Exercise 1.1.10

X , $E[X] = \mu$

$Var[X] = \sigma^2$

$$E[X^2] = Var[X] + (E[X])^2$$

$$\rightarrow Var[X] = E[X^2] - (E[X])^2$$

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

$Var[Y] = \sigma^2$

$E[Y] = \mu$

$$E[Y^2] = \mu^2 + \sigma^2$$

<u>Ex 1.2.5</u> $X_1, X_2, \dots, X_n$

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

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

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

Prop 1.2.6

independent $X_1, X_2, \ldots, X_n$

$E[X_i] = \mu$

$Var[X_i] = \sigma^2$

① $E[\bar{X}] = \mu$

② $Var[\bar{X}] = \frac{\sigma^2}{n}$

③ $E[S^2] = \sigma^2$; $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

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

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

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

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

(linearity of E)

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

(Ex. 1.1.10)

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

$$Var[\bar{X}] = \frac{\sigma^2}{n}$$

$$\therefore \ (Ex.\ 1.1.10) \quad E[\bar{X}^2] = (\mu)^2 + \left(\frac{\sigma^2}{n}\right)$$

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

$$= \frac{1}{n-1}\left[n(\mu^2+\sigma^2) - n\mu^2 - \sigma^2\right]$$

$$= \frac{1}{n-1}\left[\cancel{n\mu^2} + n\sigma^2 - \cancel{n\mu^2} - \sigma^2\right]$$

$$= \frac{1}{n-1}\left[n\sigma^2 - \sigma^2\right]$$

$$= \frac{1}{\cancel{n-1}}\left[\cancel{(n-1)}\sigma^2\right]$$

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

MARKOV INEQUALITY

Theorem 1.3.1

If a random variable X can only take nonnegative values, then

$$\boxed{\text{for all } a > 0, \quad P(X \geq a) \leq \frac{E[X]}{a}}$$

constant

$X \geq 0$ nonnegative

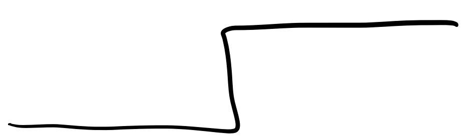
$\cancel{X > 0}$ ~~positive~~

Proof:

Fix $a > 0$

Define the new random variable Y_a

$$Y_a = \begin{cases} 0 & \text{if } X < a \\ a & \text{if } X \geq a \end{cases}$$



$X \geq 0 \quad Y_a = 0$

$X \geq Y_a$

$$\Rightarrow \quad Y_a \leq X \quad \text{for all } a, X$$

$$\Rightarrow E[Y_a] \leq E[X]$$

$$E[Y_a] = \sum_{y \in \mathrm{Im}(Y_a)} y \, P(Y_a = y)$$

$$= 0 \cdot P(Y_a = 0) + a \, P(Y_a = a)$$

$$= \underbrace{\cancel{0 \cdot P(X < a)}}_{=0} + a \, P(X \geq a)$$

$$\Rightarrow E[Y_a] = a \, P(X \geq a)$$

$$E[Y_a] \leq E[X]$$

$$\Rightarrow \quad a \, P(X \geq a) \leq E[X]$$

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

$$(a > 0)$$

Chebyshev's inequality

Theorem 1.3.4

If X is a random variable
with mean $E[X] = \mu$
and variance $Var[X] = \sigma^2$
then for all $c > 0$

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

Event $|X - \mu| \geq c$

$$\Leftrightarrow (|X - \mu|)^2 \geq c^2$$

$$\Leftrightarrow (X - \mu)^2 \geq c^2$$

Let $Y = (X - \mu)^2$

$Y \geq 0$

Apply Markov's inequality to Y, with c^2
$c^2 > 0$

$$P(Y \geq c^2) \leq \frac{E[Y]}{c^2}$$ (Markov's inequality)

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

$$\Leftrightarrow P\left(|X-\mu| \geq c\right) \leq \frac{Var[X]}{c^2}$$

$$\Leftrightarrow P\left(|X-\mu| \geq c\right) \leq \frac{\sigma^2}{c^2}$$

$\square$.

let $c = k\sigma > 0$; $k > 0$

$$\Leftrightarrow P\left(|X-\mu| \geq k\sigma\right) \leq \frac{\cancel{\sigma^2}}{k^2 \cancel{\sigma^2}}$$

$$\leq 1/k^2$$

$$\Leftrightarrow P\left(|X-\mu| < k\sigma\right) \geq 1 - 1/k^2$$

$$1 - P\left(|X-\mu| \geq k\sigma\right) \geq 1 - 1/k^2$$

$$P\left(|X-\mu| < k\sigma\right) \geq 1 - 1/k^2$$

Example 1.3.6

Suppose a country is taking part in a vote and an unknown proportion p of voters supports option A. $(A, B, C, \ldots)$

Suppose we interview a sample of n voters and $\hat{p}$ proportion of this sample support option A.

How close is $\hat{p}$ to p?

Let us label our sample of voters $i = 1, 2, \ldots, n$ and let X_i

$x_i = 1$ if voter will support A

$x_i = 0$ otherwise

$$X_i \sim \mathrm{Bern}(\underset{\text{unknown}}{p})$$

$$\hat{p} = \frac{\sum x_i}{n} = \bar{x}$$

$E[X_i] = p$

$Var[X_i] = p(1-p)$

$X_1, \dots, X_n$, $E[X_i] = \mu$, $Var[X_i] = \sigma^2$

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

$Var[\bar{X}] = \frac{\sigma^2}{n}$

Prop 1.2.6 : $\bar{X}$

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

$$Var[\bar{X}] = \frac{p(1-p)}{n} = \frac{\sigma^2}{n}$$

Using Chebyshev's inequality

$$P(|\bar{X} - p| \geq \epsilon) \leq \frac{\left(\frac{p(1-p)}{n}\right)}{\epsilon^2} \quad (\epsilon > 0)$$

$$p(1-p) \leq \frac{1}{4} \quad \text{(earlier result)}$$

$$\Rightarrow P(|\bar{X} - p| \geq \epsilon) \leq \frac{1}{4n\epsilon^2}$$

Example $\epsilon = 0.1$

$n = 100$

$$P(|\bar{X} - p| \geq 0.1) \leq \frac{1}{4 \cdot 100 \cdot (0.1)^2}$$

$$\leq 0.25$$

$\hat{p}$