Lecture 11

Recap:

- Linearity of Expectation:

  ⇒ We can decompose some discrete RV's as sums of indicator variables or smaller discrete events

- Independence and Expectation/Variance:

  ⇒ Let $X_1, \dots, X_n$ be n independent RV's then:

  ① $E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i)$

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

What if $X_1, \dots, X_n$ are dependent?

8.4

Covariance

- Let X and Y be RV defined on the same sample space s.t.

  $E(X) = \mu_X$
  $E(Y) = \mu_Y$ then:

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

Ex) What is $Cov(X,X)$?

$$\Rightarrow Cov(X,X) = E[(X-E(X))(X-E(X))] = Var(X)$$

- Alternative Form:

$$\left.\underline{Cov(X,Y) = E(XY) - \mu_X \mu_Y}\right\} \text{ useful for computation}$$

Proof:

Expand $E[(X-\mu_X)(Y-\mu_Y)]$ and simplify

- $Cov(X,Y) = E[(X-\mu_x)(Y-\mu_x)]$

$\Rightarrow$ What if $Cov(X,Y) > 0$?

- When $X-\mu_x > 0$ (X is above average) then $Y-\mu_x > 0$ on average.

  $\Rightarrow$ X and Y are <u>positively</u> correlated

$\Rightarrow$ What if $Cov(X,Y) < 0$?

- When $X-\mu_x > 0$ (X is above average) then $Y-\mu_x < 0$ on average.

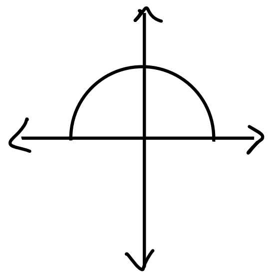
  $\Rightarrow$ X and Y are <u>negatively</u> correlated

$\Rightarrow$ What if $Cov(X,Y) = 0$?

$\Rightarrow$ As X changes Y changes independently, the negative changes cancel the positive ones.

Ex)

Let $f_{X,Y}(x,y) = \frac{2}{\pi} I((x,y) \text{ s.t. } y \geq 0 \text{ and } x^2+y^2 \leq 1)$



① Do we expect negative or positive correlation?

② What is $Cov(X,Y)$?

①

② $Cov(X,Y) = E(XY) - E(X)E(Y)$

◦ $E(XY) = \frac{2}{\pi} \int_{-1}^{1} \int_{0}^{\sqrt{1-x^2}} xy \, dy\, dx$

$= \frac{2}{\pi} \int_{-1}^{1} \frac{xy^2}{2} \Big|_{0}^{\sqrt{1-x^2}} dx = \frac{1}{\pi} \int_{-1}^{1} x(1-x^2)\, dx$

$= \frac{1}{\pi} \int_{-1}^{1} x - x^3 \, dx = \frac{1}{\pi} \left(\frac{x^2}{2} - \frac{x^4}{4} \right) \Big|_{-1}^{1}$

$= \frac{1}{\pi} \left(\frac{1}{2} - \frac{1}{4} - \frac{1}{2} + \frac{1}{4} \right)$

$= 0$

$\circ$ $E(Y) = \frac{2}{\pi}\int_{-1}^{1}\int_{0}^{\sqrt{1-x^2}} y \, dy\,dx = \frac{2}{\pi}\int_{-1}^{1} \frac{y^2}{2}\Big|_0^{\sqrt{1-x^2}} dx = \frac{1}{\pi}\int_{-1}^{1} 1-x^2 dx$

$= \frac{1}{\pi}\left(x - \frac{x^3}{3}\right)\Big|_{-1}^{1} \quad \frac{1}{\pi}\left(1-\frac{1}{3}+1-\frac{1}{3}\right) = \frac{4}{3\pi}$

$= \frac{4}{3\pi}$

$\circ$ $E(X) = \frac{2}{\pi}\int_{-1}^{1}\int_{0}^{\sqrt{1-x^2}} x \, dy\,dx = \frac{2}{\pi}\int_{-1}^{1} x(1-x^2)^{\frac{1}{2}} dx$

$U = 1-x^2 \quad du = -2x\,dx$

$= \frac{2}{\pi}\int -\frac{1}{2}(u)^{\frac{1}{2}} du = -\frac{1}{\pi} u^{3/2}\left(\frac{2}{3}\right)$

$= \frac{-2}{3\pi}(1-x^2)^{3/2}\Big|_{-1}^{1} = 0.$

So $E(X)\cdot E(Y) = 0$

$\Rightarrow Cov(X,Y) = E(XY) - E(X)\cdot E(Y) = 0 - 0\left(\frac{1}{3\pi}\right) = 0.$

So X and Y are uncorrelated.

Thm

If X and Y are independent then:

$Cov(X, Y) = 0$

★ Converse is not necessarily true ★

Covariance of Indicators

- Let I_A be an indicator variable for event A
- Let I_B be an indicator variable for event B

$$\Rightarrow Cov(I_A, I_B) = E(I_A \cdot I_B) - E(I_A) \cdot E(I_B)$$

$$I_A \cdot I_B = \begin{cases} 1 & \text{if } A \cap B \\ 0 & \text{else} \end{cases}$$

$$\begin{aligned} Cov(I_A, I_B) &= P(A \cap B) - P(A) \cdot P(B) \\ &= P(B) \cdot P(B|A) - P(A) \cdot P(B) \\ &= P(B)\underbrace{\left(P(B|A) - P(B)\right)} \end{aligned}$$

- If A increases the chances of B

  $\Rightarrow P(B|A) > P(B)$ so $Cov > 0$
- Similar for $Cov < 0$, $Cov = 0$

Ex)

Imagine you roll 2 dice.

