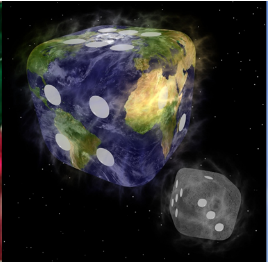
CHAPTER 8: LIMIT THEOREMS

Lecturer: Nguyen Minh Quan, PhD
Department of Mathematics
HCMC International University



CONTENTS

1. Chebyshev's Inequality and the Weak Law of Large Numbers
2. The Central Limit Theorem
3. The Strong Law of Large Numbers

Markov's inequality

We start this section by proving a result known as Markov's inequality.

Markov's inequality

If X is a random variable that takes only nonnegative values, then, for any value $a > 0$,

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

Proof

For $a > 0$, let

$$I = \begin{cases} 1, & \text{if } X \geqslant a \\ 0, & \text{otherwise} \end{cases}$$

Thus, $I \leqslant \frac{X}{a}$. Therefore,

$$E(I) \leqslant \frac{E(X)}{a}$$

But $E(I) = P(X \geqslant a)$. This proves the result.

Chebyshev's inequality

Chebyshev's inequality

If X is a random variable with finite mean μ and variance σ^2, then, for any value $k > 0$,

$$P\left(|X-\mu| \geqslant k\right) \leqslant \frac{\sigma^2}{k^2}$$

Proof

Since $(X-\mu)^2$ is a nonnegative random variable, we can apply Markov's inequality (with $a = k^2$) to obtain

$$P\left((X-\mu)^2 \geqslant k^2\right) \leqslant \frac{E\left[(X-\mu)^2\right]}{k^2}$$

$$P\left(|X-\mu| \geqslant k\right) \leqslant \frac{E\left[(X-\mu)^2\right]}{k^2} = \frac{\sigma^2}{k^2}$$

Markov's and Chebyshev's inequalities

The importance of Markov's and Chebyshev's inequalities is that they enable us to derive bounds on probabilities when only the mean, or both the mean and the variance, of the probability distribution are known.

Markov's and Chebyshev's inequalities

Example

Suppose that it is known that the number of items produced in a factory during a week is a random variable with mean 50.
(a) What can be said about the probability that this week's production will exceed 75?
(b) If the variance of a week's production is known to equal 25, then what can be said about the probability that this week's production will be between 40 and 60?

Solution (a) By Markov's inequality, $P\left(X > 75\right) \leqslant \frac{E(X)}{75} = \frac{50}{75} = \frac{2}{3}$.
(b) By Chebyshev's inequality,

$$P\left(|X - 50| \geqslant 10\right) \leqslant \frac{\sigma^2}{10^2} = \frac{1}{4}$$

$$P\left(|X - 50| < 10\right) \leqslant 1 - \frac{1}{4} = \frac{3}{4}$$

Markov's and Chebyshev's inequalities

If X is a normal random variable with mean μ and variance σ^2, Chebyshev's inequality states that

$$P\left(|X-\mu| > 2\sigma\right) \leqslant \frac{1}{4}$$

whereas the actual probability is given by

$$P\left\{|X-\mu| > 2\sigma\right\} = P\left(\left|\frac{X-\mu}{\sigma}\right| > 2\right) = 2\left[1-\Phi\left(2\right)\right] \approx 0.0456$$

The weak law of large numbers

The weak law of large numbers

Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables, each having finite mean $E[X_i] = \mu$. Then, for any $\epsilon > 0$,

$$P\left\{\left|\frac{X_1 + X_2 + ... + X_n}{n} - \mu\right| \geqslant \varepsilon\right\} \to 0 \text{ as } n \to \infty$$

Proof

Use the Chebyshev's inequality:

$$P\left\{\left|\frac{X_1 + X_2 + ... + X_n}{n} - \mu\right| \geqslant \varepsilon\right\} \leqslant \frac{\sigma^2}{n\varepsilon^2}$$

The Central Limit Theorem

The Central Limit Theorem

Let $X_1, X_2, ..., X_n$ be a sequence of independent and identically distributed random variables each having mean μ and variance σ^2. Then for n large, the distribution of $X_1 + \cdots + X_n$ is approximately normal with mean $n\mu$ and variance $n\sigma^2$.

It follows from the central limit theorem that $\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}$ is approximately a standard normal random variable; thus, for n large,

$$P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} < x\right) \simeq P\left(Z < x\right)$$

where Z is a standard normal random variable.

The Central Limit Theorem

Example

The number of students who enroll in a psychology course is a Poisson random variable with mean 100. The professor in charge of the course has decided that if the number enrolling is 120 or more, he will teach the course in two separate sections, whereas if fewer than 120 students enroll, he will teach all of the students together in a single section. What is the probability that the professor will have to teach two sections?

Solution If X denotes the number of students that enroll in the course, we have

$$\begin{aligned} P(X \geqslant 120) &= P(X \geqslant 119.5) \text{ (the continuity correction)} \\ &= P\left\{\frac{X-100}{\sqrt{100}} \geqslant \frac{119.5-100}{\sqrt{100}}\right\} \\ &\approx 1-\Phi(1.95) \approx 0.0256 \end{aligned}$$

The Central Limit Theorem

Example

An insurance company has $25,000$ automobile policy holders. If the yearly claim of a policy holder is a random variable with mean 320 and standard deviation 540, approximate the probability that the total yearly claim exceeds 8.3 million.

Solution

Let X denote the total yearly claim. Number the policy holders, and let X_i denote the yearly claim of policy holder i. With $n = 25,000$, we have from the central limit theorem that $X = \sum_{i=1}^{n} X_i$ will have approximately a normal distribution with mean $320 \times 25,000 = 8 \times 10^6$ and standard deviation $540\sqrt{25000} = 8.5381 \times 10^4$.

