Ch. 8 : Limit Theorems

We'll start with a couple inequalities which can be useful, for example, when the mean and possibly the variance of a distribution is known (but perhaps the probability distribution is not known)

Proposition 2.1 (Textbook, p. 367) **Markov's Inequality**

If X is a random variable with $X \geq 0$, then for any $a > 0$

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

Proof: (p. 367-368)

Suppose $a > 0$. Let X be a random variable with $X \geq 0$.

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

Observe
$$\frac{X}{a} \begin{cases} \geq 1 & \text{if } X \geq a \\ \geq 0 & 0 \leq X < a \end{cases}$$

So
$$\frac{X}{a} \geq I$$

Then
$$E\left[\frac{X}{a}\right] = \frac{1}{a}E[X] \geq E[I] = 1 \cdot P\{X \geq a\} + 0 \cdot P\{X < a\}$$

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

Proposition 2.2 (Textbook, p. 368) Chebyshev's Inequality

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

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

Proof: (p. 368)

Suppose $k>0$. Suppose X is a R.V. with finite mean μ and variance σ^2. Observe that $(X-\mu)^2$ is a R.V. with $(X-\mu)^2 \geq 0$. Then by Markov's Ineq.

$$P\{(X-\mu)^2 \geq k^2\} \leq \frac{E[(X-\mu)^2]}{k^2}$$

But $E[(X-\mu)^2] = \sigma^2$ $(Var(X))$

and $P\{(X-\mu)^2 \geq k^2\} = P\{|X-\mu| \geq k\}$

So

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

EX

Given $X \geq 0$, $E[X] = 75$, $\sigma^2 = 10$

Markov Ineq

a) $P\{X \geq 90\} \leq \frac{E[X]}{90} = \frac{75}{90} = \frac{5}{6}$

So $P\{X < 90\} > \frac{1}{6}$

May or may not be helpful information

Chebyshev Ineq

b) $P\{|X - 75| \geq 10\} \leq \frac{\sigma^2}{k^2} = \frac{10}{10^2} = \frac{1}{10}$

$P\{|X - 75| < 10\} \geq \frac{9}{10}$

$P\{65 < X < 85\} \geq \frac{9}{10}$

Recall

Ch. 5 result (Notes, p. 177)

DeMoivre-Laplace Limit Thm

X_B = Discrete Binomial Random Variable

$$p(i) = P\{X_B = i\} = \binom{n}{i} p^i (1-p)^{n-i}$$

$$E[X_B] = np$$

$$Var(X_B) = np(1-p)$$

X_B = # of successes that occur when n independent trials, each with success probability, p, are performed

Then, as $n \to \infty$

$$P\left\{a \le \frac{X_B - np}{\sqrt{np(1-p)}} \le b\right\} \to \Phi(b) - \Phi(a)$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

$$= P\{Z_1 \le x\}$$

standard normal

observe:

$$Z_1 = \frac{X - \mu}{\sigma}$$

$$E[X] = \mu \quad Var(X) = \sigma^2$$

Recall Also

Suppose $X_1, X_2, \dots X_n$ are independent and identically distributed R.V. each having c.d.f. F

expected value $E[X_i] = \mu$

variance $Var(X_i) = \sigma^2$

Define

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

S^2 = sample variance $= \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$

$E[\bar{X}] = \mu$

$Var(\bar{X}) = \frac{\sigma^2}{n}$

$E[S^2] = \sigma^2$

Weak Law of Large Numbers (Thm 2.1, p. 369)

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 $\epsilon > 0$

$$P\left\{ \left| \frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right| \geq \epsilon \right\} \to 0 \quad \text{as } n \to \infty$$

if you assume also $Var(X_i) = \sigma^2$ then use Chebyshev inequality

Strong Law of Large Numbers (Thm 4.1, p. 378)

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, with probability 1

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

That is,

$$P\left\{ \lim_{n \to \infty} \left[\frac{X_1 + X_2 + \cdots + X_n}{n} \right] = \mu \right\} = 1$$

Recall definition of sample mean (p. 283 textbook)

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

for independent + identically distributed R.V. X_i.

