Probability Theory

Chapter 7: Properties of Expectation

Lecturer

Dr. Monjed H. Samuh

Applied Mathematics & Physics Department
Palestine Polytechnic University
(monjedsamuh@ppu.edu)

Term 191

Table of Contents

Chebyshev's Inequality and the Weak Law of Large Numbers

Theorem (Markov's Inequality)

Suppose that X is a random variable taking only non-negative values. Then, for any $a > 0$ we have

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

Proof:

Consider the random variable

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

Note that because $X \geq 0$,

$$I \leq \frac{X}{a}$$

Therefore,

$$E\left(\frac{X}{a}\right) \geq E(I) = P(X \geq a)$$

Example

Let X be $U(0,1)$. Then for any $a > 0$ Markov's inequality gives

$$P(X \geq a) \leq \frac{1}{2a},$$

Whereas

$$P(X \geq a) = \int_a^1 dx = 1 - a, \quad 0 < a < 1.$$

It is easy to show that

$$\frac{1}{2a} > 1 - a.$$

Chebyshev's Inequality and the Weak Law of Large Numbers

Theorem (Chebyshev's Inequality)

Let X be a random variable with finite mean μ and variance σ^2, then for any $k > 0$ we have

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

Proof:

Note that $(X-\mu)^2 \geq 0$, so we may use Markov's inequality. Apply with $a = k^2$ to find that

$$P\left((X-\mu)^2 \geq k^2\right) \leq \frac{E(X-\mu)^2}{k^2} = \frac{\sigma^2}{k^2}.$$

But $(X-\mu)^2 \geq k^2$ if and only if $|X-\mu| \geq k$, so

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

Chebyshev's Inequality and the Weak Law of Large Numbers

- If $k = a\sigma$ in Chebyshev's inequality, then

$$P\left(|X-\mu| \geq a\sigma\right) \leq \frac{1}{a^2}.$$

  This gives that probability of being a deviates from the expected value drops like $\frac{1}{a^2}$.

- Of course, we also have

$$P\left(|X-\mu| < a\sigma\right) \geq 1 - \frac{1}{a^2}.$$

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.

1. What can be said about the probability that this week's production will exceed 75? By Markov's inequality,

$$P(X > 75) \leq \frac{E(X)}{75} = \frac{50}{75}.$$

2. 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? By Chebyshev's inequality,

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

Hence,

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

so the probability that this week's production will be between 40 and 60 is at least 0.75.

Chebyshev's Inequality and the Weak Law of Large Numbers

Example

If X is uniformly distributed over the interval $(0, 10)$, then, since $E(X) = 5$ and $Var(X) = \frac{25}{3}$, it follows from Chebyshev's inequality that

$$P(|X - 5| > 4) \leq \frac{25}{3 \times 16} = 0.52.$$

whereas the exact result is

$$P(|X - 5| > 4) = 0.2.$$

Example

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

$$P(|X - \mu| > 2\sigma) \leq \frac{1}{4}.$$

whereas the actual probability is given by

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

Chebyshev's Inequality and the Weak Law of Large Numbers

- Although Chebyshev's inequality is correct, the upper bound that it provides is not particularly close to the actual probability.

- Chebyshev's inequality is often used as a theoretical tool in proving results.

Chebyshev's Inequality and the Weak Law of Large Numbers

Theorem

If $Var(X) = 0$, then $P(X = E(X)) = 1$.

Proof:

- *Let $E_n = \{|X - \mu| < \frac{1}{n}\}$, then $\{E_n,\ n \geq 1\}$ forms a decreasing sequence of events.*
- *Therefore,*

$$\lim_{n\to\infty} E_n = \bigcap_{n=1}^{\infty} E_n = \bigcap_{n=1}^{\infty} \left\{|X - \mu| < \frac{1}{n}\right\} = \{X = \mu\}.$$

- *By continuity property of the probability function*

$$\lim_{n\to\infty} P(E_n) = P\left(\lim_{n\to\infty} E_n\right).$$

- *So we have*

$$\lim_{n\to\infty} P\left(|X - \mu| < \frac{1}{n}\right) = P(X = \mu).$$

Theorem

Proof (Cont.):

- *Now, by Chebyshev's inequality*

$$P\left(|X-\mu| \geq \frac{1}{n}\right) \leq \frac{\sigma^2}{\left(\frac{1}{n}\right)^2} = 0,$$

implying that, $\forall n \geq 1$

$$P(E_n) = P\left(|X-\mu| < \frac{1}{n}\right) = 1.$$

- *Thus,*

$$P(X=\mu) = \lim_{n\to\infty} P(E_n) = 1.$$

Chebyshev's Inequality and the Weak Law of Large Numbers

Theorem (The Weak Law of Large Numbers)

