CE 211
Mathematics for Engineers

Lecture 7

Functions of Random Variables

Previously on Mathematics for Engineers

Definition (Probability Density Function)

The probability density function (PDF) of a continuous random variable is denoted as $f_X(x)$ and is defined as

$$f_X(x)dx = \mathbb{P}\big(X \in [x, x + dx]\big)$$

Thus, the probability that the random variable lies in a subset S is given by

$$\mathbb{P}(X \in S) = \int_{x \in S} f_X(x)dx$$

Definition (Cumulative Distribution Function)

The cumulative distribution function (CDF) of a random variable X is denoted by $F_X(x)$ and is defined as

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(x)dx$$

Previously on Mathematics for Engineers

Definition (Expectation)

The expected value of a random variable X is denoted by $\mathbb{E}(X)$ or μ_X and is defined as

$$\mathbb{E}(X) = \int_{x \in R_X} x f_X(x) dx$$

Definition (Variance)

The variance of a random variable X is denoted by $V(X)$, $\text{Var}(X)$, or σ_X^2 and is defined as

$$V(X) = \mathbb{E}\left((X - \mu_X)^2\right) = \int_{x \in R_X} (x - \mu_X)^2 f_X(x) dx$$

Previously on Mathematics for Engineers

Definition

If X and Y are two random variables, the joint PDF is given by

$$f_{X,Y}(x, y)dydx = \mathbb{P}(X \in [x, x + dx], Y \in [y, y + dy])$$

The probability that $(X, Y) \in S$ is thus given by

$$\int \int_{(x,y)\in S} f(x, y)dydx$$

Definition

The CDF of a jointly distributed continuous random variable is

$$F_{X,Y}(x, y) = \mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)dydx$$

Previously on Mathematics for Engineers

We can also define **marginal densities** of X and Y by replacing sums with integrals. The marginal density of X is

$$\mathbb{P}(X \in [x, x + dx]) = \int_y f(x, y)dy$$

The marginal density of Y is

$$\mathbb{P}(Y \in [y, y + dy]) = \int_x f(x, y)dx$$

Definition (Expectation)

Suppose X and Y are random variables with joint PDF $f_{X,Y}(x, y)$. Then, the expectation of $g(X, Y)$ is defined as

$$\mathbb{E}(g(X, Y)) = \int_{x \in R_X} \int_{y \in R_Y} g(x, y) f_{X,Y}(x, y) dxdy$$

Previously on Mathematics for Engineers

The idea of conditional probability of events can be extended similarly to random variables to define conditional random variables and their PMFs and PDFs. In the continuous setting, we can write

$$f_{X|Y}(x|y) = \frac{\mathbb{P}(X \in [x, x+dx], Y \in [y, y+dy])}{\mathbb{P}(Y \in [y, y+dy])} = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Definition

Two random variables X and Y are said to be independent if $p_{X,Y}(x,y) = p_X(x)p_Y(y)$ or $f_{X,Y}(x,y) = f_X(x)f_Y(y)$

In other words, knowing the value of realization of one of the random variables does not affect the density of the other random variable and vice versa.

We can also define independence using CDFs, i.e., X and Y are independent if

$$F_{X,Y}(x,y) = F_X(x)F_Y(y)$$

Previously on Mathematics for Engineers

Definition (Covariance)

The covariance of two random variables X and Y is denoted by $\mathrm{Cov}(X, Y)$ and is defined as

$$\mathrm{Cov}(X, Y) = \mathbb{E}\Big((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\Big)$$

What happens when $X = Y$? The above equation is a bit unwieldy but there is an easier way to compute the covariance.

$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Definition (Correlation)

The correlation coefficient of two random variables X and Y is denoted using $\rho_{X,Y}$ and is defined as

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Lecture Outline

The Problem

The Problem

Envelope Paradox

You are to choose between two identical sealed envelopes containing money. One of them has twice the amount in the other.

After choosing an envelope, you are given the option of switching to the other envelope. Should you switch?

The Problem

Suppose, the amount in the envelope you selected is A. Then, the other envelope either contains $A/2$ or $2A$ with equal probability.

Hence, the expected value of the amount in the other envelope is

$$\frac{1}{2}A/2 + \frac{1}{2}2A = \frac{5}{4}A$$

Therefore, you must switch. But the same argument can be applied again! What's wrong with this answer?