*

Central Limit Theorem (Thm 3.1, p.370)

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

$$\frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sigma\sqrt{n}} = \frac{\frac{X_1+X_2+\ldots+X_n}{n} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$

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

That is, for $-\infty < a < \infty$

$$P\left\{\left(\frac{(X_1+X_2+\ldots+X_n)-n\mu}{\sigma\sqrt{n}}\right) \leq a\right\} \to \boxed{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-\frac{x^2}{2}}\,dx} = \Phi(a)$$

as $n \to \infty$

see pp. 189-190 in textbook

Generalizes DeMoivre-LaPlace Limit Thm p.177

Recall Section 5.4.1 on the Normal Approximation to the Binomial Distribution - DeMoivre-LaPlace Thm p.194 textbook

- Recall, for independent $X_1, X_2, \ldots, X_n$

$$E[X_1 + \ldots + X_n] = E[X_1] + \ldots + E[X_n] = n \cdot \mu$$

$$Var(X_1 + \ldots + X_n) = Var(X_1) + \ldots + Var(X_n)$$
$$= n\sigma^2$$

so $\dfrac{(X_1+X_2+\ldots+X_n) - n\mu}{\sqrt{n\sigma^2}}$ is written in standard normal form.

$Z_i = \dfrac{X - \mu}{\sqrt{Var(X)}}$

$Z_i = \dfrac{X - \mu}{\sigma}$

EXAMPLES - see
Ex 3b (p.375)
Ex 3c (p.376)
Ex 3d (p.376)
Ex 3e (p.377)

EXAMPLE 3c (textbook, p.376)

10 fair dice are rolled. Use the Central Limit Theorem to find the approximate probability that the sum of the 10 dice is between 30 and 40, inclusive.

Sol: Let X_i denote the value of the i^{th} die $i = 1, 2, \ldots, 10$. Define $X = X_1 + X_2 + \ldots + X_{10}$.

- $E[X_i] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6}$

$$= \frac{1}{6}(1+2+3+4+5+6)$$

$$\mu = \quad = \frac{21}{6} = \boxed{\frac{7}{2}}$$

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

$$= 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} - \left(\frac{7}{2}\right)^2$$

$$= \frac{1}{6}[1 + 4 + 9 + 16 + 25 + 36] - \left(\frac{7}{2}\right)^2$$

$$\sigma^2 = \quad = \frac{91}{6} - \frac{49}{4} = \frac{182 - 147}{12} = \boxed{\frac{35}{12}}$$

(251.1)

Note $E[X] = E[\sum X_i] = \sum_i E[X_i] = n \cdot \frac{7}{2}$

$$\underbrace{Var(X) = Var(\sum X_i) = \sum_i Var(X_i)}_{\text{if } X_i\text{'s are pairwise indep.}} = n \cdot \frac{35}{12}$$

see notes p. 242

$X = X_1 + X_2 + \dots + X_n$

n	E[X]	Var(X)
1	3.5	$\frac{35}{12} \approx 2.92$
2	7	$2 \cdot \frac{35}{12} \approx 5.8$
3	10.5	≈ 8.75
4	14	≈ 11.67
5	17.5	≈ 14.58
6	21	≈ 17.5
⋮		
10	35	$35 \cdot \frac{10}{12} = \frac{350}{12}$

*

By the Central limit thm, and writing

$$X = X_1 + X_2 + \dots + X_{10}$$

continuous (normal)

$$P\{30 \le X \le 40\} \approx P\{29.5 \le X \le 40.5\}$$

discrete

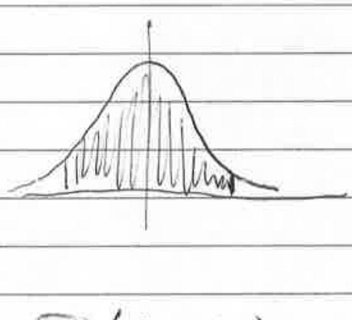
$$= P\left\{ \frac{29.5 - n\mu}{\sigma\sqrt{n}} \le \frac{X - n\mu}{\sigma\sqrt{n}} \le \frac{40.5 - n\mu}{\sigma\sqrt{n}} \right\}$$

$$\longrightarrow \left\langle \mu = \frac{7}{2} \quad n = 10 \quad \sigma^2 = \frac{35}{12} \right\rangle$$

