Probability Theory Lecture 07

Padmini Mukkamala

Budapest University of Technology and Economics

padmini.mvs@gmail.com

November 8, 2021

Overview

Definitions

Joint Probability Mass Function

Given two discrete random variables X and Y, the **joint probability mass function** is a function, $p_{X,Y} : \mathcal{R}^2 \to [0,1]$, such that,

$$\sum_x \sum_y p_{X,Y}(x,y) = 1$$

Here we think of $p_{X,Y}(x,y)$ as the probability that $X = x$ and $Y = y$.
Where its obvious, the subscript X, Y is dropped and it is written as $p(x,y)$.

Definitions contd.

Marginal Probability Mass Functions

The probability mass functions of X, $p_X(x)$ and of Y, $p_Y(y)$ are called the marginal probability mass functions.

Notice that we can obtain the marginal pmfs from the joint pmf as follows:

$$p_X(x) = \sum_y p_{X,Y}(x, y)$$

$$p_Y(y) = \sum_x p_{X,Y}(x, y)$$

Examples

Suppose that a letter is randomly chosen from the set $\{e, x, a, m\}$. We define three random variables as follows:

$$X(letter) = \begin{cases} 1 & \text{letter is e or m} \\ 0 & \text{otherwise} \end{cases},$$

$$Y(letter) = \begin{cases} 1 & \text{letter is e} \\ 2 & \text{letter is x} \\ 3 & \text{letter is a} \\ 4 & \text{letter is m} \end{cases},$$

$$Z(letter) = \begin{cases} 1 & \text{letter is x or m} \\ 0 & \text{otherwise} \end{cases}$$

be three random variables.

Examples

Value of X	$p_X(x)$
0	$\frac{1}{2}$
1	$\frac{1}{2}$

Value of Y	$p_Y(y)$
1	$\frac{1}{4}$
2	$\frac{1}{4}$
3	$\frac{1}{4}$
4	$\frac{1}{4}$

Value of Z	$p_Z(z)$
0	$\frac{1}{2}$
1	$\frac{1}{2}$

Table: Marginal pmfs of X, Y and Z

Examples

$X \backslash Y$	1	2	3	4	$p_X(x)$
0	0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$
1	$\frac{1}{4}$	0	0	$\frac{1}{4}$	$\frac{1}{2}$
$p_Y(y)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

$X \backslash Z$	0	1	$p_X(x)$
0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
$p_Z(z)$	$\frac{1}{2}$	$\frac{1}{2}$	1

Table: Joint pmfs of X, Y and X, Z

Examples cont.

$Z \backslash Y$	1	2	3	4	$p_Z(z)$
0	$\frac{1}{4}$	0	$\frac{1}{4}$	0	$\frac{1}{2}$
1	0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
$p_Y(y)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

Table: Joint pmf of Z, Y

Expected Value

The expected value of any function $g(X, Y)$ is given by

$$\sum_x \sum_y g(x, y) p_{X,Y}(x, y)$$

For example, if we wanted to compute $E(Y)$ and $E(XY)$, we must employ the following sums:

$$E(Y) = \sum_x \sum_y y p_{X,Y}(x, y)$$

$$E(XY) = \sum_x \sum_y xy p_{X,Y}(x, y)$$

Expected Value

For $E(Y)$,

$X \backslash Y$	1	2	3	4	
0	1×0	$2 \times \frac{1}{4}$	$3 \times \frac{1}{4}$	4×0	
1	$1 \times \frac{1}{4}$	2×0	3×0	$4 \times \frac{1}{4}$	
Total					$\frac{5}{2}$

Expected value (LOTUS)

For $E(XY)$,

$X \backslash Y$	1	2	3	4	
0	0	0	0	0	
1	$\frac{1}{4}$	0	0	$4 \times \frac{1}{4}$	
Total					$\frac{5}{4}$

Table: Finding $E(XY)$

Independence

Definition

Two random variables X and Y are said to be independent if $p_{X,Y}(x,y) = p_X(x)p_Y(y)$, $\forall x, y \in \mathcal{R}$.

In the case of discrete random variables, when two variables are independent, the joint pmf table's rows and columns are multiples of the marginal pmfs of the variables. This is because, if X and Y are independent, then for any constant c, $p_{X,Y}(x,c) = p_X(x)p_Y(c)$. So in the corresponding column for $Y = c$, the probabilities in the table for $p_{X,Y}$ would be a multiple of the column containing the marginal probability mass function p_X.

Independence - Example

Which of the variables X, Y, Z are independent?