Sum of Independent Random Variables

In several problems, we are interested in sums of independent random variables. For example,

- Suppose the number of men and women who arrive at a polling booth are independently Poisson distributed with rates λ_1 and λ_2. What is the distribution of the number of the total number of voter arrivals.
- The return from two stocks is independently normally distributed with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2. What is the distribution of the total return?

Sum of Independent Random Variables

Binomial Distributed Random Variables

Suppose $X \sim \text{Bin}(n, p)$ and $Y \sim \text{Bin}(m, p)$. Then, the random variable $Z = X + Y$ also has a Binomial distribution with parameters $n + m$ and p.

This is expected since the trails in the Binomial distribution are all independent. One can also show this mathematically using the techniques discussed in the next few slides.

Sum of Independent Random Variables

Poisson Random Variables

Suppose $X \sim \text{Pois}(\lambda_1)$ and $Y \sim \text{Pois}(\lambda_2)$. What is the distribution of $X + Y$?
The random variable $Z = X + Y$ is Poisson distributed with parameter $\lambda_1 + \lambda_2$.

$$
\begin{aligned}
\mathbb{P}(Z = z) &= \mathbb{P}(X + Y = z) \\
&= \sum_{k=0}^{z} \mathbb{P}(X = k, Y = z - k) \\
&= \sum_{k=0}^{z} \mathbb{P}(X = k)\mathbb{P}(Y = z - k) \\
&= \sum_{k=0}^{z} \frac{e^{-\lambda_1}\lambda_1^k}{k!}\frac{e^{-\lambda_2}\lambda_2^{z-k}}{(z-k)!} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{z} \frac{\lambda_1^k}{k!}\frac{\lambda_2^{z-k}}{(z-k)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{z!} \sum_{k=0}^{z} \frac{z!}{(z-k)!k!}\lambda_1^k\lambda_2^{z-k} = \frac{e^{-(\lambda_1+\lambda_2)}}{z!}(\lambda_1+\lambda_2)^z
\end{aligned}
$$

Sum of Independent Random Variables

Uniform Random Variables

Suppose $X \sim U(0,1)$ and $Y \sim U(0,1)$. Define $Z = X + Y$. All three random variables X, Y, and Z are continuous random variables.

What is the support of Z? The most convenient way to derive the density and distribution of Z is to work with its CDF.

$$
\begin{aligned}
F_Z(z) = \mathbb{P}(Z \le z) &= \mathbb{P}(X + Y \le z) \\
&= \int_{-\infty}^{\infty} \mathbb{P}(X \le z - y)\mathbb{P}\big(Y \in [y, y + dy]\big)\, dy \\
&= \int_{-\infty}^{\infty} F_X(z - y) f_Y(y) dy
\end{aligned}
$$

We can also derive the same result in a slightly different manner using joint densities and independence.

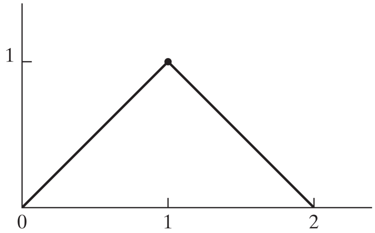
$$
\begin{aligned}
F_Z(z) &= \mathbb{P}(X + Y \leq z) \\
&= \int\int_{x+y\leq z} f_{X,Y}(x,y)dxdy \\
&= \int\int_{x+y\leq z} f_X(x)f_Y(y)dxdy \\
&= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{z-y} f_X(x)dx\right)f_Y(y)dy = \int_{-\infty}^{\infty} F_X(z-y)f_Y(y)dy
\end{aligned}
$$

The CDFs can be differentiated to get the PDFs of the sums.

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

Use this expression to derive the PDF of the sum of two uniforms between 0 and 1. If $z \in [0,1]$, $f_Z(z) = \int_0^z dy$ and if $z \in [1,2]$, $f_Z(z) = \int_{z-1}^1 dy$. Thus,

$$f_Z(z) = \begin{cases} z & \text{if } 0 \le z \le 1 \\ 2-z & \text{if } 1 < z \le 2 \\ 0 & \text{otherwise} \end{cases}$$



Suppose X and Y are two standard normal random variables. What is the distribution of $Z = X + Y$? Using the earlier formula,

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-y)^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy \end{aligned}$$

which when simplified leads to

$$\frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}(z)^2}$$

which is the PDF of a normal random variable with mean 0 and variance 2.

