Math 461 Fall 2021

Renming Song

University of Illinois at Urbana-Champaign

September 15, 2021

Outline

Outline

HW3 is due Friday, 09/17, before the end of class. You can either submit a hard copy or electronically as a pdf file via the HW3 folder in the course Moodle page.

Solutions to HW2 is on my homepage.

HW3 is due Friday, 09/17, before the end of class. You can either submit a hard copy or electronically as a pdf file via the HW3 folder in the course Moodle page.

Solutions to HW2 is on my homepage.

Outline

One of the most important concepts in probability theory is that of the expectation of a random variable.

Definition

Suppose X is a discrete random variable with mass function $p(\cdot)$. If

$$\sum_{x:p(x)>0} |x|p(x) < \infty$$

then we say that X has finite expectation and we define

$$\mathrm{E[X]} = \sum_{\mathrm{x:p(x)>0}} \mathrm{xp(x)}$$

to be the expectation of X. If $\sum_{x:p(x)>0} |x|p(x) = \infty$, the expectation of X is undefined.

One of the most important concepts in probability theory is that of the expectation of a random variable.

Definition

Suppose X is a discrete random variable with mass function $p(\cdot)$. If

$$\sum_{x:p(x)>0} |x|p(x) < \infty$$

then we say that X has finite expectation and we define

$$\mathrm{E[X]} = \sum_{\mathrm{x:p(x)>0}} \mathrm{xp(x)}$$

to be the expectation of X. If $\sum_{x:p(x)>0} |x|p(x) = \infty$, the expectation of X is undefined.

Example 1

Suppose that A is an event and define

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathrm{E}[\mathrm{I}] = 0 \cdot \mathrm{P}(\mathrm{A}^c) + 1 \cdot \mathrm{P}(\mathrm{A}) = \mathrm{P}(\mathrm{A})$.

Example 2

If X is the outcome when we roll a fair die, then

$$p(1) = \cdots = p(6) = \frac{1}{6}.$$

Thus $\mathrm{E}[\mathrm{X}] = \frac{7}{2}$.

Example 1

Suppose that A is an event and define

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathrm{E}[\mathrm{I}] = 0 \cdot \mathrm{P}(\mathrm{A}^c) + 1 \cdot \mathrm{P}(\mathrm{A}) = \mathrm{P}(\mathrm{A})$.

Example 2

If X is the outcome when we roll a fair die, then

$$p(1) = \cdots = p(6) = \frac{1}{6}.$$

Thus $\mathrm{E}[\mathrm{X}] = \frac{7}{2}$.

Example 3

2 fair dice are rolled. Let X be the sum of the two dice. Then $\mathrm{E}[X] = 7$.

Example 5

Suppose that

$$P(X = 2^{n-1}) = 2^{-n}, \quad n = 1, 2, \ldots.$$

Since

$$\sum_{n=1}^{\infty} 2^{n-1} 2^{-n} = \infty,$$

$\mathrm{E}[X]$ is not defined.

Example 3

2 fair dice are rolled. Let X be the sum of the two dice. Then $\mathrm{E}[\mathrm{X}] = 7$.

Example 5

Suppose that

$$P(X = 2^{n-1}) = 2^{-n}, \quad n = 1, 2, \ldots.$$

Since

$$\sum_{n=1}^{\infty} 2^{n-1} 2^{-n} = \infty,$$

$\mathrm{E}[\mathrm{X}]$ is not defined.

Outline

Suppose that X is a discrete random variable with mass function $p(\cdot)$. Suppose φ is a function on $\mathbb{R}$. Then $\varphi(X)$ is a discrete random variable. We want to find $\mathrm{E}[\varphi(\mathrm{X})]$. Before presenting the general theorem, let's looking at an example first.

Example 1

Suppose that

$$P(X=-1)=\frac{1}{5}, \quad P(X=0)=\frac{1}{2}, \quad P(X=1)=\frac{3}{10}.$$

Find $\mathrm{E}[\mathrm{X}^2]$.

