*

EXAMPLE (4a, but see also Example 2c, notes p. 234)

Suppose $X_1, X_2, \ldots, X_n$ are independent and identically-distributed R.V.'s & each having cumulative distribution function F, expected value $E[X_i] = \mu$ for $i = 1, 2, \ldots, n$ and variance $Var(X_i) = \sigma^2$.

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

$X_i - \bar{X}$ = deviation (of X_i) for $i = 1, 2, \ldots, n$

random variable $\longrightarrow$ $S^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$ is called the sample variance

a) Find $Var(\bar{X})$

b) $E[S^2]$

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

see pp. 234 - 235

$Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$

recall $Var(aX) = a^2 Var(X)$

$E[\bar{X}^2] - (E[\bar{X}])^2$

$= \left(\frac{1}{n}\right)^2 Var\left(\sum_{i=1}^{n} X_i\right)$

using * from previous page

$= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} Var(X_i)$

$= \left(\frac{1}{n}\right)^2 n \cdot \sigma^2 = \boxed{\frac{\sigma^2}{n} = Var(\bar{X})}$

*

b)

$$E[S^2] = E\left[\sum_{i=1}^{n} \frac{(X_i-\bar{X})^2}{n-1}\right]$$

$$= E\left[\frac{1}{(n-1)} \sum_{i=1}^{n} \left(\underbrace{X_i-\mu} + \underbrace{\mu-\bar{X}}\right)^2\right]$$

$$= E\left[\frac{1}{(n-1)} \sum_{i=1}^{n} (X_i-\mu)^2 + 2(\mu-\bar{X})(X_i-\mu) + (\mu-\bar{X})^2\right]$$

$$= E\left[\frac{1}{(n-1)}\left(\sum_{i=1}^{n} (X_i-\mu)^2 + 2(\mu-\bar{X})\sum_{i=1}^{n}(X_i-\mu) + \sum_{i=1}^{n}(\mu-\bar{X})^2\right)\right]$$

$$= E\left[\frac{1}{n-1}\left(\sum_{i=1}^{n} (X_i-\mu)^2 + 2(\mu-\bar{X})\left(n\bar{X}-n\mu\right) + n(\mu-\bar{X})^2\right)\right]$$

$$= E\left[\frac{1}{n-1}\left(\sum_{i=1}^{n} (X_i-\mu)^2 + 2n(\mu-\bar{X})(\bar{X}-\mu) + n(\mu-\bar{X})^2\right)\right]$$

$$= E\left[\frac{1}{n-1}\left(\sum_{i}^{n} (X_i-\mu)^2 - n(\bar{X}-\mu)^2\right)\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n} E(X_i-\mu)^2 - \frac{n}{n-1}E\left[(\bar{X}-\mu)^2\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n} Var(X_i) - \frac{n}{n-1}Var(\bar{X})$$

$$= \frac{1}{n-1}(n\cdot\sigma^2) - \frac{n}{n-1}\left(\frac{\sigma^2}{n}\right) = \frac{n-1}{n-1}\sigma^2 = \boxed{\sigma^2}$$

$E[S^2]$ $\parallel$ $\boxed{\sigma^2}$

*

EXAMPLE (4b, Textbook, p. 308.)

Varience of Binomial R.V. (see also notes, p. 235)

Let X be a binomial R.V. with parameters n and p.

$$\left[\begin{array}{l} p(i) = \binom{n}{i} p^i (1-p)^{n-i} \qquad i = 0, 1, \ldots, n \\ \quad = P\{X = i\} \end{array}\right.$$

i = # of successes in n total trials

p = prob. of success

Let $X = X_1 + X_2 + \ldots + X_n$

see also p. 235

$$X_i = \begin{cases} 1 & \text{if } i^{\text{th}} \text{ trial is a success} \\ 0 & \text{otherwise } (i^{\text{th}} \text{ trial is failure}) \end{cases}$$

Note: The X_i's are independent Bernoulli R.V.'s

$E[X_i] = 1 \cdot p + 0 \cdot (1-p) = p$

$$Var(X) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = n \cdot Var(X_i)$$

for any i

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

note: $X_i^2 = X_i$

see notes pp. 123 - 125

so

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

$$= p - p^2$$

see Textbook, p. 132 Section 4.6.1

So $\boxed{Var(X) = n \cdot (p - p^2) = np(1-p)}$ ← as we've seen earlier

*

Ch. 8 : Limit Theorems

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

$$\boxed{\begin{array}{l}\textbf{Proposition 2.1} \text{ (Textbook, p. 367)} \quad \text{Markov's Inequality} \\ \text{If } X \text{ is a random variable with } X \geq 0 \text{, then for} \\ \text{any } a > 0 \qquad P\{X \geq a\} \leq \dfrac{E[X]}{a}\end{array}}$$

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

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

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

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

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

*

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

$$\frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sigma\sqrt{n}}$$

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

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

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

as $n \to \infty$

see pp. 189-190 in textbook

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

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

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

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

$$Z_1 = \frac{X - \mu}{\sqrt{Var(X)}}$$

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

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 $X = X_1+X_2+\ldots+X_{10}$.

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

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

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

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

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

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

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

~~[illegible]~~

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

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

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

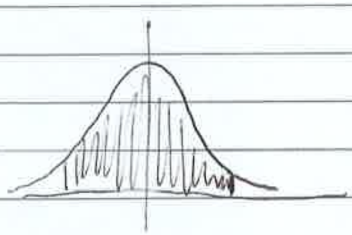
$$= P\left\{ \frac{29.5 - n\mu}{\sigma\sqrt{n}} \le \frac{X - n\mu}{\sigma\sqrt{n}} \le \frac{40.5 - n\mu}{\sigma\sqrt{n}} \right\}$$

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

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

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

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



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

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

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

$$\approx 0.692$$

Book Section 5.4

see notes p. 167

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy$$

$$= \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right]$$

EXAMPLE 3d (Textbook, p. 376)

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

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

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

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

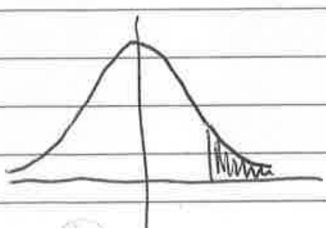
- $\mathrm{Var}(X_i) = \int_0^1 x^2\,dx - \left(\frac{1}{2}\right)^2 = \frac{1}{3}x^3\Big|_0^1 - \frac{1}{4}$

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

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

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

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



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

$$\approx \boxed{0.1367}$$

To compute this exactly, we'd need

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

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