This result can be generalized to more than two random variables with different means and variances.

If $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, for $i = 1, \ldots, n$, then

$$\sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

Sum of Independent Random Variables

Convolutions

The method described so far is a general approach to finding the distribution of sum(s) of random variables.

The distribution function of the sum Z, $F_Z(z)$ is said to be the **convolution** of the distribution functions of X and Y.

However, it is not necessary that the sum of random variables have the same distribution or even a known distribution.

Lecture Outline

Other Functions

Other Functions

Introduction

Several situations may require us to derive the PDF and CDF of other functions of random variables. For example,

- The arrival time of two travellers at a bus stop is independent and uniformly distributed in $[0, 5]$. What is the distribution of the duration of their overlap?
- The lifetime of two light bulbs is independent and exponentially distributed with means λ_1 and λ_2. What is the PDF function of the duration of a series connection of these light bulbs.

Independence allows us to simplify calculations but if not assumed, we can derive the required distributions using joint PMFs/PDFs.

Consider the exponential random variable example. Let $X_1 \sim \exp(\lambda_1)$ and $X_2 \sim \exp(\lambda_2)$ indicate the lifetime of the two bulbs. Here we are interested in the PDF of the random variable $X = \min\{X_1, X_2\}$.

Again, just as done in the case of convolutions, it is easier to work with the CDF.

$$\begin{aligned} 1 - F_X(x) &= \mathbb{P}(X \geq x) \\ &= \mathbb{P}(X_1 \geq x, X_2 \geq x) \\ &= \mathbb{P}(X_1 \geq x)\mathbb{P}(X_2 \geq x) \\ &= e^{-\lambda_1 x} e^{-\lambda_2 x} \\ &= e^{-(\lambda_1+\lambda_2)x} \end{aligned}$$

Hence, X is exponentially distributed with a rate $\lambda_1 + \lambda_2$.

Other Functions

Order Statistics

Imagine a function of several random variables, representing a specific transformation such as increasing or decreasing order.

Specifically, let $X_1, X_2, \ldots, X_n$ be continuous, independent and identically distributed (iid) random variables with a density function f.

Define a new vector of random variables $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ such that $X_{(j)}$ indicates the jth smallest element. That is $X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$. $X_{(2)}$ is the second smallest element and so on.

These new random variables are called order statistics and note that they always satisfy $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$. What is the PDF and CDF of the 1st and nth order statistic? jth order statistic?

Other Functions

Order Statistics

Is the jth order statistic a discrete or a continuous random variable? What is its support?

The CDF of $X_{(j)}$, $F_{X_{(j)}}(x)$ is the probability with which $X_{(j)}$ is less than or equal to x. For this to happen, at least j random variables must be less than or equal to x and the remaining $n - j$ random variables must be greater than x.

$$F_{X_{(j)}}(x) = \sum_{k=j}^{n} \binom{n}{k} F(x)^k (1 - F(x))^{n-k}$$

Differentiating this expression, we can show that the PDF of the jth order statistic is

$$f_{X_{(j)}}(x) = \frac{n!}{(n-j)!(j-1)!} F(x)^{j-1} (1 - F(x))^{n-j} f(x)$$

Other Functions

Order Statistics

The PDF can also be derived from the following argument. For $X_{(j)}$ to belong to a dx around x, $j-1$ random variables should be less than or equal to x, $n-j$ should be greater than x, and exactly one random variable should be in $[x, x+dx]$.

This occurs with probability $F(x)^{j-1}(1-F(x))^{n-j}f(x)$. But there are several ways of selecting three groups of sizes $j-1$, $n-j$, and 1. Hence, the PDF is

$$\begin{aligned} f_{X_{(j)}}(x) &= \binom{n}{j-1, n-j, 1} F(x)^{j-1}(1-F(x))^{n-j}f(x) \\ &= \frac{n!}{(n-j)!(j-1)!} F(x)^{j-1}(1-F(x))^{n-j}f(x) \end{aligned}$$

Order statistics have several applications. For example, in auctions, they can be used to determine how likely one can win with a bid and the expected amount they might have to pay.

Other Functions

General Functions of Univariate Random Variables

In addition to specific functions, we can also find the PDF of generic functions of random variable under some conditions.

We have already discussed what happens to expectations and variances under such transformations but did not discuss the PDF.

We will only look at the continuous univariate case but the extensions to multivariate joint distributions exist and are similar.

