cse 107 11-18-25

mid2: Thur 11/20

hw6: due tonight (11/18) 10:00 PM

## 4.2 Covariance & Correlation

**Defn**

The covariance of r.v.s $X, Y$ is

$$\mathrm{cov}(X,Y) = E[(X-E[X])(Y-E[Y])]$$

observe

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

Defn

- We say $X, Y$ are <u>uncorrelated</u> iff $\mathrm{Cov}(X,Y) = 0$.

- if $\mathrm{Cov}(X,Y) > 0$, we say $X, Y$ are <u>positively correlated</u>.

- if $\mathrm{Cov}(X,Y) < 0$, we say $X, Y$ are <u>negatively correlated</u>.

Theorem

$$\mathrm{Cov}(X,Y) = E[X \cdot Y] - E[X] \cdot E[Y]$$

Proof

$$\mathrm{Cov}(X,Y) = E\left[(X - \overbrace{E[X]}^{\mu_X})(Y - \overbrace{E[Y]}^{\mu_Y})\right]$$

$$= E\left[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y\right]$$

$$= E[XY] - \cancel{\mu_X \mu_Y} - \cancel{\mu_Y \mu_X} + \cancel{\mu_X \mu_Y}$$

$$= E[XY] - E[X] \cdot E[Y]$$

∎

<u>note</u>

$$X, Y \text{ indep.} \Rightarrow E[XY] = E[X]E[Y]$$

$$\Rightarrow \mathrm{Cov}(X,Y) = 0$$

$$\Rightarrow X, Y \text{ uncorrelated.}$$

<u>note</u>: Converse is false!

See ex. 4.13 on P. 218–219.

<u>Defn</u>

The <u>correlation coefficient</u> $\rho(X,Y)$

is

$$\rho(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\cdot\operatorname{Var}(Y)}}$$

It can be shown that

$$-1 \le \rho(X,Y) \le 1$$

Perfect neg. correlation ($-1$)

Perfect Pos. correlation ($1$)

Ex

Toss a coin $n$ times, where $P(\text{head}) = p$. let

$$X = \#\text{heads}$$
$$Y = \#\text{tails}$$

compute $\rho(X, Y)$.

solution

observe $X + Y = n$. so

$$n = E[n] = E[X+Y] = E[X] + E[Y]$$

and

$$\mathrm{Var}(Y) = \mathrm{Var}(-X+n) = (-1)^2 \mathrm{Var}(X) = \mathrm{Var}(X)$$

also

$$(X-E[X])+(Y-E[Y])=(X+Y)-E[X+Y]$$

$$= n - n = 0$$

$$\therefore \quad Y-E[Y] = -(X-E[X])$$

Thus

$$\mathrm{Cov}(X,Y) = E[(X-E[X])(Y-E[Y])]$$

$$= -E[(X-E[X])^2]$$

$$= -\mathrm{Var}(X)$$

and

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)\cdot Var(Y)}}$$

$$= \frac{-Var(X)}{\sqrt{Var(X)^2}}$$

$$= -\frac{Var(X)}{Var(X)} = \boxed{-1}$$

<u>Exercise</u> investigate: are $X, Y$ independent? ans. no!!

Exercise

Prove the following.

(1) $Cov(aX+b, Y) = a\,Cov(X,Y)$

(2) $Cov(X,Y) = Cov(Y,X)$

(3) if $a>0$, then $\rho(aX+b,Y) = \rho(X,Y)$

(4) $Cov(aX+bY, Z) = a\,Cov(X,Z) + b\,Cov(Y,Z)$

# Variance of a Sum

Let $X, Y$ be r.v.s (not necessarily independent). Then

$$\boxed{\begin{aligned} &Var(X+Y) \\ &= Var(X) + Var(Y) + 2\,Cov(X,Y) \end{aligned}}$$

more generally:

$$\boxed{\begin{aligned} &Var\left(\sum_{i=1}^{n} X_i\right) \\ &= \sum_{i=1}^{n} Var(X_i) + \sum_{\substack{(i,j) \\ i \neq j}} Cov(X_i, X_j) \end{aligned}}$$

## Proof of case $n=2$

Let $U = X - E[X] \quad \therefore E[U] = 0$

$V = Y - E[Y] \quad \therefore E[V] = 0$

also $E[U^2] = E[(X - E[X])^2] = Var(X)$

$E[V^2] = \cdots = Var(Y)$

Thus

$$Var(X+Y) = Var(U + E[X] + V + E[Y])$$

$$= Var(U+V)$$

$$= E[((U+V) - 0)^2]$$

$$= E\left[(U+V)^2\right]$$

$$= E\left[U^2 + V^2 + 2UV\right]$$

$$= E\left[U^2\right] + E\left[V^2\right] + 2E\left[UV\right]$$

$$= Var(X) + Var(Y) + 2\,Cov(X,Y)$$

$\blacksquare$

See ex. 4.15 on p. 221.