CE 211
Mathematics for Engineers

Lecture 8

Beyond Expectations

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

The idea of conditional probability of events can be extended similarly to random variables to define conditional random variables and their PMFs and PDFs.

For example, in the discrete setting, we can define the PMF of a conditional random variable as $p_{X|Y}(x|y)$ as

$$p_{X|Y}(x|y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}$$

That is, we are conditioning X with events associated with $Y = y$.

Likewise, in the continuous setting, we can write

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

Lecture Outline

Lecture Outline

The Problem

The Problem

Necktie Paradox

Two men (say A and B) are each gifted a necktie by their wives. They do not know how much each of their neckties cost. They decide to bet on who has the cheaper necktie.

They agree to consult their wives and find out the cost of the neckties. If say A has the more expensive neck tie, he has to give it to B and vice versa.

Each of them reason as follows: One could win or loose with equal probability. If I loose, I loose an amount equal to the cost of my necktie. If I win, I will get more than the cost of my necktie.

In expectation, I gain to participate in the bet. But then, by the same logic, the other man also gains from the bet!

Lecture Outline

Conditional Expectation

For the conditionally distributed random variables discussed in previous class, we can extend the idea of expectation. For the discrete case,

$$\mathbb{E}(X|Y=y) = \sum x\mathbb{P}(X=x|Y=y) = \sum x p_{X|Y}(x|y)$$

For the continuous case,

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

Conditional Expectation

Example

Imagine a biased coin whose probability of heads is not known and assumed to be a random variable with uniform distribution on $[0, 1]$. Suppose, the coin is tossed n times and let X be the number of heads observed. What is the expected number of heads?

Let Y be the random variable that indicates the probability of heads. Given a realization of Y, we can calculate the expected number of heads using Binomial distribution.

$$\mathbb{E}(X|Y = y) = ny$$

Thus, **conditional expectation is a function of the random variable** Y. Since, functions of random variables are random variables, $\mathbb{E}(X|Y)$ is a random variable! In this example, $\mathbb{E}(X|Y) = nY$. So, $\mathbb{E}_Y(\mathbb{E}(X|Y)) = n\mathbb{E}_Y(Y) = n/2$

Conditional Expectation

Random Variable

Think of the symbol $\mathbb{E}(X|Y)$ as $g(Y)$ and $\mathbb{E}(X|Y=y)$ as $g(y)$ and recall that $\mathbb{E}(g(Y)) = \sum_y g(y)p_Y(y)$ or $\mathbb{E}(g(Y)) = \int_y g(y)f_Y(y)dy$.

Thus, the expectation of conditional expectation for the discrete case is

$$\mathbb{E}(\mathbb{E}(X|Y)) = \sum_y \mathbb{E}(X|Y=y)p_Y(y)$$

One can write $\mathbb{E}_Y(\mathbb{E}(X|Y))$ instead of $\mathbb{E}(\mathbb{E}(X|Y))$ for additional emphasis.

For the continuous case,

$$\mathbb{E}(\mathbb{E}(X|Y)) = \int_y \mathbb{E}(X|Y=y)f_Y(y)dy$$

Note: For the continuous case, $\mathbb{E}(X|Y=y)$ can be interpreted as $\mathbb{E}(X|Y \in [y, y+dy])$.

Conditional Expectation

Law of Iterated Expectations

Claim

$$\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X)$$

Proof.

For discrete random variables,

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(X|Y)) &= \sum_y \mathbb{E}(X|Y=y)p_Y(y) \\
&= \sum_y \sum_x x\mathbb{P}(X=x|Y=y)\mathbb{P}(Y=y) \\
&= \sum_y \sum_x x\mathbb{P}(X=x, Y=y) \\
&= \sum_x \sum_y x\mathbb{P}(X=x, Y=y) \\
&= \sum_x x \sum_y \mathbb{P}(X=x, Y=y) = \sum_x x p_X(x)
\end{aligned}$$

■

Conditional Expectation

Law of Iterated Expectations

This result is called the law of iterated expectations, law of total expectation, or the tower rule.

The event-version of the tower rule can also be defined. If $A_1, A_2, \ldots, A_n$ is a partition of the sample space, then

$$\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X|A_i)\mathbb{P}(A_i)$$

Moment Generating Functions

Moment Generating Functions

Introduction

Moment generating functions is an alternate way of characterizing distributions of random variables instead of the PDF and CDFs.

Several results in probability can be derived relatively easily using generating functions, particularly those involving expected values and sums of random variables.

Moment Generating Functions

Introduction

Definition

The moment generating function (MGF) of a random variable X is denoted by $M_X(t)$ and is defined as

$$M_X(t) = \mathbb{E}(e^{tX})$$

By definition of expectation, we can also write the moment generating function as

$$M_X(t) = \sum e^{tx} p_X(x)$$

or for the continuous case

$$M_X(t) = \int e^{tx} f_X(x) dx$$

Moment Generating Functions

Introduction

For a random variable X, the nth order moment is defined as $\mathbb{E}(X^n)$. Hence, the mean is the first-order moment.

