Stat 230: Probability

Lecture 28

Jeremy VanderDoes

University of Waterloo

Wednesday, July 13th

Review

Example

Suppose X and Y have joint probability function given by the following table:

	f(x,y)	y = 0	y = 1	y = 2
x	-1	.15	.1	.1
	2	.2	.15	k

Compute Cov(2X, Y).

Review

Last time we talked about:

(1) Functions of joint variables
(2) Expectation Functions of joint variables
(3) Covariance
(4) Correlation

For today:

(1) Correlation
(2) Linear combinations of random variables
(3) Indicator random variables

Review

- Monday: Quiz 6
- Exam

Correlation

Definition

The **correlation** *of X and Y is denoted $corr(X, Y)$, and is defined by*

$$corr(X, Y) = \rho = \frac{Cov(X, Y)}{SD(X)SD(Y)}$$

It follows from the Cauchy-Schwarz inequality that $-1 \leq corr(X, Y) \leq 1$, and if $|corr(X, Y)| = 1$, $X = aY + b$.

Correlation

Definition

We say that X and Y are uncorrelated if $Cov(X, Y) = 0$ (or $corr(X, Y) = 0$).

Remark

If X and Y are independent, then X and Y are uncorrelated.

Remark

$Cov(X, X) = Var(X)$

Correlation

Example

Suppose X and Y have joint probability function given by the following table:

	f(x,y)	x = 0	x = 1	x = 2
y	0	0.2	0.3	0.1
	2	0.25	0.13	0.02

Compute $Corr(X, Y)$.

Linear Combination

Definition

Suppose that $X_1, ..., X_n$ *are jointly distributed RVs with joint probability function* $f(x_1, ..., x_n)$*. A* **linear combination** *of the RVs* $X_1, ..., X_n$ *is any random variable of the form*

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

where $a_1, ..., a_n \in \mathbb{R}$*. If* $\mathbf{X} = (X_1, ..., X_n)^\top$, $\mathbf{a} = (a_1, ..., a_n)^\top$*, then a linear combination is*

$$\mathbf{X}^\top \mathbf{a}$$

Linear Combination

Remark

Some "famous" linear combinations (ones you will see in Stat 231) are

(1) *The total*

$$T = \sum_{i=1}^{n} X_i \qquad 1 \leq i \leq n \qquad (a_i = 1)$$

(2) *The sample mean*

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

Linear Combination

Remark

Expected Value of a Linear Combination:

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

Follows directly from the linearity property of the expected value.

Example

Suppose $X \sim N(1, 1)$ and $Y \sim U(0, 1)$. Compute $E(2X - 4Y)$.

Linear Combination

Remark

Variance of a Linear Combination:

$$Var(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i) + \sum_{1 \le i \ne j \le n} a_i a_j Cov(X_i, X_j). \quad (1)$$

$$= \sum_{i=1}^{n} a_i^2 Var(X_i) + 2 \sum_{1 \le i < j \le n} a_i a_j Cov(X_i, X_j) \quad (2)$$

If $X_1, X_2, ..., X_n$ are mutually uncorrelated ($Cov(X_i, X_j) = 0,$ for $i \ne j$),

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

An important case where this holds is when $X_1, X_2, ..., X_n$ are independent.

Linear Combination

Example

Suppose $X \sim N(3,4)$, and $Y \sim U(0,1)$, and $Cov(X,Y) = -0.1$, compute $Var(2X - Y)$.

Indicator

Definition

Let $A \subset S$ be an event. We say that $\mathbb{1}_A$ is the **indicator** *random variable of the event A. $\mathbb{1}_A$ is defined by:*

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \omega \in A, \\ 0 & \omega \in \bar{A} \end{cases}$$

Such variables are often termed Bernoulli Random Variables.