Stat 401 - 10/30/17 Study Guide Available (Exam Materials) ①
HW: Sec 2.8: 2, 5, 6 Sec 3.2: 1, 2, 8, 13a
Sec 3.1: 1, 4, 5, 6 ↳ round mean to an integer.

## Sec 2.8 - Linear Combinations of RVs

Setup: Let $X_1, \ldots, X_n$ be $n$ RVs

Let $\vec{X} = (X_1, \ldots, X_n)^T$ ← transpose.
be a random vector.

Let $T = T(X_1, \ldots, X_n)$ be a function of the $X_i$'s.

def] Linear Combo (in this sec)

We take $T = a_1 X_1 + \cdots + a_n X_n$

$$= \sum_{i=1}^{n} a_i X_i$$

where $a_i$'s are constants.

When $T$ is of this form, it is said to be linear.

Thm] $\mathbb{E}[T] = \sum_{i=1}^{n} a_i \mathbb{E}[X_i]$

Proof] $\mathbb{E}[T] = \mathbb{E}[a_1 X_1 + a_2 X_2 + \cdots + a_n X_n]$

$$= a_1 \mathbb{E}[X_1] + a_2 \mathbb{E}[X_2] + \cdots + a_n \mathbb{E}[X_n]$$

$$= \sum_{i=1}^{n} a_i \mathbb{E}[X_i]$$

□

<u>Thm</u>) Suppose $T = \sum_{i=1}^{n} a_i X_i$

Suppose $W = \sum_{j=1}^{m} b_j Y_j$

Then

$$\mathrm{Cov}(T, W) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j \, \mathrm{Cov}(X_i, Y_j)$$

Proof sketch

go to def of cov.

plug in $T = \Sigma \quad W = \Sigma$

Simplify

$\vdots$

def of cov

<u>Corollary</u>) Let $T = \sum_{i=1}^{n} a_i X_i$

then

$$\mathrm{Var}(T) = \mathrm{Cov}(T, T)$$

Proof)

$$\mathrm{Cov}(X, X) = \mathbb{E}[X \cdot X] - \mathbb{E}[X]\,\mathbb{E}[X]$$
$$= \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$
$$= \mathrm{Var}[X]$$

$$\mathrm{Cov}(T, T) = \sum \sum a_i a_j \,\mathrm{Cov}(X_i, X_j)$$
$$= \sum a_i^2 \,\mathrm{Var}(X_i) + 2 \sum_{i<j} a_i a_j \,\mathrm{Cov}(X_i, X_j)$$
$$= \mathrm{Var}(T)$$

Corollary) If $X_1, \ldots, X_n$ are indep. RVs, then

$$\mathrm{Var}(T) = \sum_{i=1}^{n} a_i^2 \,\mathrm{Var}(X_i)$$

and $\mathrm{Cov}(X_i, X_j) = 0$

def) Sample Mean.

Let $X_1, \ldots, X_n$ be iid RVs with common mean $\mu$ and variance $\sigma^2$.

Often written as

$$X_1, \ldots, X_n \overset{iid}{\sim} (\mu, \sigma^2)$$

the sample mean is defined as

$$\overline{X} = n^{-1} \sum_{i=1}^{n} X_i$$

ex) Let $X_1, \ldots, X_n \overset{iid}{\sim} (\mu, \sigma^2)$

Find $\mathbb{E}[\overline{X}]$. Find $Var[\overline{X}]$.

