Probability Theory Lecture 08

Padmini Mukkamala

Budapest University of Technology and Economics

Overview

Definition

Joint Probability Density Function

Random variables X and Y, are said to be jointly continuous, if there exists a non-negative Riemann integrable function $f_{X,Y}(x,y) : \mathcal{R}^2 \to \mathcal{R}$, such that $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$.
Here, $f_{X,Y}(x,y)$ is said to be the joint probability density function.

Joint Cumulative Distribution Function

For a set A, $P((x,y) \in A) = \int \int_A f_{X,Y}(x,y)dxdy$. In particular, the joint cumulative distribution function is as follows:

$$F_{X,Y}(a,b) = \int_{-\infty}^{b} \int_{-\infty}^{a} f_{X,Y}(x,y)dxdy$$

Example

The following is Example 1c from the textbook.

Given $f_{X,Y}(x, y) = \begin{cases} 2e^{-x}e^{-2y} & \text{if } 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$

What is $P(X > 1, Y < 1)$?

Example

The following is Example 1c from the textbook.

Given $f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y} & \text{if } 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$

What is $P(X > 1, Y < 1)$?

What is $P(X < Y)$?

Example

The following is Example 1c from the textbook.

Given $f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y} & \text{if } 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$

What is $P(X > 1, Y < 1)$?

What is $P(X < Y)$?

What is $P(X < a)$?

Example

The following is Example 1c from the textbook.

Given $f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y} & \text{if } 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$

What is $P(X > 1, Y < 1)$?

What is $P(X < Y)$?

What is $P(X < a)$?

What is $P(X/Y < 1)$?

Marginal PDFs

Given two random variables X, Y and their joint PDF $f_{X,Y}(x,y)$, the marginal PDFs are as follows:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$

and

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx$$

Many jointly continuous variables

Joint Probability Density Function

Random variables $X_1, X_2, ...X_n$ are said to be jointly continuous, if there exists a non-negative Riemann integrable function $f_{X_1,X_2,...X_n}(x_1, x_2, ..., x_n) : \mathcal{R}^n \to \mathcal{R}$, such that

$$\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f_{X_1,X_2,..X_n}(x_1, x_2, ..., x_3) dx_1 dx_2 ... dx_n = 1$$

The above function is called the joint probability density function.

Many jointly continuous variables

Joint Cumulative Distribution Function

For a set A,
$P((x_1, x_2, .., x_n) \in A) = \int \dots \int_A f_{X_1, X_2, ..X_n}(x_1, x_2, .., x_n) dx_1 dx_2 \dots dx_n$.
In particular, the joint cumulative distribution function is as follows:

$$F_{X_1, X_2, .., X_n}(a_1, a_2, .., a_n) = \int_{-\infty}^{a_n} .. \int_{-\infty}^{a_1} f_{X_1, X_2, .., X_n}(x_1, x_2, .., x_n) dx_1 dx_2 .. dx_n$$

Independence

Intuitively, we expect that if two random variables X, Y are independent, then for any events A, B, $P(X \in A \text{ and } Y \in B) = P(X \in A)P(Y \in B)$. It can be shown that it is enough to check this for events A and B of the kind $(-\infty, a]$ and $(-\infty, b]$. Notice that in this case, $P(X \in A) = F_X(a)$ and $P(X \in B) = F_Y(b)$.
In other words, two random variables are independent if

$$F_{X,Y}(a, b) = F_X(a)F_Y(b), \qquad \forall a, b \in \mathcal{R}$$

Independence

Without proof, we will further note that this is true when the following holds:

PDF definition of Independence

Two random variables X and Y are said to be independent if

$$f_{X,Y}(x, y) = f_X(x)f_Y(y), \qquad \forall x, y \in \mathcal{R}$$

Examples

Buffon's Needle

Consider parallel lines spaced $2L$ apart. We drop a matchstick of length L. What is the probability that the matchstick intersects the lines?

Note, here if we take the distance of the center of the matchstick from the closest line, and the angle it makes, they determine the matchstick completely, and also that these two parameters are independent.

Then the density function should be $f_{X,\theta}(x,\theta) = \frac{1}{\pi L}$.

Examples cont.

Example 2f from book: Let

$$f_{X,Y}(x,y) = \begin{cases} 24xy & \text{if } 0 < x < 1, 0 < y < 1, 0 < x+y < 1 \\ 0 & \text{otherwise} \end{cases}.$$

Are the random variables X, Y independent?

Examples cont.

Example 2f from book: Let

$$f_{X,Y}(x,y) = \begin{cases} 24xy & \text{if } 0 < x < 1, 0 < y < 1, 0 < x + y < 1 \\ 0 & \text{otherwise} \end{cases}.$$

Are the random variables X, Y independent?

Consider X, Y to have uniform distribution over a unit square. Are they independent?

Examples cont.

Example 2f from book: Let

$$f_{X,Y}(x,y) = \begin{cases} 24xy & \text{if } 0 < x < 1, 0 < y < 1, 0 < x + y < 1 \\ 0 & \text{otherwise} \end{cases}.$$

Are the random variables X, Y independent?

Consider X, Y to have uniform distribution over a unit square. Are they independent?

What if they had uniform distribution over a unit circle? Are they still independent?

Independence

Another way to think of independence, then, is:

Multiplication definition of Independence

Two random variables X and Y are said to be independent if there exist functions $g(x)$ and $h(y)$, such that,

$$f_{X,Y}(x,y) = g(x)h(y), \qquad -\infty < x < \infty, -\infty < y < \infty$$

Expected Value

Definition

Let X, Y be two random variables. Then, for any function $g(X, Y)$, the expected value is defined as

$$E(g(X, Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y)dxdy$$

provided that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |g(x, y)| f_{X,Y}(x, y)dxdy$ is defined.