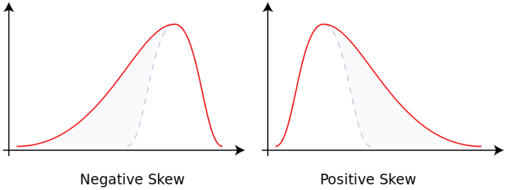
The moment generating function gets its name from a property which allows us to derive the higher order moments of random variables by differentiating it multiple times.

Moment Generating Functions

Skewness and Kurtosis

Standardized higher-order moments have some special features. For instance, the Skewness of a random variable is defined as

$$\gamma = \mathbb{E}\left(\left(\frac{X-\mu}{\sigma}\right)^3\right)$$



The sign of skewness tells us if it is left-skewed (see left figure) or right-skewed (see right figure). What can you say about the skewness of normal, exponential, lognormal distributions?

The standardized-fourth moment is called Kurtosis and reflects the shape of the tails of the PDF of a random variable.

$$k = \mathbb{E}\left(\left(\frac{X-\mu}{\sigma}\right)^4\right)$$

Moment Generating Functions

Derivatives

Using the expansion of exponential functions,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

and linearity of expectation operator,

$$M_X(t) = \mathbb{E}(e^{tX}) = 1 + t\mathbb{E}(X) + \frac{t^2\mathbb{E}(X^2)}{2!} + \dots$$

Using the above expansion, find $\frac{d^3}{dt^3} M_X(t)$ at $t = 0$.

Generalizing this, the nth derivative of the moment generating function evaluated at $t = 0$ gives the nth-order moment.

Moment Generating Functions

Common Distributions

Derive the moment generating functions for the following distribution:

- Poisson distribution $e^{\lambda(e^t-1)}$
- Uniform distribution $\frac{e^{bt}-e^{at}}{t(b-a)}$
- Standard normal distribution $e^{t^2/2}$
- Normal distribution $e^{\mu t+\sigma^2 t^2/2}$

Compute the first and second derivatives of MGFs.

Consider the problem of adding the random variables that we discussed in the last class.

Suppose X and Y are two independent random variables with moment generating functions $M_X(t)$ and $M_Y(t)$. What is the moment generating function of $Z = X + Y$?

$$
\begin{aligned}
M_Z(t) = \mathbb{E}\big(e^{t(X+Y)}\big) &= \int\int e^{t(x+y)} f(x,y)dydx \\
&= \int\int e^{tx} f(x)e^{ty} f(y)dydx \\
&= \int e^{tx} f(x)dx \int e^{ty} f(y)dy \\
&= \mathbb{E}\big(e^{tX}\big)\mathbb{E}\big(e^{tY}\big) = M_X(t)M_Y(t)
\end{aligned}
$$

Moment Generating Functions

Sums of Independent Random Variables

We can thus use this result to identify if the resulting random variable has the same distribution. Check the results from last class for

- Poisson distribution
- Uniform distribution
- Standard normal distribution

Note that just as expectation, the MGF may not always exist.

Probability Inequalities

Many applications and proofs require finding bounds on the probabilities. Let us discuss a few popular inequalities.

Claim (Markov's Inequality)

Let X be non-negative and assume $\mathbb{E}(X)$ is finite. For any $t > 0$,

$$\mathbb{P}(X > t) \leq \frac{\mathbb{E}(X)}{t}$$

Proof.

$$\begin{aligned}
\mathbb{E}(X) &= \int_0^\infty xf(x)dx = \int_0^t xf(x)dx + \int_t^\infty xf(x)dx \\
&\geq \int_t^\infty xf(x)dx \geq \int_t^\infty tf(x)dx \\
&= t\int_t^\infty f(x)dx = t\mathbb{P}(X \geq t)
\end{aligned}$$

∎

Probability Inequalities

Chebyshevs inequality

Claim (Chebyshev's Inequality)

Let X be a random variable with finite μ and σ^2. For any $t > 0$,

$$\mathbb{P}\Big(|X - \mu| \geq t\Big) \leq \frac{\sigma^2}{t^2}$$

Proof.

Using Markov' inequality

$$\begin{aligned} \mathbb{P}\Big(|X - \mu| \geq t\Big) &= \mathbb{P}\Big(|X - \mu|^2 \geq t^2\Big) \\ &\leq \frac{\mathbb{E}\left((X - \mu)^2\right)}{t^2} = \frac{\sigma^2}{t^2} \end{aligned}$$

■

In addition, if $Z = \frac{X-\mu}{\sigma}$,

$$\mathbb{P}\big(|Z| \geq t\big) \leq \frac{1}{t^2}$$

Lecture Outline

A Solution

A Solution

Necktie Paradox

From mindyourdecisions.com,

Resolution 3: never bet a gift from your wife

Of course, it should be clear the game is not really a zero-sum game but a negative sum game.

Upon learning the bet, and that their husbands would wager their thoughtful gifts, both wives will be angry. Clearly, there will be no winners, and the only safe bet is to avoid this game entirely. Consider yourself forewarned.

Your Moment of Zen