Math 461 Spring 2024

Renming Song

University of Illinois Urbana-Champaign

April 15, 2024

Outline

Outline

Solution to Test 2 is is on my homepage now. The distribution for Test 2 scores is also available on my homepage.

HW10 is due Friday, 04/19, before the end of class.

Solution to Test 2 is is on my homepage now. The distribution for Test 2 scores is also available on my homepage.

HW10 is due Friday, 04/19, before the end of class.

Outline

When we are given the distribution of a random variable X, we can find the probability of any event defined in terms of X. Suppose that we are only given the expectation and variance of X, then, in general, we can not find the probability of events defined in terms of X exactly.

But for some events, we can still get some meaningful estimates on their probabilities. Let's first look at the case of a non-negative random variable X. Suppose that we only know $E[X]$.

When we are given the distribution of a random variable X, we can find the probability of any event defined in terms of X. Suppose that we are only given the expectation and variance of X, then, in general, we can not find the probability of events defined in terms of X exactly.

But for some events, we can still get some meaningful estimates on their probabilities. Let's first look at the case of a non-negative random variable X. Suppose that we only know $E[X]$.

Markov inequality

Suppose that X is a non-negative random variable, then for any $a > 0$,

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

Proof

Define a random variable

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

Then $I \leq X/a$. Thus

$$P(X \geq a) = E[I] \leq E[X/a] = \frac{E[X]}{a}.$$

Markov inequality

Suppose that X is a non-negative random variable, then for any $a > 0$,

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

Proof

Define a random variable

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

Then $I \leq X/a$. Thus

$$P(X \geq a) = E[I] \leq E[X/a] = \frac{E[X]}{a}.$$

Note that the Markov inequality gives a trivial bound when $a \leq E[X]$. It only gives a non-trivial bound for $a > E[X]$.

As a consequence of the Markov inequality, we have the following

Chebyshev inequality

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

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

Note that the Markov inequality gives a trivial bound when $a \leq E[X]$. It only gives a non-trivial bound for $a > E[X]$.

As a consequence of the Markov inequality, we have the following

Chebyshev inequality

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

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

Note that the Markov inequality gives a trivial bound when $a \leq E[X]$. It only gives a non-trivial bound for $a > E[X]$.

As a consequence of the Markov inequality, we have the following

Chebyshev inequality

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

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

Proof

Since $(X-\mu)^2$ is a non-negative random variable with mean σ^2, we can apply the Markov inequality with $a=\epsilon^2$ to get

$$P(|X-\mu|\geq\epsilon)=P((X-\mu)^2\geq\epsilon^2)\leq\frac{\sigma^2}{\epsilon^2}.$$

Note that the Chebyshev inequality gives a trivial bound when $\epsilon^2\leq\sigma^2$. It only gives a non-trivial bound for $\epsilon^2>\sigma^2$.

Proof

Since $(X-\mu)^2$ is a non-negative random variable with mean σ^2, we can apply the Markov inequality with $a=\epsilon^2$ to get

$$P(|X-\mu| \geq \epsilon) = P((X-\mu)^2 \geq \epsilon^2) \leq \frac{\sigma^2}{\epsilon^2}.$$

Note that the Chebyshev inequality gives a trivial bound when $\epsilon^2 \leq \sigma^2$. It only gives a non-trivial bound for $\epsilon^2 > \sigma^2$.

Example 1

Suppose that it is known that the number of items produced in a certain factory during a week is a random variable X with mean 50.

(i) What can be said about the probability that this week's production will be at least 75?

(ii) If the variance of a week's production is known to be 25, then what can be said about the probability that this week's production will be between 40 and 60?

$$P(X \geq 75) \leq \frac{50}{75} = \frac{2}{3}.$$

Example 1

Suppose that it is known that the number of items produced in a certain factory during a week is a random variable X with mean 50.

(i) What can be said about the probability that this week's production will be at least 75?

(ii) If the variance of a week's production is known to be 25, then what can be said about the probability that this week's production will be between 40 and 60?

$$P(X \geq 75) \leq \frac{50}{75} = \frac{2}{3}.$$

$$P(40 \le X \le 60) = P(|X - 50| \le 10) = 1 - P(|X - 50| \ge 11) \ge 1 - \frac{25}{11^2} = \frac{96}{121}.$$

Chebyshev's inequality, although very simple, is very useful. For example, it can be used to prove the following very important result, the weak law of large numbers.

$$P(40 \leq X \leq 60) = P(|X - 50| \leq 10) = 1 - P(|X - 50| \geq 11)$$
$$\geq 1 - \frac{25}{11^2} = \frac{96}{121}.$$

Chebyshev's inequality, although very simple, is very useful. For example, it can be used to prove the following very important result, 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 with common (finite) mean $E[X_1] = \mu$. Then, for any $\epsilon > 0$,

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

I will give a proof under the additional assumption that the random variables $X_1, X_2, \ldots$ have a finite variance σ^2. The proof in the general case is more difficult.

Theorem (the weak law of large numbers)

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables with common (finite) mean $E[X_1] = \mu$. Then, for any $\epsilon > 0$,

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

I will give a proof under the additional assumption that the random variables $X_1, X_2, \ldots$ have a finite variance σ^2. The proof in the general case is more difficult.

Proof of the weak law of large numbers

Note that

$$E\left[\frac{X_1 + \cdots + X_n}{n}\right] = \mu$$

and

$$\mathrm{Var}\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{\sigma^2}{n}.$$

It follows from Chebyshev's inequality that for any $\epsilon > 0$,

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

Outline

The central limit theorem is one the most important results in probability theory. In Chapter 5, we have already seen a special case of this result. Here is the general result

Central limit theorem

Suppose that $X_1, X_2, \ldots$ are independent and identically distributed random variables with common mean μ and common variance σ^2. Then the distribution of

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

tends to the standard normal distribution as $n \to \infty$. That is, for any $a \in \mathbb{R}$,

$$P\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le a\right) \to \Phi(a), \quad \text{as } n \to \infty.$$

The central limit theorem is one the most important results in probability theory. In Chapter 5, we have already seen a special case of this result. Here is the general result

Central limit theorem

Suppose that $X_1, X_2, \dots$ are independent and identically distributed random variables with common mean μ and common variance σ^2. Then the distribution of

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

tends to the standard normal distribution as $n \to \infty$. That is, for any $a \in \mathbb{R}$,

$$P\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le a\right) \to \Phi(a), \quad \text{as } n \to \infty.$$

Note that the generality of the theorem above. The common distributions of $X_1, X_2, \ldots$ can be discrete, can be continuous, and can be neither discrete nor continuous. It can be regarded as universality law. I will try to give a proof this this result next time and give some applications.