MODULE 7

LIMITING DISTRIBUTIONS

LECTURE 40

Topics

7.1 CONVERGENCE IN DISTRIBUTION AND PROBABILITY

7.1.1 Poisson Approximation to Binomial distribution

7.2 THE WEAK LAW OF LARGE NUMBERS (WLLN) AND THE CENTRAL LIMIT THEOREM (CLT)

7.2.1 Random Walk

7.2.2 Justification of Relative Frequency Method of Assigning Probabilities

Proposition 1.1

Let $\{c_n\}_{n\geq 1}$ be a sequence of positive real numbers such that $\lim_{n\to\infty} c_n = c \in \mathbb{R}$. Then

$$\lim_{n\to\infty}\left(1+\frac{c_n}{n}\right)^n = e^c.$$

Proof. We know that

$$x-\frac{x^2}{2} \leq \ln(1+x) \leq x, \forall x > 0$$

$$\Rightarrow c_n - \frac{{c_n}^2}{2n} \leq n\ln\left(1+\frac{c_n}{n}\right) \leq c_n, \qquad n = 1, 2, \ldots$$

$$\Rightarrow \lim_{n\to\infty}\left[n\ln\left(1+\frac{c_n}{n}\right)\right] = c \qquad \text{(on taking limits on both sides)}$$

$$\Rightarrow \lim_{n\to\infty}\left[\ln\left(1+\frac{c_n}{n}\right)^n\right] = c$$

$$\Rightarrow \lim_{n\to\infty}\left(1+\frac{c_n}{n}\right)^n = e^c. \blacksquare$$

7.1.1 Poisson Approximation to Binomial distribution

Example 1.13

Let $X_n \sim \text{Bin}(n, \theta_n)$, where $\theta_n \in (0,1), n = 1, 2, \ldots,$ and let $\lim_{n\to\infty}(n\theta_n) = \theta > 0$. Show that $X_n \xrightarrow{d} X$, as $n \to \infty$, where $X \sim \text{P}(\theta)$, the Poisson distribution with mean θ.

Solution. Note that the m.g.f. of X is

$$M(t) = e^{\theta(e^t - 1)}, t \in \mathbb{R},$$

and the m.g.f. of X_n is

$$\begin{aligned} M_n(t) &= (1 - \theta_n + \theta_n e^t)^n \\ &= \left(1 + \frac{c_n(t)}{n}\right)^n, \qquad t \in \mathbb{R}, \end{aligned}$$

where $c_n(t) = n\theta_n(e^t - 1), t \in \mathbb{R},\ n = 1, 2, \ldots$. Clearly $\lim_{n\to\infty} c_n(t) = \theta(e^t - 1), \forall t \in \mathbb{R}$. Now using Proposition 1.1 we get

$$\lim_{n\to\infty} M_n(t) = e^{\theta(e^t - 1)} = M(t),\ \forall t \in \mathbb{R}.$$

Using Theorem 1.5(i) we conclude that $X_n \xrightarrow{d} X \sim P(\theta)$, as $n \to \infty$. ■

7.2 THE WEAK LAW OF LARGE NUMBERS (WLLN) AND THE CENTRAL LIMIT THEOREM (CLT)

Let $\{X_n\}_{n\geq 1}$ be a sequence of i.i.d. random variables and let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i, n = 1,2, \ldots,$ be the corresponding sequence sample means. In this section we will study the convergence behavior of the sequence $\{\bar{X}_n\}_{n\geq 1}$ of sample means.

Theorem 2.1

Let $\{X_n\}_{n\geq 1}$ be a sequence of i.i.d. random variables and let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i, n = 1,2, \ldots$.

(i) (**WLLN**) Suppose that $E(X_1) = \mu$ is finite. Then $\bar{X}_n \xrightarrow{p} \mu$, as $n \to \infty$.
(ii) (**CLT**) suppose that $0 < \text{Var}(X_1) = \sigma^2 < \infty$. Then

$$Z_n \overset{\text{def}}{=} \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Z \sim N(0,1), \text{as } n \to \infty.$$

Proof.

