Prob. & Stats. - Week 9 Outline

November 12, 2017

1 LLN (continued)

1.1 Compact form of LLN

Statement:

$$\overline{X}_n \xrightarrow[n\to\infty]{} E(X) \tag{1}$$

where "$\xrightarrow[n\to\infty]{}$" has the double meaning

- Weak form (*convergence in probability*):

$$P(|\overline{X}_n - \mu| > \epsilon) \xrightarrow[n\to\infty]{} 0,\ \forall \epsilon > 0 \tag{2}$$

  if $X_1, \ldots, X_n$ uncorrelated and with common finite mean μ and variance σ^2.

- Strong form (*covergence almost sure*):

$$P\big(\{\lim_{n\to\infty} \overline{X}_n = \mu\}\big) = 1 \tag{3}$$

  if $X_1, \ldots, X_n$ are IID with finite common mean μ.

1.2 LLN for frequencies and its applications

- **Statement:** The *empirical frequency* of an event C

$$\overline{Y}_n(C) \;=\; \frac{\sum_{i=1}^n Y_i}{n} \;=\; \frac{\text{\# of times in C}}{n}$$

  satisfy

$$\overline{Y}_n(C) \xrightarrow[n\to\infty]{} P(X \in C) \tag{4}$$

  in the senses (2)–(3).

- Important observation: The *only* hypothesis needed is that $X_1, \ldots, X_n$ be IID (in fact, for the weak version they only need to be uncorrelated). This version of the law is valid *for all* random variables X, with or without finite means or variances.

- **Application to distribution functions:** If $X_i \sim X$ fir all i,

$$\overline{Y}_n(C) \xrightarrow[n\to\infty]{} F_X$$

in the senses (2)–(3).

- **Application to probability masses:** If X is a discrete random variable,

$$\overline{Y}_n(a) \;=\; \frac{\#\text{ of times } X_i = a}{n} \xrightarrow[n\to\infty]{} p_X(a)$$

in the senses (2)–(3).

- **Application to probability densities:** If X is a continuous random variable,

$$\begin{aligned} \overline{Y}_n(a,h) \quad &= \quad \frac{\#\text{ of times } X_i = [a-h, a+h]}{n} \\ &\xrightarrow[n\to\infty]{} \quad \int_{a-h}^{a+h} f_X(x)\, dx \qquad (5) \\ &\approx \quad f_x(a)\, 2h\;. \qquad (6) \end{aligned}$$

The last approximation holds for h sufficiently small. As a consequence,

$$\frac{\overline{Y}_n(a,h)}{2h} \;\approx\; f_X(a)\;.$$

This approximation is subjected to two types of errors

  - *Stochastic error:* Due to the use of inite n in (5). It can be reduced increasing n.
  - *Discretization error:* Due to the use of non-zero h in (6). It can be reduced decreasing h.

2 The Central Limit Theorem (CLT)

2.1 The theorem

- **Theorem.** Let $X_1, \ldots, X_n$ be IID with common finite mean μ and variance σ^2. Define

$$Z_n \;=\; \sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right)\;.$$

Then

$$P(Z_n \in C) \xrightarrow[n\to\infty]{} P(Z \in C) \quad , \quad \text{with } Z \sim N(0,1) \qquad (7)$$

for each C.

- Observations

– The convergence (9) is termed *convergence in law*, and (9) is often written as
$$\sqrt{n}\left(\frac{\overline{X}_n-\mu}{\sigma}\right) \xrightarrow[n\to\infty]{} N(0,1) \quad \text{in law} \tag{8}$$

– Equivalently, the CLT states that
$$\frac{X_1+\cdots+X_n-n\,\mu}{\sqrt{n}\,\sigma} \xrightarrow[n\to\infty]{} N(0,1) \quad \text{in law} \tag{9}$$

2.2 Consequences

- Convergence (9) can be interpreted, for practical purposes, as
$$Z_n \;\approx\; Z \quad \text{(in law)}\;. \tag{10}$$

- Solving in (10) for $\overline{X}_n$ in terms of Z_n and using Exercise 1 of Problem set 10, we can write the CLT in the form
$$\overline{X}_n \;\approx\; N(\mu\,,\sigma^2/n) \tag{11}$$
well "≈" is understood as "in law

- Likewise, from (9),
$$X_1+\cdots+X_n \;\approx\; N(n\mu\,,n\sigma^2) \tag{12}$$

2.3 Normal approximation of the binomial law

If $X \sim \text{Bin}(n,p)$ then
$$\frac{X-np}{\sqrt{np(1-p)}} \xrightarrow[n\to\infty]{} N(0,1) \quad \text{in law}\;. \tag{13}$$

Equivalent expressions (if understood in law):
$$\frac{X-np}{\sqrt{np(1-p)}} \;\approx\; N(0,1)\;. \tag{14}$$

and
$$X \;\approx\; N\big(np\,,np(1-p)\big)\;. \tag{15}$$