Probability Calculus

Anna Janicka

lecture XII, 18.12.2018

CHEBYSHEV INEQUALITIES

CONVERGENCE

LAWS OF LARGE NUMBERS

Plan for Today

1. Conditional expectation as a predictor
2. Chebyshev Inequalities
3. Types of convergence of Random Variables
   - convergence almost surely
   - convergence in probability
4. Laws of Large Numbers
   - Weak LLN
   - Strong LLN

Conditional Expectation as an approximation

1. Theorem:

Let $X, Y : \Omega \to \mathbb{R}$ *be random variables such that* $\mathbb{E}Y^2 < \infty$. *Then, the function* $\varphi^* : \mathbb{R} \to \mathbb{R}$, *such that* $\varphi^*(x) = \mathbb{E}(Y|X = x)$, *satisfies:*

$$\mathbb{E}(Y - \varphi^*(X))^2$$

$$= \min\{\mathbb{E}(Y - \varphi(X))^2 : \varphi \textit{ is a Borel function} : \mathbb{R} \to \mathbb{R}\}.$$

Chebyshev Inequality

1. Sometimes we are only interested in inequalities of the type

$$\mathbb{P}(X \geq x) \leq \alpha.$$

2. Chebyshev inequality

Let X be a nonnegative integrable random variable, and let $\varepsilon > 0$. We have:

$$\mathbb{P}(X \geqslant \varepsilon) \leqslant \frac{\mathbb{E}X}{\varepsilon}$$

Chebyshev Inequality – derivates

3. Chebyshev inequality for $|X|^p$, $(X - \mathbb{E}X)^2$ and $e^{\lambda X}$

Let X be a random variable.

- **Markov Inequality:** *For any $p > 0$ such that $\mathbb{E}|X|^p$ exists, and any $\varepsilon > 0$,*
  $\mathbb{P}(|X| \geqslant \varepsilon) \leqslant \frac{\mathbb{E}|X|^p}{\varepsilon^p}$.
- **Chebyshev-Bienaymé Inequality:** *For any $\varepsilon > 0$, if the random variable X^2 is integrable,*
  $\mathbb{P}(|X - \mathbb{E}X| \geqslant \varepsilon) \leqslant \frac{\mathrm{Var}(X)}{\varepsilon^2}$.
- **Exponential Chebyshev Inequality:**
  Let us assume that $\mathbb{E}e^{pX} < \infty$ for a given value $p > 0$. Then, for any $\lambda \in [0, p]$ and for any $\varepsilon > 0$,
  $\mathbb{P}(X \geqslant \varepsilon) \leqslant \frac{\mathbb{E}e^{\lambda X}}{e^{\lambda \varepsilon}}$.

Bernstein Inequality

4. Bernstein inequality:

Let S_n be a random variable from a binomial distribution with parameters n and p. Then, for any $\varepsilon > 0$, *we have* $\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant 2e^{-2\varepsilon^2 n}$.

also:

$$\mathbb{P}\left(\frac{S_n}{n} \geqslant p + \varepsilon\right) \leqslant e^{-2\varepsilon^2 n}$$

$$\mathbb{P}\left(\frac{S_n}{n} \leqslant p - \varepsilon\right) \leqslant e^{-2\varepsilon^2 n}$$

5. Examples

Comparison of inequalities

ε	n	Chebyshev-Bienaymé	Bernstein
0,1	100	0,25	0,2707
0,1	1000	0,025	4,1223E-09
0,05	100	1	1,2131
0,05	1000	0,1	0,0135
0,05	10000	0,01	3,8575E-22
0,01	100	25	1,9604
0,01	1000	2,5	1,6375
0,01	10000	0,25	0,2707
0,01	100000	0,025	4,1223E-09

Types of convergence:

1. Almost sure convergence

A sequence $(X_n)_{n\geqslant 1}$ *of random variables over* Ω *converges* **almost surely** *to* X, *if* $\mathbb{P}(\lim_{n\to\infty} X_n = X) = 1$.

Equivalently, we may say that there exists a subset $\Omega' \subset \Omega$ *such that* $\mathbb{P}(\Omega') = 1$, *such that for any* $\omega \in \Omega'$, *we have* $\lim_{n\to\infty} X_n(\omega) = X(\omega)$.

Denoted by $X_n \xrightarrow{a.s.} X$.

An alternative formulation:

$$\lim_{n\to\infty} \mathbb{P}(\sup_{k\geqslant n} |X_k - X| > \varepsilon) = 0.$$

Types of convergence – cont.

2. Convergence in probability

A sequence $(X_n)_{n\geqslant 1}$ *of random variables over* Ω *converges* **in probability** *to* X, *if for any* $\varepsilon > 0$, *we have that* $\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$.

Equivalently, for any $\varepsilon > 0$,
$\lim_{n\to\infty} \mathbb{P}(|X_n - X| \leqslant \varepsilon) = 1$.

Denoted by $X_n \xrightarrow{\mathbb{P}} X$ *or* $\operatorname{plim}_{n\to\infty} X_n = X$.

3. Almost sure convergence $\Rightarrow$ convergence in probability

Properties of limits of RV

4. Theorem

Let $(X_n)_{n\geqslant 1}$ *and* $(Y_n)_{n\geqslant 1}$ *be sequences of random variables. If* $(X_n)_{n\geqslant 1}$ *converges to* X *and* $(Y_n)_{n\geqslant 1}$ *converges to* Y *almost surely (/in probability)), then* $X_n \pm Y_n \to X \pm Y$ *and* $X_n \cdot Y_n \to XY$ *almost surely (/in probability).*

Weak Laws of Large Numbers

1. Weak Law of Large Numbers for the Bernoulli Scheme

Let $X_1, X_2, \ldots$ *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* $\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) = 0$.

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

Weak Laws of Large Numbers – cont.

2. Weak Law of Large Numbers for uncorrelated random variables

Let $X_1, X_2, \ldots$ be uncorrelated random variables with a common upper bound to their variances. Then, the sequence $(X_n)_{n\geqslant 1}$ 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* $\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n - \mathbb{E}S_n}{n}\right| > \varepsilon\right) = 0$.

Weak Laws of Large Numbers – cont. (2)

Examples

- independent events
- variances without bounds → NO
- correlated RV → NO
- embarrassing question

1. Strong Law of Large Numbers for the Bernoulli Scheme

Let $X_1, X_2, \ldots$ *be a sequence of independent random variables, such that*
$\mathbb{P}(X_n = 1) = p = 1 - \mathbb{P}(X_n = 0), \quad n = 1, 2, \ldots.$
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*
$\lim_{n\to\infty} \frac{S_n(\omega)}{n} = p.$

Strong Laws of Large Numbers – cont.

2. Kolmogorov’s Strong Law of Large Numbers

Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed integrable random variables. Then, $\frac{S_n}{n} \xrightarrow[n\to\infty]{a.s.} \mathbb{E}X_1$.

flaw: we do not know the rate of convergence

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

Applicationf of the SLLN:

1. Convergence of the sample mean

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow[n\to\infty]{a.s.} \mathbb{E}X_1.$$

2. Convergence of the sample variance

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

Applications of the SLLN – cont.

3. Convergence of sample distributions: for

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

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

4. Convergence of sample CDFs: for

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

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

Applications of SLLN – cont. (2)

5. Glivenko–Cantelli Theorem

Let $X_1, X_2, \ldots$ be independent random variables from a distribution with a CDF F. Then, $\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| \xrightarrow[n\to\infty]{a.s.} 0.$