Let $Y=X^2$. Then $P(Y=0)=P(Y=1)=\frac{1}{2}$. Thus $\mathrm{E}[\mathrm{X}^2]=\frac{1}{2}$.

Suppose that X is a discrete random variable with mass function $p(\cdot)$. Suppose φ is a function on $\mathbb{R}$. Then $\varphi(X)$ is a discrete random variable. We want to find $\mathrm{E}[\varphi(\mathrm{X})]$. Before presenting the general theorem, let's looking at an example first.

Example 1

Suppose that

$$P(X=-1)=\frac{1}{5}, \quad P(X=0)=\frac{1}{2}, \quad P(X=1)=\frac{3}{10}.$$

Find $\mathrm{E}[\mathrm{X}^2]$.

Let $Y = X^2$. Then $P(Y=0) = P(Y=1) = \frac{1}{2}$. Thus $\mathrm{E}[\mathrm{X}^2] = \frac{1}{2}$.

Suppose that X is a discrete random variable with mass function $p(\cdot)$. Suppose φ is a function on $\mathbb{R}$. Then $\varphi(X)$ is a discrete random variable. We want to find $\mathrm{E}[\varphi(\mathrm{X})]$. Before presenting the general theorem, let's looking at an example first.

Example 1

Suppose that

$$P(X=-1)=\frac{1}{5}, \quad P(X=0)=\frac{1}{2}, \quad P(X=1)=\frac{3}{10}.$$

Find $\mathrm{E}[\mathrm{X}^2]$.

Let $Y=X^2$. Then $P(Y=0)=P(Y=1)=\frac{1}{2}$. Thus $\mathrm{E}[\mathrm{X}^2]=\frac{1}{2}$.

What we did in the previous example was: first find the mass function of $\varphi(X)$ and then use the definition of expectation. We could also the following proposition

Proposition

Suppose that X is a discrete random variable with mass function $p(\cdot)$ and that φ is a function on $\mathbb{R}$. If

$$\sum_x |\varphi(x)|p(x) < \infty,$$

then $\varphi(X)$ has finite expectation and

$$\mathrm{E}[\varphi(\mathrm{X})] = \sum_{\mathrm{x}} \varphi(\mathrm{x})\mathrm{p}(\mathrm{x}).$$

What we did in the previous example was: first find the mass function of $\varphi(X)$ and then use the definition of expectation. We could also the following proposition

Proposition

Suppose that X is a discrete random variable with mass function $p(\cdot)$ and that φ is a function on $\mathbb{R}$. If

$$\sum_{x} |\varphi(x)| p(x) < \infty,$$

then $\varphi(X)$ has finite expectation and

$$\mathrm{E}[\varphi(\mathrm{X})] = \sum_{\mathrm{x}} \varphi(\mathrm{x})\mathrm{p}(\mathrm{x}).$$

I am not going to give a proof of the above proposition. Let's apply this proposition to the previous example:

$$\mathrm{E}[\mathrm{X}^2] = (-1)^2 \cdot \frac{1}{5} + 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{3}{10} = \frac{1}{2},$$

which coincides with the answer we found before.

Example 2

A coin is tossed 4 times. Suppose that the probability of Heads is $\frac{2}{3}$ on each toss. Let X be the total number of Heads. Find $\mathrm{E}[\sin(\frac{\pi \mathrm{X}}{2})]$.

$$\mathrm{E}[\sin(\frac{\pi \mathrm{X}}{2})] = \sum_{k=0}^{4} \sin(\frac{k\pi}{2}) \binom{4}{k} (\frac{2}{3})^k (\frac{1}{3})^{4-k}$$

$$= 4 \cdot \frac{2}{3} \cdot (\frac{1}{3})^3 - 4 \cdot (\frac{2}{3})^3 \cdot \frac{1}{3}.$$

I am not going to give a proof of the above proposition. Let's apply this proposition to the previous example:

$$\mathrm{E}[X^2] = (-1)^2 \cdot \frac{1}{5} + 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{3}{10} = \frac{1}{2},$$

which coincides with the answer we found before.