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having the finite mean $E(X_i) = \mu$. Then for any $\epsilon > 0$ we have

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \epsilon\right) \longrightarrow 0 \text{ as } n \to \infty.$$

Proof:

1. *We will also assume that σ^2 is finite to make the proof easier.*

2. *We know that*

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

3. *Therefore, a direct application of Chebyshev's inequality shows that*

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2} \longrightarrow 0 \text{ as } n \to \infty.$$

The Central Limit Theorem (CLT)

Theorem (The Central Limit Theorem)

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having mean μ and variance σ^2. Then the distribution of

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

tends to the standard normal as $n \to \infty$.

$$P\left(\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n}\sigma} \leq a\right)$$

The Central Limit Theorem (CLT)

The following theorem is integral to the proof of the Central Limit theorem (not provided due to limited time).

Theorem

Let $Z_1, Z_2, \ldots$ be a sequence of random variables with distribution functions F_{Z_n} and moment generating functions M_{Z_n}, $n \geq 1$; and let Z be a random variable with a distribution function F_Z and a moment generating function M_Z.

$$M_{Z_n}(t) \longrightarrow M_Z(t), \quad \forall t,$$

then

$$F_{Z_n}(t) \longrightarrow F_Z(t), \quad \forall t,$$

at which $F_Z(t)$ is continuous.

If we let Z be a standard normal random variable, then, since $M_Z(t) = e^{\frac{1}{2}t^2}$, it follows that if

$$M_{Z_n}(t) \longrightarrow e^{\frac{1}{2}t^2} \text{ as } n \to \infty,$$

then

$$F_{Z_n}(t)(t) \longrightarrow \Phi(t) \text{ as } n \to \infty.$$

The Central Limit Theorem (CLT)

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?

- The exact solution

$$P(X \geq 120) = e^{-100} \sum_{x=120}^{\infty} \frac{100^x}{x!}$$

  does not readily yield a numerical answer.

- However, by recalling that a Poisson random variable with mean 100 is the sum of 100 independent Poisson random variables, each with mean 1.

- We can make use of the central limit theorem to obtain an approximate solution.

The Central Limit Theorem (CLT)

Example (Cont.)

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

$$\begin{aligned} P(X \geq 120) &= P(X > 119.5) \\ &= P\left(\frac{X-\mu}{\sigma} > \frac{119.5-100}{\sqrt{100}}\right) \\ &\approx 1 - \Phi(1.95) \\ &\approx 0.0256. \end{aligned}$$

The Central Limit Theorem (CLT)

Example

Let X_i, $i = 1, \ldots, 10$, be independent random variables, each uniformly distributed over $(0, 1)$. Calculate an approximation to $P\left(\sum_{i=1}^{10} X_i > 6\right)$.

Since $E(X_i) = \frac{1}{2}$ and $Var(X_i) = \frac{1}{12}$, we have, by the CLT,

$$
\begin{aligned}
P\left(\sum_{i=1}^{10} X_i > 6\right) &= P\left(Z > \frac{6-5}{\sqrt{10\left(\frac{1}{12}\right)}}\right) \\
&\approx 1 - \Phi(\sqrt{1.2}) \\
&\approx 0.1367.
\end{aligned}
$$

The Central Limit Theorem (CLT)

Example

An instructor has 50 exams that will be graded in sequence. The times required to grade the 50 exams are independent, with a common distribution that has mean 20 minutes and standard deviation 4 minutes. Approximate the probability that the instructor will grade at least 25 of the exams in the first 450 minutes of work.

1. If we let X_i be the time that it takes to grade exam i, then $X = \sum_{i=1}^{25} X_i$ is the time it takes to grade the first 25 exams.
2. Because the instructor will grade at least 25 exams in the first 450 minutes of work if the time it takes to grade the first 25 exams is less than or equal to 450, we see that the desired probability is $P(X \leq 450)$.
3. To approximate this probability, we use the CLT.
4. Now, $E(X) = \sum_{i=1}^{25} E(X_i) = 25 \times 20 = 500$,
5. Consequently, with Z being a standard normal random variable, we have

$$P(X \leq 450) = P\left(Z \leq \frac{450 - 500}{\sqrt{400}}\right)$$
$$\approx 1 - \Phi(2.5) = 0.006.$$

Problems and Exercises

PROBLEMS

PAGE 390:

$$1, 4, 9, 20, 24$$

THEORETICAL EXERCISES

PAGE 392:

$$1, 3$$

SELF-TEST PROBLEMS AND EXERCISES

PAGE 393:

$$1, 3, 4, 8$$