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 $E(XY) - \mu_X \mu_Y$.

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.
- $Cov(aX + b, Y) = aCov(X, Y)$

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.
- $Cov(aX + b, Y) = aCov(X, Y)$
- $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$

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).

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.
- $Var(aX + b) = a^2 Var(X)$. Then, $\sigma_{aX+b} = |a|\sigma_X$.

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 $Var(X + Y) = Var(X) + Var(Y)$.

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 $Var(X + Y) = Var(X) + Var(Y)$.
- $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y)$.

Linear Algebra

Expectation

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

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 $\boldsymbol{AX}$.

Convolution

Definition

Let X, Y be two **independent** random variables. Then, the **convolution** of X and Y is the random variable $Z = X + Y$.

Convolution

If X and Y are discrete and independent random variables, let $Z = X + Y$. If we try to find the probability mass function of Z, then,

$P(Z = z) = P(X + Y = z)$

$= \sum_y P(X = z - y \text{ and } Y = y)$

Using Independence, this is

$= \sum_y P(X = z - y)P(Y = y)$

Discrete Case

The PMF of the convolution of two discrete independent random variables is given by

$$p_Z(z) = \sum_y p_X(z - y)p_Y(y)$$

Convolution

If X and Y are continuous and independent random variables, let $Z = X + Y$. If we try to find the cumulative distribution function of Z, then,

$F_Z(z) = P(Z \leq z) = P(X + Y \leq z)$

$= \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_{X,Y}(x, y) dxdy = \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_X(x) f_Y(y) dxdy$

$= \int_{-\infty}^{\infty} \left(f_Y(y) \int_{-\infty}^{z-y} f_X(x) dx\right) dy = \int_{-\infty}^{\infty} f_Y(y) F_X(z - y) dy$

We differentiate with respect to z to get, $f_Z(z) = \frac{d}{dz} F_Z(z)$

$= \frac{d}{dz} \int_{-\infty}^{\infty} f_Y(y) F_X(z - y) dy = \int_{-\infty}^{\infty} f_Y(y) f_X(z - y) dy$

Convolution

Continuous Case

The PDF and CDF for the convolution $Z = X + Y$ of two independent continuous random variables X, Y are as follows:

$$F_Z(z) = \int_{-\infty}^{\infty} F_X(z-y) f_Y(y) dy$$

and

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy$$

Convolution

Using Moment Generating Functions

Let X, Y be independent random variables. Then we know that the MGF of the convolution is given by $M_X(t)M_Y(t)$. Sometimes, this gives us a one step answer to finding the convolution, in particular, if this MGF has a recognizable 'standard' form.

Examples of Convolution

Let $X, Y \sim U(0,1)$. Then $Z = X + Y$ has range $[0,2]$. Its PDF is given by,

Examples of Convolution

Let $X, Y \sim U(0,1)$. Then $Z = X + Y$ has range $[0,2]$. Its PDF is given by,

$$f_Z(z) = \int_{\max(0,z-1)}^{\min(z,1)} dxdy$$

This gives us two integrals,
$f_Z(z) = \int_0^z dy$, when $0 \leq z < 1$
and
$f_Z(z) = \int_{z-1}^1 dy$, when $1 \leq z \leq 2$

Examples of Convolution

Uniform Distributions

Convolution of two $U(0,1)$ is:

$$f_Z(z) = \begin{cases} z & 0 \leq z < 1 \\ 2 - z & 1 \leq z \leq 2 \end{cases}$$

Examples of Convolution

Uniform Distributions

Convolution of two $U(0,1)$ is:

$$f_Z(z) = \begin{cases} z & 0 \leq z < 1 \\ 2 - z & 1 \leq z \leq 2 \end{cases}$$

Note: MGF is not particularly useful in this case, as the product of MGFs is not an easy recognizable MGF.

Examples of Convolution cont.

Bernoulli distributions

Let $X, Y \sim 1(p)$ be independent and let $Z = X + Y$. Then, the PMF of Z is given by,

Examples of Convolution cont.

Bernoulli distributions

Let $X, Y \sim 1(p)$ be independent and let $Z = X + Y$. Then, the PMF of Z is given by,

$$p_Z(z) = \begin{cases} (1-p)^2 & z = 0 \\ 2(1-p)p & z = 1 \\ p^2 & z = 2 \end{cases}$$

Note: MGF was the faster way here!

Examples of Convolution cont.

Binomial distributions

Let $X \sim Bin(n, p)$ and $Y \sim Bin(m, p)$ be two independent random variables. Then their convolution is

Examples of Convolution cont.

Binomial distributions

Let $X \sim Bin(n, p)$ and $Y \sim Bin(m, p)$ be two independent random variables. Then their convolution is $Bin(n + m, p)$

Note: Again MGFs are much faster here!

Examples of Convolution cont.

Poisson distributions

Let $X \sim Pois(\lambda_1)$ and $Y \sim Pois(\lambda_2)$ be two independent random variables. Then their convolution is

Examples of Convolution cont.

Poisson distributions

Let $X \sim Pois(\lambda_1)$ and $Y \sim Pois(\lambda_2)$ be two independent random variables. Then their convolution is $Pois(\lambda_1 + \lambda_2)$

Note: Again MGFs are much faster here!

Examples of Convolution cont.

Normal distributions

Let $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ be two independent random variables. Then their convolution is

Examples of Convolution cont.

Normal distributions

Let $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ be two independent random variables. Then their convolution is $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Note: Again MGFs are much faster here! Infact, in this particular case, the integrals are formidable if we use the direct definition of convolution.

The End