(i) As the proof for the case $\mathrm{Var}(X_1) = \infty$ is quite involved, for simplicity, we assume that $\mathrm{Var}(X_1) = \sigma^2 < \infty$. Then

$$E(\bar{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = E(X_1) = \mu$$

and

$$\mathrm{Var}(\bar{X}_n) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{\sigma^2}{n} \xrightarrow{n\to\infty} 0.$$

Using Theorem 1.4 it follows that $\bar{X}_n \xrightarrow{p} \mu$, as $n \to \infty$.

(ii) For simplicity we will assume that the common m.g.f. $M(\cdot)$ of $X_1, X_2, \ldots$ is finite in an interval $(-a, a)$ for some $a > 0$. Then, by Theorem 3.4, Module 3, $\mu_r' = E(X_1^r)$ is finite for each $r \in \{1, 2, \cdots\}$ and $\mu_r' = E(X_1^r) = M^{(r)}(0) = \left[\frac{d^r}{dt^r} M(t)\right]_{t=0}, r = 1, 2, \ldots$. Let $Y_i = \frac{X_i - \mu}{\sigma}, i = 1, \ldots, n$. Then $Y_1, Y_2, \ldots$ are i.i.d. random variables with mean 0 and variance 1. Let $M_Y(\cdot)$ denote the common m.g.f. of $Y_1, Y_2, \ldots$, so that

$$M_Y(t) = e^{-\frac{\mu t}{\sigma}} M\left(\frac{t}{\sigma}\right), \qquad -a\sigma < t < a\sigma,$$

$$M_Y^{(1)}(0) = -\frac{\mu}{\sigma} + \frac{M^{(1)}(0)}{\sigma} = 0 = E(Y_1)$$

and
$$M_Y^{(2)}(0) = \left(\frac{\mu}{\sigma}\right)^2 M(0) - \frac{2\mu}{\sigma^2} M^{(1)}(0) + \frac{1}{\sigma^2} M^{(2)}(0) = 1 = E(Y_1^2).$$

Let $\psi_2 : (-a\sigma, a\sigma) \to \mathbb{R}$ be such that

$$M_Y(t) = M_Y(0) + tM_Y^{(1)}(0) + \frac{t^2}{2}\left(M_Y^{(2)}(0) + \psi_2(t)\right), \qquad t \in (-a\sigma, a\sigma) \qquad (2.1)$$

i.e.,
$$\psi_2(t) = \frac{M_Y(t) - M_Y(0) - tM_Y^{(1)}(0)}{t^2/2} - M_Y^{(2)}(0), \quad t \in (-a\sigma, a\sigma), \ t \neq 0.$$

Using L' Hospital rule $(0/0$ form$)$ we get

$$\lim_{t\to 0} \psi_2(t) = \lim_{t\to 0} \frac{M_Y^{(1)}(t) - M_Y^{(1)}(0)}{t} - M_Y^{(2)}(0)$$

$$= M_Y^{(2)}(0) - M_Y^{(2)}(0)$$

$$= 0. \qquad (2.2)$$

The m.g.f. of $Z_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i$ is

$$M_n(t) = E\left(e^{\frac{t}{\sqrt{n}} \sum_{i=1}^{n} Y_i}\right)$$

$$= E\left(\prod_{i=1}^{n} e^{\frac{tY_i}{\sqrt{n}}}\right)$$

$$= \prod_{i=1}^{n} E\left(e^{\frac{tY_i}{\sqrt{n}}}\right) \qquad (Y_i\text{s are independent})$$

$$= \left[M_Y\left(\frac{t}{\sqrt{n}}\right)\right]^n \qquad (Y_i\text{s are i.i.d.})$$

$$= \left[M_Y(0) + \frac{t}{\sqrt{n}} M_Y^{(1)}(0) + \frac{t^2}{2n}\left(M_Y^{(2)}(0) + \psi_2\left(\frac{t}{\sqrt{n}}\right)\right)\right]^n \qquad (\text{using } (2.1))$$

$$= \left[1 + \frac{t^2}{2n}\left(1 + \psi_2\left(\frac{t}{\sqrt{n}}\right)\right)\right]^n, t \in (-\sqrt{n}a\sigma, \sqrt{n}a\sigma), n = 1, 2, \ldots.$$

Now using (2.2) and Proposition 1.1 we get

$$\lim_{n\to\infty} M_n(t) = e^{\frac{t^2}{2}} = K(t), \text{ say}, \qquad t \in \mathbb{R}.$$

Note that $K(t), t \in \mathbb{R}$, is the m.g.f. of $Z \sim N(0,1)$. Using Theorem 1.5 (i) we conclude that $Z_n \xrightarrow{d} Z \sim N(0,1)$, as $n \to \infty$. ■

Remark 2.1

(i) The WLLN implies that the sample mean, based on a random sample from any parent distribution, can be made arbitrarily close to the population mean in probability by choosing sufficiently large sample size.

