

## 13 2025-02-10 | Week 5 | Lecture 12

Today's lecture is on sections 3.3 and 3.4 in the textbook

### 13.1 Expectation

The *expected value* of a discrete random variable is

$$\mu = \mathbb{E}[X] = \sum_x x \cdot p(x)$$

where the  $x$  in the summation runs over all possible values of  $X$ .

In words, the expected value is the *long-run average*, meaning that if you repeated an experiment many times (independently), the long-run average would converge to  $\mathbb{E}[X]$ .

Textbook uses the notation  $\mu$  or  $\mu_X$  for  $\mathbb{E}[X]$ .

The following is a simple but important fact:

**Proposition 39** (Linearity of Expectation). *Let  $a, b$  be numbers and  $X$  be a random variable. Then*

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

We saw this earlier, when we looked at the expected value of playing the lottery once ( $-\$1$ ) and of playing 10 times ( $-\$10$ ), etc.

**Theorem 40** (Expectation of a function of  $X$ ). *Let  $X$  be a discrete random variable with pmf  $p(x)$ . Let  $h : \mathbb{R} \rightarrow \mathbb{R}$  be any function. Then*

$$\mathbb{E}[h(X)] = \sum_x h(x)p(x)$$

*provided that the sum on the right hand side is absolutely convergent. As usual, the summation runs over all possible values of  $X$ .*

**Example 41.** Suppose  $Y$  is a discrete random variable with pmf

$$p(-2) = \mathbb{P}[Y = -2] = 0.1$$

$$p(-1) = \mathbb{P}[Y = -1] = 0.3$$

$$p(1) = \mathbb{P}[Y = 1] = 0.4$$

$$p(2) = \mathbb{P}[Y = 2] = 0.2$$

and such that  $\mathbb{P}[Y = y] = 0$  if  $y \notin \{-2, -1, 1, 2\}$ . Find the following quantities:

(a)  $\mathbb{E}[Y]$

(b)  $\mathbb{E}[3Y + 7]$

(c)  $\mathbb{E}[Y^3 + 2Y]$

(d)  $\mathbb{E}[e^X]$

*Solution to (a):*

$$\begin{aligned}\mathbb{E}[Y] &= -2(0.1) + (-1)(0.3) + (1)(0.4) + (2)(0.2) \\ &= -.2 - .3 + .4 + .4 \\ &= .3\end{aligned}$$

*Solution to (b):*

$$\begin{aligned}\mathbb{E}[3Y + 7] &= 3\mathbb{E}[Y] + 7 && \text{by Proposition 39} \\ &= 3(0.3) + 7 && \text{by our answer to part (a).} \\ &= 7.9.\end{aligned}$$

*Solution to (c):* Here we will apply Theorem 40 with  $h(x) = x^3 + 2x$ :

$$\begin{aligned}\mathbb{E}[Y^3 + 2Y] &= \sum_{y \in \{-2, -1, 1, 2\}} h(y)p(y) \\ &= h(-2)p(-2) + h(-1)p(-1) + h(1)p(1) + h(2)p(2) \\ &= (-12)(0.1) + (-3)(0.3) + (3)(0.4) + (12)(0.2) \\ &= 1.5.\end{aligned}$$

*Solution to (d):* Here we will apply Theorem 40 with  $h(x) = e^x$ :

$$\begin{aligned}\mathbb{E}[e^Y] &= \sum_{y \in \{-2, -1, 1, 2\}} h(y)p(y) \\ &= \sum_{y \in \{-2, -1, 1, 2\}} e^y p(y) \\ &= e^{-2}p(-2) + e^{-1}p(-1) + e^1p(1) + e^2p(2) \\ &= e^{-2} \cdot 0.1 + e^{-1} \cdot 0.3 + e^1 \cdot 0.4 + e^2 \cdot 0.2 \\ &\approx 2.689.\end{aligned}$$

End of Example 41.  $\square$

## 13.2 Variance

The *variance* of a random variable  $X$  is

$$\text{Var}(X) := \sum_x (x - \mu)^2 p(x)$$

where  $\mu = \mathbb{E}[X]$  and  $x$  ranges over all possible values that  $X$  can take. The variance of a random variable is often denoted  $\sigma^2$ .

**Important formula:** There is a shortcut for computing variance

$$\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

Variance is a measure of how likely the value of a random variable is to be far from its expected value. In an ideal world, we would measure this by

$$\mathbb{E}[|X - \mu|] = (\text{expected distance of } X \text{ from its mean})$$

But unfortunately, the absolute value function is mathematically difficult to work with, so instead we use

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

which is mathematically easier to work with because it doesn't have an absolute value.

We'll do some examples next time