Expected value of XY

Let X and Y be two independent random variables. Then,

$$E(XY) = E(X)E(Y)$$

.

Proof: Recall, $E(1_A) = P(A)$. We also note that, $1_A 1_B = 1_{A \cap B}$ and if A, B are independent, $E(1_A 1_B) = E(1_{A \cap B}) = P(A \cap B) = P(A)P(B) = E(1_A)E(1_B)$. So the above property holds for independent indicator random variables. Hence, the only thing left to do is to see that any discrete finite random variable can be written as a linear combination of indicator random variables. This can be done as follows:

Expected value of XY

We will use $1_{X=k}$ to denote the indicator random variable of the event $X = k$. Similarly, we have $1_{Y=l}$. We note that:
$X = \sum_k k 1_{X=k}$ and similarly $Y = \sum_l l 1_{Y=l}$.
Exercise: Please finish the proof.

Examples

$E(X) = \frac{1}{2}$, $E(Y) = \frac{5}{2}$, $E(Z) = \frac{1}{2}$.
Is $E(XY) = E(X)E(Y)$? Are they independent?

Examples

$E(X) = \frac{1}{2}$, $E(Y) = \frac{5}{2}$, $E(Z) = \frac{1}{2}$.
Is $E(XY) = E(X)E(Y)$? Are they independent?
IMPORTANT: if $E(XY) = E(X)E(Y)$ it does not imply that X, Y are independent. The above example is one such example.

Covariance

Definition

Let X, Y be two random variables. Then, the covariance of X and Y, denoted by $Cov(X, Y)$ is defined as $E((X - \mu_X)(Y - \mu_Y))$. Because of Linearity of expectation, this is $E(XY) - \mu_X\mu_Y$.

Covariance

Definition

Let X, Y be two random variables. Then, the covariance of X and Y, denoted by $Cov(X, Y)$ is defined as $E((X-\mu_X)(Y-\mu_Y))$. Because of Linearity of expectation, this is $E(XY)-\mu_X\mu_Y$.

What is the covariance of two independent variables?

Covariance

Definition

Let X, Y be two random variables. Then, the covariance of X and Y, denoted by $Cov(X, Y)$ is defined as $E((X - \mu_X)(Y - \mu_Y))$. Because of Linearity of expectation, this is $E(XY) - \mu_X\mu_Y$.

What is the covariance of two independent variables?
If two random variables X, Y are independent, then
$E(XY) = \mu_X\mu_Y$, so $Cov(X, Y) = 0$.
Is the opposite true? If X, Y are two random variables such that $Cov(X, Y) = 0$, does it imply that X, Y are independent?

Covariance

Definition

Let X, Y be two random variables. Then, the covariance of X and Y, denoted by $Cov(X, Y)$ is defined as $E((X - \mu_X)(Y - \mu_Y))$. Because of Linearity of expectation, this is $E(XY) - \mu_X\mu_Y$.

What is the covariance of two independent variables?
If two random variables X, Y are independent, then
$E(XY) = \mu_X\mu_Y$, so $Cov(X, Y) = 0$.
Is the opposite true? If X, Y are two random variables such that $Cov(X, Y) = 0$, does it imply that X, Y are independent? No.

Covariance - Examples

Covariance of X, Y?

$X \backslash Y$	1	2	3	4	$p_X(x)$
0	0	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$
1	$\frac{1}{4}$	0	0	$\frac{1}{4}$	$\frac{1}{2}$
$p_Y(y)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

$E(X) = \frac{1}{2}$, $E(Y) = \frac{1}{4}(1+2+3+4) = \frac{5}{2}$
$E(XY) = 1 \times \frac{1}{4} + 4 \times \frac{1}{4} = \frac{5}{4}$
So $Cov(X, Y) = 0$.

Covariance Examples cont.

$Z \backslash Y$	1	2	3	4	$p_Z(z)$
0	$\frac{1}{4}$	0	$\frac{1}{4}$	0	$\frac{1}{2}$
1	0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
$p_Y(y)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

$E(Z) = \frac{1}{2}$, $E(Y) = \frac{1}{4}(1+2+3+4) = \frac{5}{2}$
$E(ZY) = 2 \times \frac{1}{4} + 4 \times \frac{1}{4} = \frac{6}{4}$
So $Cov(Z, Y) = \frac{1}{4}$.

Covariance Examples cont.