(ii) The CLT states that, irrespective of the nature of the parent distribution, the probability distribution of a normalized version of the sample mean, based on a random sample of large size, is approximately normal. For this reason the normal distribution is quite important in the field of Statistics. ■

7.2.1 Random Walk

Example 2.1

Consider a drunkard, who having missed his bus from the bus stand, starts walking towards his residence. Every second he either moves half a meter forward or half a meter backward from his current position, each with probability $\frac{1}{2}$. Assuming that steps are taken independently, find the (approximate) probability that after fifteen minutes the drunkard will be within 30 meters form the bus stand.

Solution. Note that in 15 minutes ($=$ 900 seconds) the drunkard will take 900 steps. Let Y_i be the size (in meters) of the i-th step, $i = 1, 2, \dots, 900$. Then $Y_1, Y_2, \dots$ are i.i.d. random variables with

$$P\left(\left\{Y_1 = -\frac{1}{2}\right\}\right) = P\left(\left\{Y_1 = \frac{1}{2}\right\}\right) = \frac{1}{2},$$

and $Y = \sum_{i=1}^{900} Y_i$ is the position of the drunkard after 15 minutes. The desired probability is

$$P(\{|Y| \leq 30\}) = P\left(\left\{-\frac{1}{30} \leq \bar{Y}_{900} \leq \frac{1}{30}\right\}\right),$$

where $\bar{Y}_{900} = \frac{1}{900}\sum_{i=1}^{900} Y_i = \frac{Y}{900}$. Note that $E(Y_1) = 0$ and $\mathrm{Var}(Y_1) = E(Y_1^2) = \frac{1}{4} = \sigma^2$, say. By the CLT

$$Z_{900} = \frac{\sqrt{900}(\bar{Y}_{900} - 0)}{1/2} \overset{\text{approx}}{\sim} N(0,1),$$

i.e.,
$$Z_{900} = 60\, \bar{Y}_{900} \overset{\text{approx}}{\sim} N(0,1).$$

The desired probability is

$$\begin{aligned} P(\{|Y| \leq 30\}) &= P(\{-2 \leq Z_{900} \leq 2\}) \\ &\overset{\text{approx .}}{=} \Phi(2) - \Phi(-2) \\ &= 2\Phi(2) - 1 \\ &= 2 \times .9772 - 1 \\ &= .9544. \blacksquare \end{aligned}$$

7.2.2 Justification of Relative Frequency Method of Assigning Probabilities

Example 2.2

Suppose that we have independent repetitions of a random experiment under identical conditions. Further suppose that we are interested in assigning probability, say $P(E)$, to an event E. To do this we repeat the random experiment a large (say N) number of times. Define

$$Y_i = \begin{cases} 1, & \text{if } i-\text{th trial results in occurrence of } E \\ 0, & \text{otherwise} \end{cases}, \quad i = 1, \dots, N.$$

Then $Y_1, Y_2, \dots$ are i.i.d. random variables with common mean $\mu = E(Y_1) = P(E)$. Also

$$f_N(E) = \text{number of times event } E \text{ occurs in first } N \text{ trials}$$

$$= \sum_{i=1}^{N} Y_i$$

and the relative frequency of event E in first N trials is

$$r_N(E) = \frac{f_N(E)}{N} = \frac{1}{N}\sum_{i=1}^{N} Y_i = \bar{Y}_N, \text{ say}\cdot$$

The WLLN implies that

$$r_N(E) = \bar{Y}_N \xrightarrow{p} \mu = P(E), \text{ as } N \to \infty.$$

Thus the WLLN justifies the relative frequency approach to assign probabilities. ■