Chapter 7

Curs 7

7.1 Legea slabă a numerelor mari

Reamintim că două variabile aleatoare X şi Y se numesc necorelate dacă $M\left(X^2\right)$, $M\left(Y^2\right) < +\infty$ şi are loc egalitatea

$$M(XY) = M(X)\,M(Y)$$

(în particular, această egalitate are loc dacă variabilele aleatoare sunt independente).

Definiţia 7.1.1 *Spunem că o familie de variabile aleatoare $(X_i)_{i\in I}$ sunt necorelate dacă $M(|X_i|) < +\infty$, $i \in I$ şi are loc*

$$M(X_i X_j) = M(X_i)\,M(X_j)\,,$$

oricare ar fi $i, j \in I$ cu $i \neq j$.

Are loc următoarea:

Lema 7.1.2 *Dacă $X_1, \ldots, X_n$ sunt variabile aleatoare necorelate, atunci*

$$\sigma^2(X_1 + \ldots + X_n) = \sigma^2(X_1) + \ldots + \sigma^2(X_n)$$

(adică dispersia sumei unor variabile aleatoare necorelate este egală cu suma dispersiilor variabilelor aleatoare respective).

Demonstraţie. Fie $\mu_\iota = M(X_i)$ şi să notăm cu $S_n = X_1 + \ldots + X_n$. Cum $M(S_n) = \sum_{i=1}^n M(X_i) = \sum_{i=1}^n \mu_i$, avem

$$\begin{aligned}
\sigma^2(S_n) &= M\left((S_n - M(S_n))^2\right) = M\left(\left(\sum_{i=1}^n X_i - \mu_i\right)^2\right) \\
&= \sum_{i=1}^n M\left((X_i-\mu_i)^2\right) + 2\sum_{1\le i<j\le n} M\left((X_i-\mu_i)(X_j-\mu_j)\right) \\
&= \sum_{i=1}^n \sigma^2(X_i) + 2\sum_{1\le i<j\le n} M\left((X_i-\mu_i)(X_j-\mu_j)\right).
\end{aligned}$$

Pentru a demonstra afirmaţia, arătăm că cea de-a doua sumă este egală cu zero. Pentru aceasta, observăm că

$$\begin{aligned}
M\left((X_i-\mu_i)(X_j-\mu_j)\right) &= M(X_iX_j) - \mu_i M(X_j) - \mu_j M(X_i) - \mu_i\mu_j \\
&= M(X_iX_j) - \mu_i\mu_j \\
&= 0,
\end{aligned}$$

deoarece variabilele aleatoare $X_1, \ldots, X_n$ sunt necorelate, încheiând demonstraţia. ■

O primă variantă a legii numerelor mari este următoarea:

Teorema 7.1.3 *Fie $X_1, X_2, \ldots$ un şir de variabile aleatoare necorelate cu medii $M(X_i) = \mu$ şi dispersii $\sigma^2(X_i) \le C < +\infty$, $i \in I$. Atunci*

$$\frac{S_n}{n} = \frac{X_1 + \ldots + X_n}{n} \underset{n\to\infty}{\to} \mu$$

în L^2 şi în probabilitate.

Demonstraţie. Să observăm că $M\left(\frac{S_n}{n}\right) = M\left(\frac{X_1+\ldots+X_n}{n}\right) = \mu$, şi folosind lema anterioară obţinem:

$$\begin{aligned}
M\left(\left(\frac{S_n}{n} - \mu\right)^2\right) &= \sigma^2\left(\frac{S_n}{n}\right) \\
&= \frac{1}{n^2}\sigma^2(S_n) \\
&= \frac{1}{n^2}\left(\sigma^2(X_1) + \ldots + \sigma^2(X_n)\right) \\
&\le \frac{1}{n^2}Cn \\
&= \frac{C}{n} \to 0,
\end{aligned}$$

adică $\frac{S_n}{n}$ converge la μ în L^2.

Pentru a demonstra convergenţa în L^2, din inegalitatea lui Cebâşev obţinem

$$P\left(\left|\frac{S_n}{n}-\mu\right|\geq\varepsilon\right)\leq\frac{1}{\varepsilon^2}\sigma^2\left(\frac{S_n}{n}\right)\to 0,$$

pentru $n\to\infty$, oricare ar fi $\varepsilon>0$, şi deci $\frac{S_n}{n}$ converge la μ în probabilitate. ■