We have already encountered this situation in a couple of instances. (Where?)

As an example, let $X \sim U(0,1)$. What is the PDF of $Y = X^n$? What is the support of Y? Again, we use the CDF to get to the PDF.

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) \\ &= \mathbb{P}(X^n \leq y) \\ &= \mathbb{P}(X \leq y^{1/n}) \\ &= F_X(y^{1/n}) \\ &= y^{1/n} \end{aligned}$$

Hence, the PDF of Y is

$$f_Y(y) = \begin{cases} \frac{1}{n} y^{1/n-1} & \text{if } 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Other Functions

General Functions of Univariate Random Variables

Claim

Let X be a continuous random variable and $g(.)$ be a strictly monotonic and differentiable function. The PDF of the random variable $Y = g(X)$ is given by

$$f_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right) \left|\frac{d}{dy} g^{-1}(y)\right| & \text{if } y = g(x) \text{ for some } x \\ 0 & \text{otherwise} \end{cases}$$

Proof.

Suppose g is strictly increasing. If $y = g(x)$ for some x,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(g(X) \leq y) \\ &= \mathbb{P}(X \leq g^{-1}(y)) \\ &= F_X(g^{-1}(y)) \end{aligned}$$

Differentiating this, we get

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \frac{d}{dy} g^{-1}(y)$$

If there is no x for which $y = g(x)$, then $F_X(x)$ is either 0 or 1 and hence its derivative is 0. The case of strictly decreasing g is left as an exercise. ■

Lecture Outline

Expectation and Variance

Expectation and Variance

Introduction

When we deal with functions of random variables, expectation can be calculated using the joint density as done in the last lecture. We can also calculate covariance terms from the definition.

However, for sums of random variables, the expressions for expectation and covariances are simpler.

Expectation and Variance

Expectation of Sums

Claim

Suppose $X_1, \ldots, X_n$ denotes a vector of random variables such that $|\mathbb{E}(X_i)| < \infty$. Then,

$$\mathbb{E}(X_1 + \ldots + X_n) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n)$$

Proof.

The following is a proof for two random variables. It can be extended to the general case using induction.

$$\begin{aligned}
\mathbb{E}(X_1 + X_2) &= \int\int (x_1 + x_2) f(x_1, x_2) dx_1 dx_2 \\
&= \int\int x_1 f(x_1, x_2) dx_2 dx_1 + \int\int x_2 f(x_1, x_2) dx_1 dx_2 \\
&= \int x_1 f(x_1) dx_1 + \int x_2 f(x_2) dx_2 \\
&= \mathbb{E}(X_1) + \mathbb{E}(X_2)
\end{aligned}$$

■

Note that for the previous result we **do not need** the random variable to

- Be independent
- Have the same distribution

Using this result, assuming a success probability of p, find the expected value of a

- Binomial random variable for n trials
- Negative binomial random variable for r successes

Claim

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

Proof.

Exercise. Use the definition of Covariance

$$\text{Cov}(X, Y) = \mathbb{E}\Big((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\Big)$$

■

Expectation and Variance

Variance of Sums

Thus, the variance of a sum of random variables can be written as

$$\begin{aligned} V\left(\sum_{i=1}^{n} X_i\right) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} V(X_i) + \sum\sum_{i \neq j} \mathrm{Cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} V(X_i) + 2\sum\sum_{i<j} \mathrm{Cov}(X_i, X_j) \end{aligned}$$

Corollary

If X_is are independent, then $V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i)$

A Solution

A Solution

Two Envelope Problem

There are a few ways to resolve this problem. First, is to use the gains and losses instead of the absolute values. Suppose your envelope has A.

Swapping can give you a gain of A or $-A$ with equal probability and hence the expected gain is 0.

A Solution

Two Envelope Problem

The problem with the earlier expectation calculation is that it is not conditioned. Since we are comparing A with the expected value in the other envelope, A must refer to the expected amount in envelope 1 and not the absolute amount.

The right way to find the expectation is that, the expected value in the other envelope (say envelope 2) is

Expected value in 2 $= 1/2$ (expected value in 2 $|$ 1 has more amount than 2) $+ 1/2$ (expected value in 2 $|$ 2 has more amount than 1)

If x and $2x$ were the amounts in both envelopes, we get Expected value in 2 $= 1/2(x + 2x)$, which is also the expected value in 1 (=A).

Your Moment of Zen