Lecture 22: Weak law of large numbers and Moment genrating fucntion

Statistics 251

Yi Sun and Zhongjian Wang

Department of Statistics
The University of Chicago

Lecture Outline

Ineqalities for tail distribution

The weak law of large numbers

Moment Generating Function: Part I

Where are we?

Ineqalities for tail distribution

The weak law of large numbers

Moment Generating Function: Part I

Markov inequalities

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

$$\mathbb{P}\{X \geq a\} \leq \frac{\mathbb{E}[X]}{a}$$

Markov inequalities

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

$$\mathbb{P}\{X \geq a\} \leq \frac{\mathbb{E}[X]}{a}$$

Proof. Consider

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

Chebyshev's inequality

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

$$\mathbb{P}\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

Chebyshev's inequality

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

$$\mathbb{P}\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

Proof. Note that $(X - \mu)^2$ is then a nonnegative random variable, we can apply Markov's inequality to it.

Corollary

If $\mathrm{Var}(X) = 0$, then

$$\mathbb{P}\{X = \mathbb{E}[X]\} = 1$$

Example: the inequalities are inaccurate

- Consider X is uniformly distributed over the interval $(0, 10)$.
- What is its mean and variance?

Example: the inequalities are inaccurate

- Consider X is uniformly distributed over the interval $(0, 10)$.
- What is its mean and variance? $\mathbb{E}[X] = 5$, $\mathrm{Var}(X) = \frac{25}{3}$.
- How to estimate $P\{|X - 5| > 4\}$?

Where are we?

Ineqalities for tail distribution

The weak law of large numbers

Moment Generating Function: Part I

Theorem

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

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

How to prove it?

- Theorem applies without any additional assumption.

How to prove it?

- Theorem applies without any additional assumption.
- But here we assume that the random variables have a finite variance σ^2.

How to prove it?

- Theorem applies without any additional assumption.
- But here we assume that the random variables have a finite variance σ^2.
- Now,

$$E\left[\frac{X_1+\cdots+X_n}{n}\right]=\mu \text{ and } \operatorname{Var}\left(\frac{X_1+\cdots+X_n}{n}\right)=\frac{\sigma^2}{n}$$

How to prove it?

- Theorem applies without any additional assumption.
- But here we assume that the random variables have a finite variance σ^2.
- Now,

$$E\left[\frac{X_1 + \cdots + X_n}{n}\right] = \mu \text{ and } \operatorname{Var}\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{\sigma^2}{n}$$

- Lastly, from Chebyshev's inequality that

$$P\left\{\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right\} \leq \frac{\sigma^2}{n\varepsilon^2}.$$

How to prove it?

- Theorem applies without any additional assumption.
- But here we assume that the random variables have a finite variance σ^2.
- Now,

$$E\left[\frac{X_1 + \cdots + X_n}{n}\right] = \mu \text{ and } \operatorname{Var}\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{\sigma^2}{n}$$

- Lastly, from Chebyshev's inequality that

$$P\left\{\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right\} \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Example: Concentration of Gamma r.v.

If $\{X_i\}$ are independent gamma random variables with parameters $(1, 1)$, approximately how large need n be so that

$$P\left\{\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - 1\right| > .01\right\} < .01?$$

Where are we?

Ineqalities for tail distribution

The weak law of large numbers

Moment Generating Function: Part I

Definition

The moment generating function $M(t)$ of the random variable X is defined for all real values of t by

$$M(t) = \mathbb{E}\left[e^{tX}\right]$$
$$= \begin{cases} \sum_x e^{tx} p(x) & \text{if } X \text{ is discrete with mass function } p(x) \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & \text{if } X \text{ is continuous with density } f(x) \end{cases}$$

Derivaties

First,

$$\begin{aligned} M'(t) &= \frac{d}{dt}\mathbb{E}\left[e^{tX}\right] \\ &= \mathbb{E}\left[\frac{d}{dt}\left(e^{tX}\right)\right] \\ &= \mathbb{E}\left[Xe^{tX}\right] \end{aligned}$$

So,

$$M'(0) =$$

Derivaties

First,

$$M'(t) = \frac{d}{dt}\mathbb{E}\left[e^{tX}\right]$$
$$= \mathbb{E}\left[\frac{d}{dt}\left(e^{tX}\right)\right]$$
$$= \mathbb{E}\left[Xe^{tX}\right]$$

So,

$$M'(0) = \mathbb{E}[X].$$

Derivaties

First,

$$M'(t) = \frac{d}{dt}\mathbb{E}\left[e^{tX}\right] = \mathbb{E}\left[\frac{d}{dt}\left(e^{tX}\right)\right] = \mathbb{E}\left[Xe^{tX}\right]$$

So,

$$M'(0) = \mathbb{E}[X].$$

What's more?

$$M^n(t) = \mathbb{E}\left[X^n e^{tX}\right] \quad n \geq 1$$

implying that

Derivaties

First,

$$M'(t) = \frac{d}{dt}\mathbb{E}\left[e^{tX}\right]$$
$$= \mathbb{E}\left[\frac{d}{dt}\left(e^{tX}\right)\right]$$
$$= \mathbb{E}\left[Xe^{tX}\right]$$

So,

$$M'(0) = \mathbb{E}[X].$$

What's more?

$$M^n(t) = \mathbb{E}\left[X^n e^{tX}\right] \quad n \geq 1$$

implying that

$$M^n(0) = \mathbb{E}\left[X^n\right] \quad n \geq 1$$