STATISTICS

Module 2 *of*
MATHEMATICS AND STATISTICS

Marina Santacroce

marina.santacroce@unibocconi.it
Department of Decision Sciences, Bocconi University

L12: Limit theorems: (weak) law of large numbers (LLN) and central limit theorem (CLT)

References: S.Ross *Introduction to* Probability and Statistics for Engineers and Scientists: Sections 4.9 and 6.3

Outline

Useful inequalities

Prop. (Markov inequality) X is a non negative random variable. Then, for any $a > 0$

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Suppose that X is continuous. Since $X \geq 0$ with probability 1,

$$\mathbb{E}(X) = \int_0^\infty xf(x)dx = \int_0^a xf(x)dx + \int_a^\infty xf(x)dx \geq a \int_a^\infty f(x)dx = a\,\mathbb{P}(X \geq a).$$

Rem. The inequality gives an *upper bound* of the probability of $X \geq a$ in terms of $\mathbb{E}(X)$.

Cor: (Chebyshev inequality) Given a random variable X with finite mean μ and variance σ^2. Then for any $k > 0$

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

Consider the event $\{|X - \mu| \geq k\} = \{(X - \mu)^2 \geq k^2\}$ and use Markov inequality for the random variable $(X - \mu)^2$. Note that $\mathbb{E}((X - \mu)^2) = \sigma^2$.

Rem. If X has $Var(X) = 0 \Rightarrow \mathbb{P}(X = \mathbb{E}(X)) = 1$.

Law of Large Numbers

Law of large numbers are theorems which state the conditions under which the sample mean converges (in some sense) to its expected value.

Theorem: "(Weak) Law of large numbers"[1][2]

Let $X_1, X_2, \dots$ be a sequence of i.i.d. random variables with finite mean $\mathbb{E}(X_i) = \mu$.Then for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left| \frac{X_1 + \cdots + X_n}{n} - \mu \right| \geq \varepsilon \right) = 0.$$

Let $Var(X_i) = \sigma^2 < \infty$. Recall

$$\mathbb{E}\left(\frac{X_1 + \cdots + X_n}{n} \right) = \mu \text{ and } Var(\frac{X_1 + \cdots + X_n}{n}) = \frac{\sigma^2}{n}.$$

From Chebyshev inequality follows that

$$\mathbb{P}\left(\left| \frac{X_1 + \cdots + X_n}{n} - \mu \right| \geq \varepsilon \right) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Thus, the result follows by taking the limit as $n \to \infty$.

---

[1] weak refers to the kind of convergence (in probability)

[2] J. Bernoulli proved it in 1713 for Bernoulli random variables, this proof is due to Khintchine

Central Limit Theorem

In the central limit theorem the distribution of sums of random variables suitable standardized are shown to converge to the standard normal.

Theorem: "Central Limit Theorem"[3]

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables, with mean $\mathbb{E}(X_i) = \mu$ and variance $Var(X_i) = \sigma^2$ (finite). Then, the distribution of

$$\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n\sigma^2}}$$

tends to the one of the standard normal, i.e.

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n\sigma^2}} \leq z\right) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du, \qquad z \in \mathbb{R}.$$

Rem. The central limit theorem has been shown under weaker hypotheses.

This results justifies the empirical evidence of real data follow (are close to) the normal distribution.

[3] The central limit theorem with Bernoulli random variables is called De Moivre Laplace theorem