1 Random Variables

- Up until now, we have discussed probability in terms of events and samples spaces of probability experiments
- Anothe way we can discuss probability is in terms of *Random Variables*
- Random Variables
  - A *Random Variable* is a numerical representation of the outcome of an experiment
- Essentially, we can think of a random variable as a numerical summary of what happened in a probability experiment

1.1 Examples

- Consider the experiment where we flip a coin
  - We know that the sample Space is $S = \{H, T\}$
  - Let $X = \#$ of observed heads in this experiment
  - If the coin lands heads up, then $X = 1$
  - If the coin lands tails up, then $X = 0$
  - This means that we can talk about the event that the coin lands heads up or tails up in terms of whether $X = 1$ or $X = 0$
- Suppose we instead flip two coins instead of just one
  - We know that the sample space would be $S = \{(H, H), (H, T), (T, H), (T, T)\}$
  - Again we let $X =\#$ of obsereved heads
  - Then if both coins come up tails, then $X = 0$
  - If either coin comes up heads, then $X = 1$
  - If both coins come up heads, then $X = 2$
- Suppose we have a spinner (like the one used in the game twister) that we can flick. Eventually the spinner will stop moving; we will consider this as stopping at a point on the circle. suppose that each point on the circle corresponds to a value between 0 and 1, where at the very top of the spinner is counted as 1 and the very bottom is counted as 0.5. the rest of the points are continuous around the spinner
- *Note:* 0 is not actually included as a possible value
  - Here, the spinner can land on any value between 0 and 1, including 1 but not including 0
  - This means that the sample space is $S = (0, 1]$
  - Let Y be the value that the spinner points to

- This means that instead of talking about what value the spinner points to, we can focus on the value of Y instead
- For example, we can talk about events in terms of Y, such as $Y = .5$, $Y < .5$, and $.35 \leq Y < .75$

1.2 The Support of a Random Variable

- *Note:* Every sample point in the original probability experiment's sample space corresponds with a potential value of the corresponding random variable
- This means that the sample space of the probability experiment is transformed into a ntoher set. We call this set the *Support* of the random vairable
- Support of a random variable
  - The *Support* of a random variable is defined to be the set of all possible values of that random variable
- The support for X in the first experiment (with one coin) would have simply been $S = \{0, 1\}$
- The support for X in the second experiment (with two coins) would have been $S = \{0, 1, 2\}$
- The support for Y in the third experiment (with the spinner) would have been $S = (0, 1]$
- *Note:* We use S for both the sample space and for the support. This is for two reasons
  1. Once we start working with random variables, we do not need to think in terms of the original probability experiment
  2. We can think of the support of the random variable as teh random variable analog of the sample space

2 Discrete Random Variables

- There are (primarily) two kinds of Random Variables
  1. Discrete Random variables : Discrete Random variables have a support that either has a finite or a countably infinite number of elements
  2. Continuous Random variables: Continuous Random Variables have a support that has an uncountably infinite numbe of elements (basically they contain an interval of numbers)
- For now we will focus on Discrete Random Variables

3 Probability Distribution Functions (PDFs)

- Like with a probability experiment, we know that among all of the possible values of the Random variable, the random variable must evaluate to one of the possible out comes
- What we often want to know is how likely is the random value to be each one of those potential values
- In other words, we want to know how the total probability is *distributed* among the different potential values
- We often summarize this with what is called a *Probability Distribution Function*, or PDF for short.
- Probability Distribution Funtion
  - A *Probability Distribution Function* for a random variable X is the function $p_X(☺)$ such that $p_X(☺) = P(X = ☺)$
- Like our Probability Rules, there are rules that all PDFs for discrete random variables follow:
- Let X be a discrete RV(Random Variable) with support S and PDF p_X. Then,
  1. $0 \leq p_X(x) \leq 1 \forall x \in S$ (Remember $\forall =$ "For all")
  2. $\sum_{☺ \in S} p_X(☺) = 1$

4 Example of a Discerete RV

- Suppose you are studying family size in the state of CT. Let X be the number of children in the next family you observed
- Suppose X has the following PDF:

$x =$	0	1	2	3
$P(X = x)$	.1	.5	.2	.2

- This table can be used to give us values for our PDFs
- For example $P(X = 0) = .1$
- Also

$$\begin{aligned} P(X < 2) &= P(X = 0) + P(X = 1) \\ &= .1 + .5 \\ &= .6 \end{aligned}$$

5 Expectation of a discrete RV

- One thing we often want to know about is what value do we *expect* the random variable to be. one way we determine this is trhough a process called *Expectation*
- Expectation
  - Let X be a random variable with support S and PDF p_X, and let $g(☺)$ be a real valued function that is defined $\forall ☺ \in S$. Then the expected value of $g(X)$ (denoted $E[g(X)]$) is defined to be

$$E[g(X)] = \sum_{y \in S} g(y) p_X(y)$$

- When $g(X)$ is just $g(X) = X$, the identity function, then it is called the *mean* of X (in additiona to being refered to as the expected value of X)
- When $g(X)$ is $g(X) = (X - E[X])^2$, then $E[g(X)]$ is also refered to the *variance* of X. The variance of X is often denoted $V[X]$.
- *Note:* The use of the terms mean and variance is not by mistake. these terms are used because the corresponding expecations are designed to capture the concepts of the mean (where is the middle) and the variance (how spread out) of the distribution of the probability of a RV in the same way that our definitions of mean and variance are meant to caputre where the location (mean) and how spread out (variance) a data set (or a set of numbers) is.

6 Expectation Theorems

- Let X be a discrete RV with support S and PDF p_X
- Let c be a real valued constant. Then,

$$E[c] = c$$

Proof.

$$\begin{aligned}
E[c] &= \sum_{x \in S} c p_X(x) \\
&= c \sum_{x \in S} p_X(x) \leftarrow \text{Pulling out a factor of } c \\
&= c \leftarrow \text{By Second rule of Discrete PDFs}
\end{aligned}$$

□

- Let c be a real valued constant and let g be a real valued function that is defined over the support of X. Then,

$$E[cg(X)] = cE[g(X)]$$

Proof.

$$\begin{aligned} E[cg(X)] &= \sum_{x\in S} cg(X)p_X(x) \\ &= c\sum_{x\in S} g(X)p_X(x) \leftarrow \text{Pulling out a factor of } c \\ &= cE[g(X)] \leftarrow \text{Definition of } E[g(X)] \end{aligned}$$

$\square$

- Let $g_1, g_2, g_3, ..., g_n$ be n real valued functions that are all defined over the support of X. Then,

$$E[\sum_{i=1}^{n} g_i(X)] = \sum_{i=1}^{n} E[g_i(X)]$$

Proof.

$$\begin{aligned} E[\sum_{i=1}^{n} g_i(X)] &= \sum_{x\in S}\sum_{i=1}^{n} g_i(X)p_X(x) \\ &= \sum_{i=1}^{n}\sum_{x\in S} g_i(X)p_X(x) \leftarrow \text{Simply rearranging the order of the terms.} \\ &\qquad\qquad \text{This is just the commutative law of addition} \\ &= \sum_{i=1}^{n} E[g_i(X)] \leftarrow \text{Defintion of } E[g_i(X)] \end{aligned}$$

$\square$

- Variance Equality

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

Note: $(E[X])^2$ is often denoted as $E^2[X]$

Proof.

$$
\begin{aligned}
V[X] &= E[(X - E[X])^2] \\
&= E[X^2 - 2E[X] + E^2[X]] \\
&= E[X^2]E[-2XE[X]] + E[E^2[X]] \leftarrow \text{By the Third Expectation Rule} \\
&= E[X^2] - 2E[X]E[X] + E^2[X] \leftarrow \text{By the First and second Expectation Rules since} \\
&\qquad\qquad\qquad\qquad\qquad\qquad -2E[X] \text{ and } E^2[X] \text{ are both constant values} \\
&= E[X^2] - 2E^2[X] + E^2[X] \\
&= E[X^2] - E^2[X]
\end{aligned}
$$

□