Example 2

A coin is tossed 4 times. Suppose that the probability of Heads is $\frac{2}{3}$ on each toss. Let X be the total number of Heads. Find $\mathrm{E}[\sin(\frac{\pi X}{2})]$.

$$\mathrm{E}[\sin(\frac{\pi X}{2})] = \sum_{k=0}^{4} \sin(\frac{k\pi}{2})\binom{4}{k}(\frac{2}{3})^k(\frac{1}{3})^{4-k}$$

$$= 4 \cdot \frac{2}{3} \cdot (\frac{1}{3})^3 - 4 \cdot (\frac{2}{3})^3 \cdot \frac{1}{3}.$$

I am not going to give a proof of the above proposition. Let's apply this proposition to the previous example:

$$\mathrm{E}[\mathrm{X}^2] = (-1)^2 \cdot \frac{1}{5} + 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{3}{10} = \frac{1}{2},$$

which coincides with the answer we found before.

Example 2

A coin is tossed 4 times. Suppose that the probability of Heads is $\frac{2}{3}$ on each toss. Let X be the total number of Heads. Find $\mathrm{E}[\sin(\frac{\pi \mathrm{X}}{2})]$.

$$\begin{aligned}\mathrm{E}[\sin(\frac{\pi \mathrm{X}}{2})] &= \sum_{k=0}^{4} \sin(\frac{k\pi}{2})\binom{4}{k}(\frac{2}{3})^k(\frac{1}{3})^{4-k} \\ &= 4 \cdot \frac{2}{3} \cdot (\frac{1}{3})^3 - 4 \cdot (\frac{2}{3})^3 \cdot \frac{1}{3}.\end{aligned}$$

Corollary

If X is a discrete random variable, and a, b are constants, then

$$\mathrm{E}[aX + b] = a\mathrm{E}[X] + b.$$

This corollary follows immediately from the proposition. Let $p(\cdot)$ be the mass function of X. Then

$$\begin{aligned}\mathrm{E}[aX + b] &= \sum_x (ax + b)p(x) = a\sum_x xp(x) + b\sum_x p(x)\\ &= a\mathrm{E}[X] + b.\end{aligned}$$

The expectation of X, $\mathrm{E}[X]$ is also called the first moment of X. For any integer $n \geq 1$, $\mathrm{E}[X^n]$, if exists, is called the n-th moment of X. If X is a discrete random variable with mass function $p(\cdot)$, then

$$\mathrm{E}[X^n] = \sum_x x^n p(x).$$

Corollary

If X is a discrete random variable, and a, b are constants, then

$$\mathrm{E}[aX + b] = a\mathrm{E}[X] + b.$$

This corollary follows immediately from the proposition. Let $p(\cdot)$ be the mass function of X. Then

$$\begin{aligned}\mathrm{E}[aX + b] &= \sum_x (ax + b)p(x) = a\sum_x xp(x) + b\sum_x p(x) \\ &= a\mathrm{E}[X] + b.\end{aligned}$$

The expectation of X, $\mathrm{E}[X]$ is also called the first moment of X. For any integer $n \geq 1$, $\mathrm{E}[X^n]$, if exists, is called the n-th moment of X. If X is a discrete random variable with mass function $p(\cdot)$, then

$$\mathrm{E}[X^n] = \sum_x x^n p(x).$$

Corollary

If X is a discrete random variable, and a, b are constants, then

$$\mathrm{E}[aX + b] = a\mathrm{E}[X] + b.$$

This corollary follows immediately from the proposition. Let $p(\cdot)$ be the mass function of X. Then

$$\begin{aligned}\mathrm{E}[aX + b] &= \sum_x (ax + b)p(x) = a\sum_x xp(x) + b\sum_x p(x) \\ &= a\mathrm{E}[X] + b.\end{aligned}$$

