{-\alpha.n}^{\alpha.n} |y| dF(y) \leq \alpha.n.\gamma$$

e pela desigualdade de Tchebyshev temos

$$P(|\overline{Y_n} - \mu_n| > \frac{\varepsilon}{2}) \leq \frac{4\sigma_n^2}{n\varepsilon^2} = \frac{4.\alpha.\gamma}{\varepsilon^2} = \frac{\delta}{3}, \quad para \quad \alpha = \frac{\delta.\varepsilon^2}{12.\gamma}.$$

Lei Fraca dos Grandes números

Em adição temos

$$P(\overline{X_n} \neq \overline{Y_n}) = \sum_{k=1}^{n} P(X_k \neq Y_k) = n.P(|X_1| \geq n.\alpha)$$

$$\leq n \int_{|X|>\alpha.n} dF(x) \leq \alpha^{-1} \int_{|X|>\alpha.n} |x| dF(x) \to 0$$

quando $n \to \infty$.
Assim $\exists n_1 \ \mid \ P(\overline{X_n} \neq \overline{Y_n}) \leq \frac{\delta}{3}, \ \ se \ \ n \geq n_1$.
Tomando $n \geq \max\{n_0, n_1\}$, temos $P(|\overline{X_n} - \mu| > \varepsilon) \leq \delta$.