Un caz particular important al teoremei anterioare este cazul în care variabilele aleatoare $X_1, X_2, \ldots$ sunt independente (şi deci sunt necorelate) şi au aceeaşi distribuţie. În acest caz, teorema anterioară afirmă că dacă $M\left(X_i^2\right)<+\infty$, atunci $S_n/n\to\mu=M\left(X_i\right)$ în L^2 şi în probabilitate. Pentru a generaliza rezultatul de mai sus în cazul în care $M\left(|X_i|\right)<+\infty$, demonstrăm mai întâi următoarea:

Teorema 7.1.4 *Pentru fiecare $n\geq 1$, fie $X_{n,1}, X_{n,2}, \ldots, X_{n,n}$ variabile aleatoare independente. Fie $(b_n)_{n\geq 1}$ un şir crescător de numere reale cu $b_n\nearrow+\infty$, şi să notăm cu $\bar{X}_{n,k}=X_{n,k}1_{\{|X_{n,k}|\leq b_n\}}$ şi $a_n=\sum_{k=1}^n M\left(\bar{X}_{n,k}\right)$. Dacă*

i) $\sum_{k=1}^n P\left(|X_{n,k}|>b_n\right)\underset{n\to\infty}{\to}0$;

ii) $\frac{1}{b_n^2}\sum_{k=1}^n M\left(\bar{X}_{n,k}^2\right)\underset{n\to\infty}{\to}0$,

atunci

$$\frac{S_n-a_n}{b_n}=\frac{X_{n,1}+\ldots X_{n,n}-a_n}{b_n}\underset{n\to\infty}{\overset{\mathcal{P}}{\to}}0.$$

Demonstraţie. Să notăm $\bar{S}_n=\bar{X}_{n,1}+\ldots+\bar{X}_{n,n}$. Pentru $\varepsilon>0$ arbitrar fixat, avem:

$$\begin{aligned}
P\left(\left|\frac{S_n-a_n}{b_n}\right|>\varepsilon\right) &\leq P\left(\left|\frac{S_n-a_n}{b_n}\right|>\varepsilon, S_n\neq\bar{S}_n\right)+P\left(\left|\frac{S_n-a_n}{b_n}\right|>\varepsilon, S_n=\bar{S}_n\right)\\
&\leq P\left(S_n\neq\bar{S}_n\right)+P\left(\left|\frac{\bar{S}_n-a_n}{b_n}\right|>\varepsilon\right)\\
&\leq P\left(\cup_{k=1}^n\left\{X_{n,k}\neq\bar{X}_n^k\right\}\right)+P\left(\left|\bar{S}_n-a_n\right|>\varepsilon b_n\right)\\
&\leq \sum_{k=1}^n P\left(X_{n,k}\neq\bar{X}_{n,k}\right)+\frac{M\left(\left(\bar{S}_n-a_n\right)^2\right)}{\left(\varepsilon b_n\right)^2}\\
&= \sum_{k=1}^n P\left(|X_{n,k}|>b_n\right)+\frac{\sigma^2\left(\bar{S}_n\right)}{\varepsilon^2 b_n^2}.
\end{aligned}$$

Cum variabilele aleatoare $X_{n,1},\ldots,X_{n,n}$ sunt independente, variabilele aleatoare $X_{n,1},\ldots,X_{n,n}$ sunt de asemena independente, şi deci avem:

$$\begin{aligned}
\sigma^2\left(\bar{S}_n\right) &= \sigma^2\left(\bar{X}_{n,1}+\ldots+\bar{X}_{n,n}\right)\\
&= \sigma^2\left(\bar{X}_{n,1}\right)+\ldots+\sigma^2\left(\bar{X}_{n,n}\right)\\
&\leq M\left(\bar{X}_{n,1}^2\right)+\ldots+M\left(\bar{X}_{n,n}^2\right).
\end{aligned}$$

Obţinem deci

$$P\left(\left|\frac{S_n - a_n}{b_n}\right| > \varepsilon\right) \le \sum_{k=1}^{n} P\left(|X_{n,k}| > b_n\right) + \frac{1}{\varepsilon^2 b_n^2}\sum_{k=1}^{n} M\left(\bar{X}_{n,k}^2\right),$$

de unde folosind ipoteza obţinem

$$\lim_{n\to\infty} P\left(\left|\frac{S_n - a_n}{b_n}\right| > \varepsilon\right) = 0,$$

adică

$$\frac{S_n - a_n}{b_n} \xrightarrow{\mathcal{P}} 0,$$

încheiând demonstraţia. ■

Folosind teorema anterioară, obţinem următoarea:

Teorema 7.1.5 *Fie $X_1, X_2, \ldots$ un şir de variabile aleatoare independente şi identic distribuite astfel încât*

$$nP\left(|X_i| > n\right) \underset{n\to\infty}{\to} 0,$$

şi să notăm $S_n = X_1 + \ldots + X_n$ şi $\mu_n = M\left(X_1 1_{\{|X_1|\le n\}}\right)$. Atunci

$$\frac{S_n}{n} - \mu_n \xrightarrow{\mathcal{P}} 0.$$

Demonstraţie. Să aplicăm teorema anterioară pentru $X_{n,k} = X_k$ şi $b_n = n$, $n, k \ge 1$.

