

Probability Calculus

Anna Janicka

lecture XIII, 8.01.2019

LAWS OF LARGE NUMBERS – CONT.

CENTRAL LIMIT THEOREM

Plan for Today

- Laws of Large Numbers – cont.
 - Strong Law of Large Numbers
 - consequences
- Central Limit Theorem
 - de Moivre-Laplace Theorem



Weak Laws of Large Numbers – reminder

1. Weak Law of Large Numbers for the Bernoulli Scheme

Let X_1, X_2, \dots be independent with distributions $\mathbb{P}(X_n = 1) = p = 1 - \mathbb{P}(X_n = 0)$. We then have that (S_n/n) converges in probability to p ; in other words, for any $\varepsilon > 0$, we have $\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{S_n}{n} - p \right| > \varepsilon \right) = 0$.

$$S_n = X_1 + X_2 + \dots + X_n$$



Weak Laws of Large Numbers – cont. – reminder

2. Weak Law of Large Numbers for uncorrelated random variables

Let X_1, X_2, \dots be uncorrelated random variables with a common upper bound to their variances. Then, the sequence $(X_n)_{n \geq 1}$ satisfies the weak law of large numbers: $\frac{S_n - \mathbb{E}S_n}{n} \xrightarrow{\mathbb{P}} 0$, i.e. for any $\varepsilon > 0$ we have $\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{S_n - \mathbb{E}S_n}{n} \right| > \varepsilon \right) = 0$.



Strong Laws of Large Numbers

1. Strong Law of Large Numbers for the Bernoulli Scheme

Let X_1, X_2, \dots be a sequence of independent random variables, such that

$$\mathbb{P}(X_n = 1) = p = 1 - \mathbb{P}(X_n = 0), \quad n = 1, 2, \dots$$

Then, the sequence (S_n/n) converges almost surely to p ; i.e., there exists an event Ω' of measure 1 such that for any $\omega \in \Omega'$, we have

$$\lim_{n \rightarrow \infty} \frac{S_n(\omega)}{n} = p.$$



Strong Laws of Large Numbers – reminder cont.

2. Kolmogorov's Strong Law of Large Numbers

Let X_1, X_2, \dots be a sequence of independent, identically distributed integrable random variables. Then,

$$\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{a.s.} \mathbb{E}X_1.$$

flaw: we do not know the rate of convergence

uses: many, e.g. verification of the probabilistic model, MC methods



Application of the SLLN:

1. Convergence of the sample mean

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow[n \rightarrow \infty]{a.s.} \mathbb{E}X_1.$$

2. Convergence of the sample variance

$$S^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \bar{X})^2 \xrightarrow[n \rightarrow \infty]{a.s.} \text{Var}X_1.$$



Applications of the SLLN – cont.

3. Convergence of sample distributions: for

$$\mu_n(A) = \frac{1_A(X_1) + 1_A(X_2) + \dots + 1_A(X_n)}{n}$$

we have $\mu_n(A) \xrightarrow[n \rightarrow \infty]{a.s.} \mathbb{E}1_A(X_1) = \mathbb{P}(X_1 \in A)$

4. Convergence of sample CDFs: for

$$F_n(t) = \frac{1_{\{X_1 \leq t\}} + 1_{\{X_2 \leq t\}} + \dots + 1_{\{X_n \leq t\}}}{n}$$

we have $F_n(t) \xrightarrow[n \rightarrow \infty]{a.s.} F(t)$



Applications of SLLN – cont. (2)

5. Glivenko–Cantelli Theorem

Let X_1, X_2, \dots be independent random variables from a distribution with a CDF F . Then,

$$\sup_{t \in \mathbb{R}} |F_n(t) - F(t)| \xrightarrow[n \rightarrow \infty]{a.s.} 0.$$



Central Limit Theorem

1. Classical version:

Let X_1, X_2, \dots be identically distributed independent random variables, such that $\mathbb{E}X_1^2 < \infty$. If by $m = \mathbb{E}X_1$ we denote the mean, and by $\sigma^2 = \text{Var}X_1$ the variance of this distribution, then for any $t \in \mathbb{R}$, we have that

$$\mathbb{P}\left(\frac{X_1 + X_2 + \dots + X_n - nm}{\sigma\sqrt{n}} \leq t\right) \xrightarrow{n \rightarrow \infty} \Phi(t),$$

where $\Phi(t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) dx$

is the CDF of the standard normal distribution.

also:

$$\mathbb{P}\left(s \leq \frac{X_1 + X_2 + \dots + X_n - nm}{\sigma\sqrt{n}}\right) \xrightarrow{n \rightarrow \infty} 1 - \Phi(s)$$



WARSAW U
Faculty of

$$\mathbb{P}\left(s \leq \frac{X_1 + X_2 + \dots + X_n - nm}{\sigma\sqrt{n}} \leq t\right) \xrightarrow{n \rightarrow \infty} \Phi(t) - \Phi(s)$$

De Moivre-Laplace Theorem

2. Theorem:

Let X_1, X_2, \dots be a sequence of independent identically distributed random variables, such that $\mathbb{P}(X_n = 1) = p = 1 - \mathbb{P}(X_n = 0)$.

Then, we have that for any $s < t$,

$$\mathbb{P} \left(s \leq \frac{X_1 + X_2 + \dots + X_n - np}{\sqrt{np(1-p)}} \leq t \right) \xrightarrow{n \rightarrow \infty} \Phi(t) - \Phi(s).$$

each inequality (both in the CLT and in dML) may be changed to strict without consequences



Central Limit Theorem

3. Examples

- boys and girls
- how many students should be accepted?
- aggregate errors
- confidence intervals





WARSAW UNIVERSITY
Faculty of Economic Sciences