By Thm $\mathbb{E}[\overline{X}] = n^{-1} \sum_{i=1}^{n} \mathbb{E}[X_i]$

$$= \mathbb{E}\left[\frac{1}{n}(X_1 + \cdots + X_n)\right]$$

$$= \frac{1}{n} \mathbb{E}[X_1 + \cdots + X_n]$$

$$= \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i]$$

$$= \frac{1}{n} \sum_{i=1}^{n} \mu$$

$$= \frac{1}{n} \left( \underbrace{\mu + \cdots + \mu}_{n \text{ times}} \right)$$

$$= \frac{1}{n} n \cdot \mu = \boxed{\mu}$$

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

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

$$= \frac{1}{n^2} \left[ \underbrace{\sigma^2 + \cdots + \sigma^2}_{n \text{ of them}} \right]$$

sample mean.

$$= \frac{1}{n^2} \cdot n \sigma^2 = \boxed{\frac{\sigma^2}{n}}$$

$$\bar{X} = \frac{1}{n} (X_1 + \cdots + X_n)$$

def) Unbiased Estimator.

An estimator $T$ (from a sample) for a parameter $\theta$ is unbiased if

$$\mathbb{E}[T] = \theta$$

Bias refers to whether an estimator tends to over/underestimate the parameter.

Note: $\bar{X}$ is an unbiased estimator of $\mu$.

sample mean

pop. mean.

def). Sample Variance

Let $X_1, \ldots, X_n \overset{iid}{\sim} (\mu, \sigma^2)$

Define $S^2$ as

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

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

sample variance

Note: $S^2$ is an unbiased estimator of $\sigma^2$

pop. variance.

Proof

Recall $Var[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

$$\sigma^2 = \mathbb{E}[X^2] - \mu^2$$

$$\Rightarrow \sigma^2 + \mu^2 = \mathbb{E}[X^2]$$

Also $Var[\bar{X}] = \mathbb{E}[\bar{X}^2] - (\mathbb{E}[\bar{X}])^2$

$$\frac{\sigma^2}{n} + \mu^2 = \mathbb{E}[\bar{X}^2]$$

Need to show $\mathbb{E}[S^2] = \sigma^2$

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

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

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

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

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

$$= \frac{1}{n-1}(n-1)\sigma^2 = \sigma^2 \checkmark$$

---

$$\mathbb{E}[n\bar{X}^2] = n\,\mathbb{E}[\bar{X}^2]$$

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

ex) Toss a 6-sided die 100 times.

Let $X_k$ = value obtained on the $k^{th}$ toss

a) State $D_{X_k} = \{1, 2, 3, 4, 5, 6\}$

b) Find $\mu_{X_k} = \mathbb{E}[X_k]$

$$= 1\left(\tfrac{1}{6}\right) + 2\left(\tfrac{1}{6}\right) + 3\left(\tfrac{1}{6}\right) + 4\left(\tfrac{1}{6}\right) + 5\left(\tfrac{1}{6}\right) + 6\left(\tfrac{1}{6}\right)$$

$$= 3.5$$

c) Find $Var[X_k]$

$$= \mathbb{E}[X_k^2] - (\mathbb{E}[X_k])^2$$

$$= \frac{1}{6}(1^2 + \cdots + 6^2) - (3.5)^2$$

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

d) Is it reasonable to assume that $X_1, \ldots X_{100}$ are independent?

$\rightarrow$ Yes.

e) Find $\mathbb{E}[\bar{X}]$, $Var[\bar{X}]$,

assuming independence and identically distributed

$$\mathbb{E}[\bar{X}] = \mu_{X_k} = 3.5$$

$$Var[\bar{X}] = \frac{Var[X_k]}{n} = \frac{35/12}{100}$$

$$= 7/240$$

f) Assume indep.

Let $Y = 3X_2 - X_1$

Suppose $\sigma_1^2 = k$ and

$\sigma_2^2 = 2$.

If $Var[Y] = 25$, find $k$.

$$25 = Var[Y]$$

$$= Var[3X_2 - X_1]$$

$$\overset{\text{indep.}}{=} (3)^2\, Var(X_2) + (-1)^2\, Var(X_1)$$

$$= 9\sigma_2^2 + \sigma_1^2$$

$$= 9(2) + \sigma_1^2$$

$$25 = 18 + \sigma_1^2$$

$$\boxed{7 = \sigma_1^2 = k}$$