Pentru aceasta, să arătăm că sunt verificate ipotezele *i)* şi *ii)* ale teoremei anterioare. Avem:

$$\sum_{k=1}^{n} P\left(|X_{n,k}| > n\right) = \sum_{k=1}^{n} P\left(|X_k| > n\right) = nP\left(|X_1| > n\right) \underset{n\to\infty}{\to} 0$$

şi

$$b_n^{-2}\sum_{k=1}^{n} M\left(\bar{X}_{n,k}^2\right) = n^{-2}\sum_{k=1}^{n} M\left(\bar{X}_{n,k}^2\right) = n^{-2}\cdot nM\left(\bar{X}_{n,1}^2\right) = n^{-1}.$$

Să arătăm în continuare că $n^{-1}M\left(\bar{X}_{n,1}^2\right) \underset{n\to\infty}{\to} 0$. Pentru aceasta, folosind Lema 7.1.6 de mai jos, avem:

$$M\left(\bar{X}_{n,1}\right) = \int_0^\infty 2xP\left(\left|\bar{X}_{n,1}\right| > x\right)dx \le \int_0^n 2xP\left(|X_1| > x\right)dx,$$

deoarece

$$\begin{aligned} P\left(\left|\bar{X}_{n,1}\right| > x\right) &= P\left(|X_1|\, 1_{\{|X_1|\le n\}} > x\right) \\ &= \begin{cases} 0, & x > n \\ P\left(|X_1| > x\right) - P\left(|X_1| > n\right), & x \le n \end{cases} \\ &\le \begin{cases} 0, & x > n \\ P\left(|X_1| > x\right), & x \le n. \end{cases} \end{aligned}$$

Să considerăm funcţia $f(x) = 2xP(|X_1| > x)$. Deoarece $0 \leq g(x) \leq 2x$ şi $\lim_{x\to\infty} g(x) = 0$ (conform ipotezei), rezultă că $M = \sup_{x\geq 0} g(x) < +\infty$. Mai mult, oricare ar fi $0 \leq K \leq \infty$ fixat, $M_K = \sup_{x>K} g(x) \leq M < +\infty$.

Pentru orice $n \geq K$ obţinem:

$$\begin{aligned} M\left(\bar{X}_{n,1}\right) &\leq \int_0^n 2xP(|X_1| > x)\,dx = \int_0^n f(x)\,dx \\ &= \int_0^K f(x)\,dx + \int_K^n f(x)\,dx \\ &\leq K \cdot M + (n-K) \cdot M_K, \end{aligned}$$

de unde, trecând la limită cu $n \to \infty$, obţinem:

$$\lim_{n\to\infty} \frac{1}{n} M\left(\bar{X}_{n,1}\right) \leq \lim_{n\to\infty} \frac{KM + (n-K)M_K}{n} = M_K,$$

oricare ar fi $K \geq 0$, şi deci trecând la limită cu $K \to \infty$, obţinem

$$\lim_{n\to\infty} \frac{1}{n} M\left(\bar{X}_{n,1}\right) \leq \lim_{K\to\infty} M_K = 0,$$

şi deci $\lim_{n\to\infty} \frac{1}{n} M\left(\bar{X}_{n,1}\right) = 0$.

Sunt deci verificate ipotezele teoremei anterioare, şi deci rezultă că

$$\frac{S_n - a_n}{b_n} = \frac{S_n - a_n}{n} \xrightarrow{\mathcal{P}} 0,$$

unde $a_n = \sum_{k=1}^n M\left(\bar{X}_{n,k}\right) = nM\left(\bar{X}_{n,1}\right) = nM\left(X_1 1_{\{|X_1| \leq n\}}\right) = n\mu_n$, adică

$$\frac{S_n}{n} - \mu_n \xrightarrow{\mathcal{P}} 0,$$

încheiând demonstraţia. ■

În demonstraţia teoremei anterioare, am utilizat următorul rezultat, util în calculul momentului unei variabile aleatoare:

Lema 7.1.6 *Dacă $X \geq 0$ este o variabilă aleatoare ne-negativă şi $p > 0$, atunci are loc egaliatatea*

$$M(X^p) = \int_0^\infty px^{p-1} P(X > x)\,dx.$$

Demonstraţie. Din teorema Fubini (pentru variabile aleatoare ne-negative), avem:

$$\begin{aligned}
\int_0^\infty px^{p-1}P(X>x)\,dx &= \int_0^\infty px^{p-1}M\left(1_{\{X>x\}}\right)dx \\
&= \int_0^\infty M\left(px^{p-1}1_{\{X>x\}}\right)dx \\
&= M\int_0^\infty px^{p-1}1_{\{X>x\}}dx \\
&= M\int_0^X px^{p-1}dx \\
&= M\left(x^p\big|_0^X\right) \\
&= M(X^p),
\end{aligned}$$

încheiând demonstraţia. ■

Cu această pregatire, putem acum demonstra varianta generală a legii slabe a numerelor mari, după cum urmează:

Teorema 7.1.7 (Legea slabă a numerelor mari) *Fie $X_1, X_2, \ldots$ variabile aleatoare independente şi identic distribuite cu medie $\mu = M(X_1) < +\infty$. Atunci*

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\mathcal{P}} \mu.$$

Demonstraţie. Din teorema convergenţei dominate, avem

$$nP(|X_1| > n) \leq M\left(|X_1|\,1_{\{|X_1|>n\}}\right) \underset{n\to\infty}{\rightarrow} 0,$$

şi

$$\mu_n = M\left(X_1 1_{\{|X_1|\leq n\}}\right) \underset{n\to\infty}{\rightarrow} M(X_1) = \mu.$$

Sunt deci verificate ipotezele teoremei anterioare, şi deci

$$\frac{S_n}{n} - \mu_n = \frac{X_1 + \ldots + X_n}{n} - \mu_n \xrightarrow{\mathcal{P}} 0\cdot$$

Cum $\mu_n \to \mu$, obţinem echivalent

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\mathcal{P}} \mu,$$

încheiând demonstraţia. ■

EXERCIŢII

Exerciţiul 7.1.1 *Fie $X_1, X_2, \ldots$ variabile aleatoare necorelate cu medii $M(X_n) = \mu_n$, $n = 1, 2, \ldots$ astfel încât $\frac{\sigma^2(X_n)}{n} \underset{n\to\infty}{\to} 0$. Să se arate că*

$$\frac{S_n}{n} - \nu_n \to 0$$

în L^2 şi în probabilitate, unde $S_n = X_1 + \ldots + X_n$ şi $\nu_n = \frac{1}{n} M(S_n)$.

Exerciţiul 7.1.2 *Să se arate că dacă $X_1, X_2, \ldots$ sunt variabile aleatoare cu medii $M(X_n) = 0$, $n = 1, 2, \ldots$ şi $M(X_n X_m) \leq f(n-m)$ pentru $n \geq m \geq 1$, unde $\lim_{k\to\infty} f(k) = 0$, atunci*

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\mathcal{P}} 0.$$

Exerciţiul 7.1.3 *Fie $f : [0,1] \to \mathbb{R}$ o funcţie măsurabilă astfel încât $\int_0^1 |f(x)|\, dx < +\infty$. Fie $U_1, U_2, \ldots$ un ;ir de variabile aleatoare aleatoare independente, uniform distribuite pe intervalul $[0,1]$, şi să notăm*

$$I_n = \frac{f(U_1) + \ldots f(U_n)}{n}.$$

i) *Să se arate că $I_n \xrightarrow{\mathcal{P}} I = \int_0^1 f(x)\, dx$;*

ii) *Dacă în plus $\int_0^1 f^2(x)\, dx$, să se utilizeze inegalitatea lui Cebâşev pentru a estima $P(|I_n - I| > \varepsilon)$, $\varepsilon > 0$.*

Exerciţiul 7.1.4 *Să se arate că dacă $X \geq 0$ este o variabilă aleatoare discretă ne-negativă ce ia numai valori numere întregi, atunci*

$$M(X) = \sum_{n=0}^{\infty} P(X \geq n).$$

Să se gasească o expresie similară pentru calculul momentului de ordin doi $M(X^2)$.

Exerciţiul 7.1.5 *Să se arate că dacă $H(x) = \int_{-\infty}^{x} h(y)\, dy$, unde $h : \mathbb{R} \to [0, +\infty)$ este o funcţie integrabilă, atunci pentru orice valoare aleatoare ne-negativă are loc*

$$M(H(X)) = \int_{-\infty}^{+\infty} h(y)\, P(X \geq y)\, dy.$$

Să se particularizeze formula obţinută pentru $H(x) = x^p$, $p > 0$, respectiv $H(x) = e^{\alpha x}$, $\alpha > 0$.