$X \backslash Z$	0	1	$p_X(x)$
0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
$p_Z(z)$	$\frac{1}{2}$	$\frac{1}{2}$	1

$E(X) = \frac{1}{2}$, $E(Z) = \frac{1}{2}$
$E(XZ) = 1 \times \frac{1}{4} = \frac{1}{4}$
So $Cov(X, Z) = 0$ (as it should be since they are independent).

Properties of Covariance

- (Commutative) $Cov(X, Y) = Cov(Y, X)$.

Properties of Covariance

- (Commutative) $Cov(X, Y) = Cov(Y, X)$.
- If X, Y are independent, $Cov(X, Y) = 0$. (Note: The converse is not true! $Cov(X, Y) = 0$ does not imply that X, Y are independent.

Properties of Covariance

- (Commutative) $Cov(X, Y) = Cov(Y, X)$.
- If X, Y are independent, $Cov(X, Y) = 0$. (Note: The converse is not true! $Cov(X, Y) = 0$ does not imply that X, Y are independent.
- $Cov(aX + b, Y) =$

Properties of Covariance

- (Commutative) $Cov(X, Y) = Cov(Y, X)$.
- If X, Y are independent, $Cov(X, Y) = 0$. (Note: The converse is not true! $Cov(X, Y) = 0$ does not imply that X, Y are independent.
- $Cov(aX + b, Y) =$

Properties of Covariance

- (Commutative) $Cov(X, Y) = Cov(Y, X)$.
- If X, Y are independent, $Cov(X, Y) = 0$. (Note: The converse is not true! $Cov(X, Y) = 0$ does not imply that X, Y are independent.
- $Cov(aX + b, Y) = aCov(X, Y)$
- $Cov(X + Y, Z) =$

Properties of Covariance

- (Commutative) $Cov(X, Y) = Cov(Y, X)$.
- If X, Y are independent, $Cov(X, Y) = 0$. (Note: The converse is not true! $Cov(X, Y) = 0$ does not imply that X, Y are independent.
- $Cov(aX + b, Y) = aCov(X, Y)$
- $Cov(X + Y, Z) =$

Properties of Covariance

- (Commutative) $Cov(X, Y) = Cov(Y, X)$.
- If X, Y are independent, $Cov(X, Y) = 0$. (Note: The converse is not true! $Cov(X, Y) = 0$ does not imply that X, Y are independent.
- $Cov(aX + b, Y) = aCov(X, Y)$
- $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$

Variance: Revisited

Recall that we defined the Variance of a random variable X as $Var(X) = E((X - \mu_X)^2)$. This we can see is $E((X - \mu_X)(X - \mu_X)) = Cov(X, X)$.

Standard deviation

We define the standard deviation of a random variable X as $\sqrt{Var(X)}$, and is denoted by σ_X (pronounced: Sigma X).

Examples

Value of X	$p_X(x)$
0	$\frac{1}{2}$
1	$\frac{1}{2}$

Value of Y	$p_Y(y)$
1	$\frac{1}{4}$
2	$\frac{1}{4}$
3	$\frac{1}{4}$
4	$\frac{1}{4}$

Value of Z	$p_Z(z)$
0	$\frac{1}{2}$
1	$\frac{1}{2}$

Table: Marginal pmfs of X, Y and Z

Examples

$Var(X) = Var(Z) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$
$Var(Y) = \frac{1}{4}(1 + 4 + 9 + 16) - \frac{25}{4} = \frac{5}{4}$
And so, $\sigma_X = \sigma_Z = \frac{1}{2}$, while $\sigma_Y = \frac{\sqrt{5}}{2}$.

Variance Properties

- $Var(X) \geq 0$. Also, if $Var(X) = 0$, then it is 'almost surely' a constant.

Variance Properties

- $Var(X) \geq 0$. Also, if $Var(X) = 0$, then it is 'almost surely' a constant.
- $Var(aX + b) =$

Variance Properties

- $Var(X) \geq 0$. Also, if $Var(X) = 0$, then it is 'almost surely' a constant.
- $Var(aX + b) =$

Variance Properties

- $Var(X) \geq 0$. Also, if $Var(X) = 0$, then it is 'almost surely' a constant.
- $Var(aX + b) = a^2 Var(X)$. Then, $\sigma_{aX+b} = |a|\sigma_X$.
- $Var(X + Y) =$

Variance Properties

- $Var(X) \geq 0$. Also, if $Var(X) = 0$, then it is 'almost surely' a constant.
- $Var(aX + b) = a^2 Var(X)$. Then, $\sigma_{aX+b} = |a|\sigma_X$.
- $Var(X + Y) =$

Variance Properties

- $Var(X) \geq 0$. Also, if $Var(X) = 0$, then it is 'almost surely' a constant.
- $Var(aX + b) = a^2 Var(X)$. Then, $\sigma_{aX+b} = |a|\sigma_X$.
- $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$. **In particular, if X, Y are independent, then** $Var(X + Y) = Var(X) + Var(Y)$.
- $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y)$.