The Central Limit Theorem

Example (Cont.)

$$P\left(X > 8.3 \times 10^6\right) = P\left(\frac{X - 8 \times 10^6}{8.5381 \times 10^4} > \frac{8.3 \times 10^6 - 8 \times 10^6}{8.5381 \times 10^4}\right)$$

$$\approx P\left(Z > 3.51\right) = 0.0002 \text{ (Z is standard normal)}$$

Applications of the CLT in binomial RV

Since a binomial random variable X having parameters (n, p) represents the number of successes in n independent trials when each trial is a success with probability p, we can express it as $X = X_1 + ... + X_n$.

$$X_i = \begin{cases} 1 \text{ if the ith trial is a success} \\ 0 \text{ otherwise} \end{cases}$$

We recall that $E(X_i) = p$ and $Var(X_i) = p(1-p)$.

Corollary

For n large $\frac{X-np}{\sqrt{np(1-p)}}$ will approximately be a standard normal random variable.

The Central Limit Theorem

Example

The ideal size of a first-year class at a particular college is 150 students. The college, knowing from past experience that, on the average, only 30 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 450 students. Compute the probability that more than 150 first-year students attend this college.

Solution

Let X denote the number of students that attend; then assuming that each accepted applicant will independently attend, it follows that X is a binomial random variable with parameters $n = 450$ and $p = 0.3$.

$$P(X > 150.5) = P\left(\frac{X - (450)(0.3)}{\sqrt{450(0.3)(0.7)}} \geqslant \frac{150.5 - (450)(0.3)}{\sqrt{450(0.3)(0.7)}}\right)$$

$$\approx P(Z > 1.59) = 0.06 \ (Z \text{ is standard normal})$$

Approximate Distribution of the Sample Mean

- Let $X_1, ..., X_n$ be a sample from a population having mean μ and variance σ^2. The central limit theorem can be used to approximate the distribution of the sample mean

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

- It follows from the central limit theorem that X will be approximately normal when the sample size n is large.
- Since the sample mean has expected value μ and standard deviation $\sigma/\sqrt{n}$, it then follows that

$$Z := \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

Approximate Distribution of the Sample Mean

Example

The weights of a population of workers have mean 167 and standard deviation 27.
(a) If a sample of 36 workers is chosen, approximate the probability that the sample mean of their weights lies between 163 and 170.
(b) Repeat part (a) when the sample is of size 144.

Solution

Let Z be a standard normal random variable.
(a) It follows from the central limit theorem that X is approximately normal with mean 167 and standard deviation $27/\sqrt{36} = 4.5$. Therefore,

$$\frac{\bar{X} - 167}{4.5} \sim N(0, 1)$$

$$P\left(163 < \bar{X} < 170\right) = P\left(\frac{163 - 167}{4.5} < \frac{\bar{X} - 167}{4.5} < \frac{170 - 167}{4.5}\right)$$

Approximate Distribution of the Sample Mean

Example

Solution (Cont.)

$$\approx P\left(-0.89 < Z < 0.67\right) \approx 0.5619$$

(b) $\frac{\bar{X}-163}{27/\sqrt{144}} \sim N\left(0,1\right) \Rightarrow \frac{\bar{X}-163}{2.25} \sim N\left(0,1\right)$

$$P\left(163 < \bar{X} < 170\right) = P\left(\frac{163-167}{2.25} < \frac{\bar{X}-167}{2.25} < \frac{170-167}{2.25}\right)$$

$$\approx P\left(-1.78 < Z < 1.33\right) \approx 0.8698$$

Thus increasing the sample size from 36 to 144 increases the probability from 0.5619 to 0.8698.

Approximate Distribution of the Sample Mean

Exercise

An astronomer wants to measure the distance from her observatory to a distant star. However, due to atmospheric disturbances, any measurement will not yield the exact distance d. As a result, the astronomer has decided to make a series of measurements and then use their average value as an estimate of the actual distance. If the astronomer believes that the values of the successive measurements are independent random variables with a mean of d light years and a standard deviation of 2 light years, how many measurements need she make to be at least 95 percent certain that her estimate is accurate to within ± 0.5 light years?

Hint

$$\frac{\bar{X} - d}{2/\sqrt{n}} \sim N(0, 1)$$

$$P\left(-0.5 < \bar{X} - d < 0.5\right) \geqslant 0.95 \Rightarrow \frac{\sqrt{n}}{4} \geqslant \phi(0.975)$$

The Strong Law of Large Numbers

The strong law of large numbers is probably the best-known result in probability theory. It states that the average of a sequence of independent random variables having a common distribution will, with probability 1, converge to the mean of that distribution.

The Strong Law of Large Numbers

Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables, each having a finite mean $\mu = E[X_i]$. Then, with probability 1,

$$\frac{X_1 + X_2 + ... + X_n}{n} \to \mu \text{ as } n \to \infty$$

The Strong Law of Large Numbers

As an application of the strong law of large numbers, suppose that a sequence of independent trials of some experiment is performed. Let E be a fixed event of the experiment, and denote by $P(E)$ the probability that E occurs on any particular trial. Letting

$$X_i = \begin{cases} 1, \text{if E occurs} \\ 0, \text{ otherwise} \end{cases}$$

We have, by the strong law of large numbers, that with probability 1,

$$\frac{X_1 + X_2 + ... + X_n}{n} \to E(X_i) = P(E)$$

That is, with probability 1, the limiting proportion of time that the event E occurs is just $P(E)$.

–THE END! THANK YOU!–