# 3.3: Expected Values of Discrete RVs

Taylor

University of Virginia

# Definition

Let $X$ be a discrete rv with range $D$ and pmf $p(x)$. The **expected value** or **mean** of X, $E(X)$ is

$$E(X) = \sum_{x \in D} x \cdot p(x)$$

We say it exists if it's finite.

# Example

In our last example we had $p(x) = p(1-p)^{x-1}$. Then

$$
\begin{aligned}
EX &= \sum x \cdot p(x) \\
&= p \sum_{x=1}^{\infty} x(1-p)^{x-1} \\
&= p \sum_{\tilde{x}=0}^{\infty} (\tilde{x}+1)(1-p)^{\tilde{x}} \\
&= p \sum_{\tilde{x}=0}^{\infty} \tilde{x}(1-p)^{\tilde{x}} + p \sum_{\tilde{x}=0}^{\infty} (1-p)^{\tilde{x}} \\
&= p \sum_{\tilde{x}=1}^{\infty} \tilde{x}(1-p)^{\tilde{x}} + p \sum_{\tilde{x}=0}^{\infty} (1-p)^{\tilde{x}}
\end{aligned}
$$

# Motivation

$$
\begin{aligned}
\ldots &= p\sum_{x=1}^{\infty} x(1-p)^x + p\sum_{x=0}^{\infty}(1-p)^x \\
&= (1-p)p\sum_{x=1}^{\infty} x(1-p)^{x-1} + p\sum_{x=0}^{\infty}(1-p)^x \\
&= (1-p)E(X) + p\sum_{x=0}^{\infty}(1-p)^x \\
&= (1-p)E(X) + p\frac{1}{1-(1-p)} \\
&= (1-p)E(X) + 1
\end{aligned}
$$

So $E(X) = (1-p)E(X) + 1$ or $E(X) = 1/p$. Note: the book has another way to do this in example 3.18

# Example

Example 3.19 is an example of a "heavy-tailed" distribution. Let $p(x) = \frac{k}{x^2}$, $x > 0$.

$$E(X) = \sum_{x=1}^{\infty} x \frac{k}{x^2} = k \sum_{x=1}^{\infty} \frac{1}{x} = \infty$$

Recall from calculus that $\sum_{x \geq 0} \frac{1}{x^p}$ converges iff $p > 1$ and diverges iff $0 \leq p \leq 1$

# Motivation

Say we have $X$. We can make a new rv $Y = h(X)$ with some function $h(\cdot)$. It would be true that $E(Y)$ could be found using the formula above, but we would need $p_Y(y)$ to do that. We would have to find that from $p_X(x)$. That's a pain. Good news though: we don't have to find the new distribution, though.

# Motivation

Say we start out with $X$ and $p_X(x)$. Then for any function $h(\cdot)$,

$$E(Y) = E[h(X)] = \sum_x h(x) p_X(x)$$

(we're assuming here that these expected values exist i.e. that they're finite)

This is called the law of the unconscious statistician (LOTUS).

# Example

Example 3.22 on page 116: Let $X$ denote the number of computers sold by a small shop. Assume the pmf is $p(0) = .1$, $p(1) = .2$, $p(2) = .3$, and $p(3) = .4$. Let $h(x)$ denote the profit. We pay \$500 per computer up front (\$1500 total), then we try to sell as many as we can for \$1000 a piece. The ones that don't get sold are bought back from the manufacturer at less than we paid (\$200 a piece).

So

$$h(X) = 1000X + 200(3 - X) - 1500 = 800X - 900$$

What's $Eh(X)$?

$$Eh(X) = [-900 \cdot .1] + [-100 \cdot .2] + [700 \cdot .3] + [1500 \cdot .4] = 700$$

# Proposition

A lot of times $h(\cdot)$ is a linear transformation. In this case

$$E[aX + b] = aE(X) + b$$

where $a$ and $b$ are constants

# Definition

Here's the definition of population variance. Let $D$ be the range of a rv $X$. Let $\mu = E(X)$ (it's easier to write it this way). Then the variance of $X$, call it $V(X)$ is:

$$V(X) = \sum_{x \in D} (x - \mu)^2 \cdot p(x) = E[(X - \mu)^2]$$

Standard deviation is just the square root of this.

This is an average again, but we're not taking the average of $X$. We're taking the average of a nonlinear transformation of $X$: $(X - \mu)^2$.

# A convenient formula

Sometimes we use this formula:

$$V(X) = E(X^2) - [E(X)]^2$$

# Proof

$$
\begin{aligned}
V(X) &= E[(X-\mu)^2] \\
&= E[X^2 - 2X\mu + \mu^2] \\
&= E[X^2] - 2E[X]\mu + \mu^2 \\
&= E[X^2] - 2\mu^2 + \mu^2 \\
&= E[X^2] - [E(X)]^2
\end{aligned}
$$

# Proposition

We also have this:

$$V(aX + b) = a^2 V(X)$$

(check)

# Another Example

A lot of special quantities are just expectations of intuitive functions

Entropy:

$$E[-\log p(X)] = \sum_x -\log p(x)p(x)$$

We're using a random variable's pmf as a transformation now.
The transformation $-\log p(X)$ measures "surprise" or "disorder."