The expectation of X, $\mathrm{E}[X]$ is also called the first moment of X. For any integer $n \geq 1$, $\mathrm{E}[X^n]$, if exists, is called the n-th moment of X. If X is a discrete random variable with mass function $p(\cdot)$, then

$$\mathrm{E}[X^n] = \sum_x x^n p(x).$$

Outline

Definition

Suppose that X is a discrete random variable with finite $\mathrm{E[X]} = \mu$. If $\mathrm{E[(X-\mu)^2]}$ exists, we call it the variance of the random variable X:

$$\mathrm{Var}(X) = \mathrm{E[(X-\mu)^2]}.$$

The square root of $\mathrm{Var}(X)$ is called the standard deviation of X:

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}.$$

$\mathrm{Var}(X)$ is always non-negative. It measures how spread-out the random variable X is from its mean. If $\mathrm{Var}(X) = 0$, then X is deterministic.

Definition

Suppose that X is a discrete random variable with finite $\mathrm{E[X]} = \mu$. If $\mathrm{E[(X - \mu)^2]}$ exists, we call it the variance of the random variable X:

$$\mathrm{Var}(X) = \mathrm{E[(X - \mu)^2]}.$$

The square root of $\mathrm{Var}(X)$ is called the standard deviation of X:

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}.$$

$\mathrm{Var}(X)$ is always non-negative. It measures how spread-out the random variable X is from its mean. If $\mathrm{Var}(X) = 0$, then X is deterministic.

Here is another formula for $\mathrm{Var}(X)$:

$$\mathrm{Var}(X) = \mathrm{E}[X^2] - (\mathrm{E}[X])^2.$$

Here is a derivation. Let $p(\cdot)$ be the mass function of X.

$$\begin{aligned}\mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2] &= \sum_x (x-\mu)^2 p(x) \\ &= \sum_x (x^2 - 2\mu x + \mu^2) p(x) \\ &= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x) \\ &= \mathrm{E}[X^2] - 2\mu \cdot \mu + \mu^2 = \mathrm{E}[X^2] - \mu^2 \\ &= \mathrm{E}[X^2] - (\mathrm{E}[X])^2.\end{aligned}$$

Here is another formula for $\mathrm{Var}(X)$:

$$\mathrm{Var}(X) = \mathrm{E}[X^2] - (\mathrm{E}[X])^2.$$

Here is a derivation. Let $p(\cdot)$ be the mass function of X.

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}[(X-\mu)^2] = \sum_x (x-\mu)^2 p(x) \\
&= \sum_x (x^2 - 2\mu x + \mu^2) p(x) \\
&= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x) \\
&= \mathrm{E}[X^2] - 2\mu \cdot \mu + \mu^2 = \mathrm{E}[X^2] - \mu^2 \\
&= \mathrm{E}[X^2] - (\mathrm{E}[X])^2.
\end{aligned}$$

Proposition

Suppose that X is a discrete random variable with finite variance. Then for any real numbers a and b,

$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X).$$

In particular,

$$\mathrm{Var}(-X) = \mathrm{Var}(X).$$

Example

Suppose that X is a randomly chosen number from $\{1, 2, \ldots, 10\}$. Find $\mathrm{Var}(X)$.

Proposition

Suppose that X is a discrete random variable with finite variance. Then for any real numbers a and b,

$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X).$$

In particular,

$$\mathrm{Var}(-X) = \mathrm{Var}(X).$$

Example

Suppose that X is a randomly chosen number from $\{1, 2, \ldots, 10\}$. Find $\mathrm{Var}(X)$.

$$p(1) = p(2) = \cdots = p(10) = \frac{1}{10}.$$

$$\mathrm{E[X]} = \frac{1}{10}\sum_{k=1}^{10} k = \frac{55}{10} = \frac{11}{2}.$$

$$\mathrm{E[X^2]} = \frac{1}{10}\sum_{k=1}^{10} k^2 = \frac{1}{10}\,\frac{10 \cdot 11 \cdot 21}{6} = \frac{77}{2}.$$

$$\mathrm{Var}(X) = \frac{77}{2} - (\frac{11}{2})^2.$$