Is the sum of the 2 dice being over 10 correlated with the second roll being a 6? If so positive or negative?

Let $S = X_1 + X_2$

So we want:

$$\text{Cov}(I(S>10), I(X_2=6))$$

$$= P(S>10, X_2=6) - P(S>10) \cdot P(X_2=6)$$

$S>10$: $\left.\begin{matrix} 5+6 \\ 6+6 \\ 6+5 \end{matrix}\right\}$ 3 combos

$S>10, X_2=6$: $\left.\begin{matrix} 5+6 \\ 6+6 \end{matrix}\right\}$ 2 combos

$$= \frac{2}{36} - \frac{3}{36} \cdot \frac{6}{36} > 0$$

So $S>10$ and $X_2=6$ are positively correlated.

Covariance Properties

Let X, Y be RV s.t. $Cov(X,Y)$ is well-defined then:

- ① $Cov(X,Y) = Cov(Y,X)$
- ② $Cov(a, X) = 0 \quad \forall a \in \mathbb{R}$
- ③ $Cov(aX, Y) = a\,Cov(X,Y) \; \forall a \in \mathbb{R}$

④ $Cov(X,X) = Var(X)$

⑤ $Cov(aX+b, Y) = a\,Cov(X,Y)$

$a, b \in \mathbb{R}$

Billinearity of Covariance

Provided all the Covariances exist:

For RV $X_1, \dots, X_m$, $Y_1, \dots, Y_n$,

and $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbb{R}$

then:

$$\mathrm{Cov}\left(\sum_{j=1}^{m} a_j X_j, \sum_{i=1}^{n} b_i Y_i\right) = \sum_{j=1}^{m}\sum_{i=1}^{n} a_j b_i \,\mathrm{Cov}(X_j, Y_i)$$

Proof

See textbook proof for Fact 8.33, pg 290

Multinomial Distribution

Let $(X_1, \dots, X_r) \sim \text{Multinom}(n, r, p_1, \dots, p_r)$

1. r categories
2. X_i is the total number of i'th outcomes, each with prob $= p_i$
3. $p_1 + \dots + p_r = 1$

Ex)

Find $\text{Cov}(X_i, X_j)$ for $i, j \in \{1, \dots, n\}$

1. Decompose X_i and X_j into indicator variables

   - Let $I_{k,i} = \begin{cases} 1 & \text{if trial } k \text{ is a success for category } i \\ 0 & \text{else} \end{cases}$

   - So $X_i = \sum_{k=1}^{n} I_{k,i}$

- If $i = j$:

$$\text{Cov}(X_i, X_i) = \text{Var}(X_i) = np_i(1 - p_i)$$

- If $i \neq j$:

$$Cov(X_i, X_j) = Cov\left(\sum_{k=1}^{n} I_{k,i}, \sum_{t=1}^{n} I_{t,j}\right)$$

$$= \sum_{k=1}^{n} \sum_{t=1}^{n} Cov(I_{k,i}, I_{t,j})$$

We know each trial is independent so
$Cov(I_{k,i}, I_{t,j}) = 0 \quad \forall \ k \neq t$

$$\Rightarrow \sum_{k=1}^{n} Cov(I_{k,i}, I_{k,j})$$

$$Cov(I_{k,i}, I_{k,j}) = \underbrace{E(I_{k,i} \cdot I_{k,j})} - E(I_{k,i}) E(I_{k,j})$$

$$= 0 - p_i p_j < 0$$

trial k cannot
be a success for
both i and j

So $Cov(X_i, X_j) = \sum_{k=1}^{n} -p_i p_j = -n p_i p_j$

Variance of Sums

- If $X_1, \dots, X_n$ are <u>independent</u> RV then:

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

- If X, Y are 2 RV then:

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

assuming finite variances and covariances

More generally:

- If $X_1, \dots, X_n$ are n RV's then:

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + \sum_{j=1}^{n} \sum_{i=1}^{n} Cov(X_i, X_j)$$

$$= \sum_{i=1}^{n} Var(X_i) + 2 \sum_{\underbrace{1 \le i < j \le n}_{\substack{i<j \\ i \neq j}}} Cov(X_i, X_j)$$

Proof:

$$Var\left(\sum_{1}^{n} X_i\right) = Cov\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} Cov(X_i, X_j)$$

$$= \sum_{i=1}^{n} Var(X_i, X_j) + \sum_{i \neq j} Cov(X_i, X_j)$$

$\square$

Ex)

Let $(X_1, \ldots, X_r) \sim \text{Multinom}(n, r, p_1, \ldots, p_r)$

Find $Var(X_1 + X_2)$.

$$Var(X_1 + X_2) = Var(X_1) + Var(X_2) + 2Cov(X_1, X_2)$$

$$= n(p_1(1-p_1) + p_2(1-p_2) - 2p_1p_2)$$

$$= n(p_1 - p_1^2 + p_2 - p_2^2 - 2p_1p_2)$$

$$= n(p_1 + p_2 - (p_1^2 + p_2^2 + 2p_1p_2))$$

$$= n(p_1 + p_2 - (p_1 + p_2)^2)$$

$$= n(p_1 + p_2)(1 - (p_1 + p_2))$$