$$= P\left\{ \frac{29.5 - 35}{\sqrt{\frac{350}{12}}} \le \frac{X - 35}{\sqrt{\frac{350}{12}}} \le \frac{40.5 - 35}{\sqrt{\frac{350}{12}}} \right\}$$

$$= P\left\{ \frac{-5.5}{5.4006} \le \frac{X - 35}{\sqrt{\frac{350}{12}}} \le \frac{5.5}{5.4006} \right\}$$

$$= P\left\{ -1.0184 \le \frac{X - 35}{\sqrt{\frac{350}{12}}} \le 1.0184 \right\}$$



$$= \Phi(1.0184) - \Phi(-1.0184)$$

$$= \Phi(1.0184) - \left(1 - \Phi(1.0184)\right)$$

$$= 2\Phi(1.0184) - 1$$

$$\approx 0.692$$

$\Phi(1.0184) \approx 0.845$

$1.690 - 1 = 0.69$

Book Section 5.4 see notes p. 167

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right]$$

To compute this exactly, we'd need

$$\underset{\sum_{i=1}^{10} x_i > 6}{\int \cdots \int} 1 \cdot dx_1\, dx_2 \cdots dx_{10}$$

$x_1 + x_2 + \cdots + x_{10} > 6$ ugh...

EXAMPLE 3d (Textbook, p. 376)

Let X_i, $i = 1, 2, \ldots, 10$ be independent random variables, each uniformly distributed on $(0,1)$. Use the central limit theorem to approximate $P\left\{ \sum_{i=1}^{10} X_i \geq 6 \right\}$

For each X_i recall (X_i uniform on $(0,1)$)

$$f(x) = \begin{cases} 1 & 0<x<1 \\ 0 & \text{otherwise} \end{cases}$$

- $E[X_i] = \int_0^1 x\,dx = \frac{1}{2}x^2 \Big|_0^1 = \boxed{1/2} = \mu$

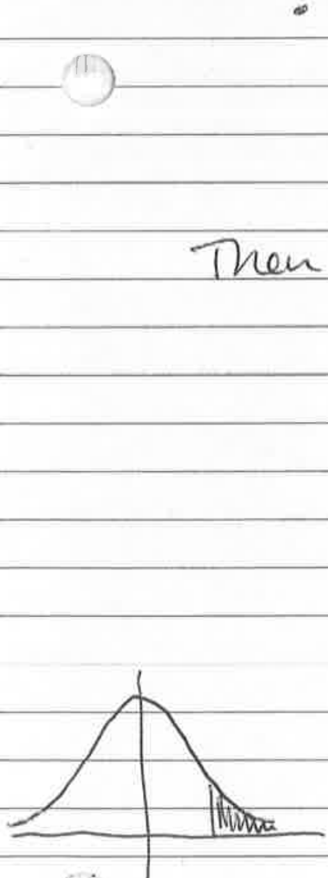
- $Var(X_i) = \int_0^1 x^2\,dx - \left(\frac{1}{2}\right)^2 = \frac{1}{3}x^3 \Big|_0^1 - \frac{1}{4}$

$$= \frac{1}{3} - \frac{1}{4} = \boxed{\frac{1}{12}} = \sigma^2$$

Then, by the central limit thm $(n=10, \mu = 1/2, \sigma^2 = \frac{1}{12})$

$$P\left\{ \sum_{i=1}^{10} X_i > 6 \right\} = P\left\{ \frac{\sum_{i=1}^{10} X_i - 10 \cdot \frac{1}{2}}{\sqrt{10 \cdot \frac{1}{12}}} > \frac{6 - 10\frac{1}{2}}{\sqrt{10\left(\frac{1}{12}\right)}} \right\}$$

$$= P\left\{ \frac{\sum_{i=1}^{10} X_i - 5}{\sqrt{\frac{10}{12}}} > \frac{1}{\sqrt{\frac{10}{12}}} \right\}$$



$$\approx 1 - \Phi\left(\sqrt{\frac{12}{10}}\right) = 1 - \Phi\left(\sqrt{1.2}\right)$$

$$\approx \boxed{0.1367}$$