Correlation coefficient

Definition

Given two random variables, X, Y, the correlation coefficient, denoted by $\rho(X, Y)$ (pronounced 'Row' X, Y), is defined as $\frac{Cov(X,Y)}{\sigma_X \sigma_Y}$.

Range of ρ

$-1 \leq \rho(X, Y) \leq 1$.
Proof: Consider $Z_1 = \frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}$ and $Z_2 = \frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}$. Now we note that $Var(Z_i) \geq 0$.

Examples

For the random variables of previous examples, $Cov(Y, Z) = \frac{1}{4}$ and $\rho(X, Y) = \frac{1}{\sqrt{5}}$, since $\sigma_Y = \frac{\sqrt{5}}{2}, \sigma_Z = \frac{1}{2}$.

Random Variable Vectors

All the discussion of multiple random variables is much more succintly expressed as vectors. Again consider X_1, X_2 as two random variables with joint pmf p_{X_1,X_2}. We could instead think of them as a random variable vector, $\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$. The joint pmf will be the same, only denoted by the vector $p_{\boldsymbol{X}}$.

Expectation

$$\mu_{\boldsymbol{X}} = E(\boldsymbol{X}) = E\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} E(X_1) \\ E(X_2) \end{pmatrix}.$$

Random Variable Vectors cont.

Covariance Matrix

The Covariance Matrix for $\boldsymbol{X}$ is denoted by $\boldsymbol{\Sigma}$ and is defined as follows: $\boldsymbol{\Sigma} = Cov(\boldsymbol{X}, \boldsymbol{X}) = E((\boldsymbol{X} - \boldsymbol{\mu_X})(\boldsymbol{X} - \boldsymbol{\mu_X})^T)$

$$= E\begin{pmatrix} (X_1-\mu_1)^2 & (X_1-\mu_1)(X_2-\mu_2) \\ (X_1-\mu_1)(X_2-\mu_2) & (X_2-\mu_2)^2 \end{pmatrix}$$

$$= \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) \\ Cov(X_1, X_2) & Var(X_2) \end{pmatrix}$$

Random Variable Vectors cont.

Properties

Let $\boldsymbol{X}$ be a random variable vector and $\boldsymbol{A}$ a matrix of constants. Then $\Sigma_{\boldsymbol{A}} = Cov(\boldsymbol{AX}, \boldsymbol{AX}) = \boldsymbol{A}Cov(\boldsymbol{X}, \boldsymbol{X})\boldsymbol{A}^T = \boldsymbol{A}\Sigma\boldsymbol{A}^T$. Where Σ is the Covariance matrix of $\boldsymbol{X}$ and $\Sigma_{\boldsymbol{A}}$ is the Covariance matrix of $\boldsymbol{AX}$.

Example: Suppose X and Y are two random variables with $Var(X) = Var(Y) = 2$ and $Cov(X, Y) = 1$. What can we say about $Cov(2X + 3Y, 3X - Y)$?

Random Variable Vectors cont.

Properties

Let $\boldsymbol{X}$ be a random variable vector and $\boldsymbol{A}$ a matrix of constants. Then $\Sigma_{\boldsymbol{A}} = Cov(\boldsymbol{AX}, \boldsymbol{AX}) = \boldsymbol{A}Cov(\boldsymbol{X}, \boldsymbol{X})\boldsymbol{A}^T = \boldsymbol{A}\Sigma\boldsymbol{A}^T$.
Where Σ is the Covariance matrix of $\boldsymbol{X}$ and $\Sigma_{\boldsymbol{A}}$ is the Covariance matrix of $\boldsymbol{AX}$.

Example: Suppose X and Y are two random variables with $Var(X) = Var(Y) = 2$ and $Cov(X, Y) = 1$. What can we say about $Cov(2X + 3Y, 3X - Y)$?

We just note that here if $\boldsymbol{X} = \begin{pmatrix} X \\ Y \end{pmatrix}$, then

$$\begin{pmatrix} 2X + 3Y \\ X - 2Y \end{pmatrix} = \boldsymbol{AX}, \text{ where, } \boldsymbol{A} = \begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix}$$

The End