Math 3280 20-11-30

Review

- Markov inequality. For a r.v. $X \geq 0$, and $a > 0$

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

- Chebyshev inequality: For $\varepsilon > 0$,

$$P\{ |X - \mu| \geq \varepsilon \} \leq \frac{Var(X)}{\varepsilon^2}$$

Prop 1. Let X be a r.v. with a finite mean μ. Suppose $Var(X) = 0$. Then

$$P\{ X = \mu \} = 1.$$

Proof. $$P\{ X \neq \mu \} = P\left\{ \bigcup_{k=1}^{\infty} |X - \mu| \geq \frac{1}{k} \right\}$$

(countable sub-addititivy)

$$\leq \sum_{k=1}^{\infty} P\left\{ |X-\mu| \geq \frac{1}{k} \right\}$$

(Chebyshev)

$$\leq \sum_{k=1}^{\infty} \frac{Var(X)}{(\frac{1}{k})^2} = 0$$

Hence $P\{X \neq \mu\} = 0$

So

$$P\{X=\mu\} = 1 - P\{X \neq \mu\} = 1.$$

$\boxtimes$

Thm 2 (The weak law of large numbers)

Let $X_1, X_2, \cdots, X_n, \cdots$ be an i.i.d sequence of r.v's, having a finite mean. Then for any $\varepsilon > 0$,

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

Pf. We prove the thm under an additional assumption that $Var(X_i) =: \sigma^2 < \infty$.

$$E\left[\frac{X_1 + \cdots + X_n}{n}\right] = \frac{1}{n} \sum_{k=1}^{n} E[X_k]$$

$$= \mu.$$

$$Var\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{1}{n^2} Var(X_1 + \cdots + X_n)$$

$$= \frac{1}{n^2} \cdot \sum_{k=1}^{n} Var(X_k)$$

(since $X_1, \cdots, X_n$ are independent)

$$= \frac{\sigma^2 \cdot n}{n^2} = \frac{\sigma^2}{n}.$$

Applying the Chebyshev inequality to $\frac{X_1+\cdots+X_n}{n}$,

we obtain

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

$$= \frac{\sigma^2}{n\varepsilon^2}$$

$\to 0$ as $n \to \infty$.

$\boxed{}$

Thm 3 (The central limit Thm).

Let $X_1, \cdots, X_n, \cdots,$ be an i.i.d. sequence of r.v.'s, each having a finite mean μ and a finite variance σ^2.

Then $\forall\, a \in \mathbb{R}$,

$$P\left\{ \frac{X_1+\cdots+X_n - n\mu}{\sqrt{n}\,\sigma} \leq a \right\} \longrightarrow \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx$$

as $n \to \infty$.

That is, the distribution of

$$\frac{X_1+\cdots+X_n - n\mu}{\sqrt{n}\,\sigma} = \frac{X_1+\cdots+X_n - n\mu}{\sqrt{Var(X_1+\cdots+X_n)}}$$

converges to the standard normal distribution, as $n \to \infty$.

The proof of the above Thm will be given in next class.

Example 4. If 10 fair dice are rolled, find the approximate prob. that the sum obtained is between 30 and 40.

Solution: Let X_i be the value obtained in the i-th roll, $i = 1, 2, \cdots, 10$.

We need to calculate

$$P\{ 29.5 \le X_1 + \cdots + X_{10} \le 40.5 \}$$

Notice that $\mu = E[X_i] = \frac{1}{6}(1+2+\cdots+6)$

$$= 7/2$$

$$E[X_i^2] = \frac{1}{6}\left(1^2 + 2^2 + \cdots + 6^2\right)$$

$$= \frac{1}{6} \cdot \frac{6 \times 7 \times 13}{6}$$

$$\sigma^2 = Var(X_i) = E[X_i^2] - E[X_i]^2$$

$$= \frac{35}{12}$$

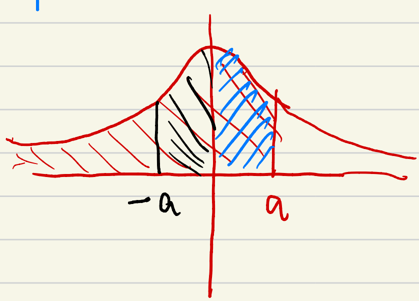
Hence $P\left\{ 29.5 \leq X_1 + \cdots + X_{10} \leq 40.5 \right\}$

$$= P\left\{ \frac{29.5 - 10 \times \frac{7}{2}}{\sqrt{350/12}} \leq \frac{X_1 + \cdots + X_{10} - 10 \times \frac{7}{2}}{\sqrt{10} \cdot \sqrt{\frac{35}{12}}} \leq \frac{40.5 - 10 \times \frac{7}{2}}{\sqrt{350/12}} \right\}$$

$$\approx P\{ -1.018 \leq Z \leq 1.018 \}$$

$$= 2 \cdot \Phi(1.018) - 1$$

$$\approx 0.692.$$



Example 5. Suppose a fair die is rolled for 100 times.

Let X_i be the value obtained in the i-th roll. Compute an approximation of

$$P\left\{ \prod_{i=1}^{100} X_i \leq a^{100} \right\}, \quad 1 < a < 6$$

Sketched solution:

$$P\left\{ \prod_{i=1}^{100} X_i \leq a^{100} \right\}$$

$$= P\left\{ \sum_{i=1}^{100} \log X_i \leq 100 \cdot \log a \right\}$$

(Letting $Y_i = \log X_i$)

$$= P\left\{ Y_1 + \cdots + Y_{100} \leq 100 \log a \right\}$$

$$= P\left\{ \frac{Y_1 + \cdots + Y_{100} - 100\, E[Y]}{10 \cdot \sqrt{\mathrm{Var}(Y)}} \leq \frac{100 \log a - 100\, E[Y]}{10 \sqrt{\mathrm{Var}(Y)}} \right\}$$

$$\approx \Phi\left(\frac{100 \log a - 100\, E[Y]}{10 \sqrt{\mathrm{Var}(Y)}} \right)$$

estimate $E[Y]$, $\mathrm{Var}[Y]$

and plug them in the expression.