Section 3.1 - Random Variable

**Definition 1** (Random Variable)**.** A random variable is a function that associates a real number with each element in the sample space.

A random variable is used when you know the possible values it could have, but not what value it will actually take.

*Example* 1. Roll a 6-sided die. What values can it take on? What value will come up? Define a random variable for this scenario.

**Definition 2** (Discrete Random Variable)**.** A random variable whose image is countable.

*Example* 2. We wait in line to see a movie. We know that the waiting time in line can be anywhere from 0 hours to maybe 3 hours. However, we don't know how long the wait is actually going to be. Define a random variable for this scenario.

**Definition 3** (Continuous Random Variable)**.** A random variable whose image is an interval and whose CDF is continuous.

Section 3.2 - Probability Mass Function (PMF)

**Definition 4** (Probability Distribution (Discrete Random Variable))**.** Writing down a formula or a table to represent all of the probabilities of the values of the random variable.

**Definition 5** (Probability Mass Function (PMF))**.** A function, $f(x)$, associated with a random variable $X$.

PMF Properties:

*Example* 3 (Probability Calculations).
Toss a coin 2 times. Let $X$ be a random variable representing the number of heads that occur.

(a) Identify the sample space $S$.

(b) Write the PMF for $X$.

(c) Find $\mathsf{P}(X \leq 1)$.

(d) Find $\mathsf{P}(X < 1)$.

(e) Find $\mathsf{P}(X \leq 2)$.

(f) Find $\mathsf{P}(X < 2)$.

(g) Find $\mathsf{P}(X \geq 1)$.

(h) Find $\mathsf{P}(X > 1)$.

*Example* 4 (Valid PMF Check).
Suppose we know

$$f(x) = c(x + 5), \quad x = 1, 2, 3.$$

What value of $c$ would make this a valid pmf?

*Example* 5. Find a formula for the probability distribution of the random variable (X) representing the outcome when a single die is rolled once.

Section 4.1 - Expected Value

*Example* 6 (Expected Value Intro Example).
You play a game where the roll of 1 fair die determine whether you win or lose.

- If you roll a 5 or 6, you win \$5.
- If you roll a 3 or 4, you win \$0 (breakeven).
- If you roll a 1 or 2, you lose \$2.

How much do you win/lose on average, if you play a really long time?

**Definition 6** (Expected Value (General Formula))**.** Given a function $g(x)$ from $\mathbb{R}$ to $\mathbb{R}$, we define

$$\mu_{g(X)} = \mathbf{E}\left[g(X)\right] = \sum_x g(x)f(x)$$

*Example* 7. Suppose $f(x)$, the pmf for a discrete random variable $X$ is given in the table below:

| $x$ | 5 | 10 | 15 |
|---|---|---|---|
| $f(x)$ | 0.50 | 0.10 | 0.40 |

(a) What is $\mathsf{P}(X = 5)$?

(b) What is $\mathsf{P}(X = 6)$?

(c) What is $\mathsf{P}(X > 8)$?

(d) What is $\mathsf{P}(X > 10)$?

(e) Compute $\mu$.

(f) Compute $\sigma$.

Section 4.2 - Variance and Standard Deviation

**Definition 7** (Variance)**.** The variance of a discrete random variable $X$ is denoted by $\sigma^2$ or $\mathbf{V}(X)$.

$$\sigma^2 = \mathbf{V}(X) = \mathbf{E}\left[(X-\mu_X)^2\right] = \sum_x (x-\mu_X)^2 f(x) = \mathbf{E}(X^2) - [\mathbf{E}(X)]^2$$

**Definition 8** (Standard Deviation)**.** The standard deviation of $X$ is denoted by $\sigma$ or $\mathbf{SD}(X)$. The standard deviation of $X$ is the square root of the variance of $X$:

$$\sigma = \mathbf{SD}(X) = \sqrt{\mathbf{V}(X)} = \sqrt{\sigma^2}$$

*Example* 8 (Citrus Farmer: Expected Value / Variance / Standard Deviation).
A citrus farmer observed the following distribution for $X$, the number of oranges per tree.

| $x$ | 25 | 30 | 35 | 40 |
|---|---|---|---|---|
| $f(x)$ | 0.1 | 0.4 | 0.3 | 0.2 |

(a) What is the expected value of $X$?

(b) What is the variance of $X$?

**Definition 9** (Variance (General Formula))**.** The variance of a function of $X$, say $g(X)$ is given by

$$\sigma^2_{g(X)} = \mathbf{V}\left[g(X)\right] = \mathbf{E}\left\{\left[g(X) - \mu_{g(X)}\right]^2\right\} = \sum_x \left[g(x) - \mu_{g(X)}\right]^2 f(x).$$

*Example* 9 (Example 8 continued).

(a) What is the expected value of $g(X) = 2X + 1$?

(b) What is the variance of $g(X) = 2X + 1$?

Section 4.3 - Linear Combinations of Random Variables

**Theorem 1** (Expected Value Linear Combinations). *Let $a, b$ be constants. Let $X$ be a random variable. Let $g$ and $h$ be functions of $X$. Then*

1. $\mathbf{E}(aX) = a\,\mathbf{E}(X)$
2. $\mathbf{E}(b) = b$
3. $\mathbf{E}(aX \pm b) = a\,\mathbf{E}(X) \pm b$
4. $\mathbf{E}\left[g(X) \pm h(X)\right] = \mathbf{E}\left[g(X)\right] \pm \mathbf{E}\left[h(X)\right]$

**Theorem 2** (Variance Linear Combinations). *Let $a, b$ be constants. Let $X$ be a random variable.*

1. $\mathbf{V}(aX) = a^2\,\mathbf{V}(X)$
2. $\mathbf{V}(X \pm b) = \mathbf{V}(X)$
3. $\mathbf{V}(aX \pm b) = a^2\,\mathbf{V}(X)$

*Example* 10 (Example 8 continued).

| $x$ | 25 | 30 | 35 | 40 |
|---|---|---|---|---|
| $f(x)$ | 0.1 | 0.4 | 0.3 | 0.2 |

(a) What is the expected value of $g(X) = 2X + 1$? Use the shorter formula.

(b) What is the variance of $g(X) = 2X + 1$? Use the shorter formula.

(c) Find the standard deviation of $g(X) = 2X + 1$.

**Section 4.3 Proofs: Theorem 1 (#1 and 2 during lecture; #3 and 4 read on own)**

1. Proof of $\mathbf{E}(aX) = a\,\mathbf{E}(X)$.

$$\begin{aligned}\mathbf{E}(aX) &= \sum_x axf(x);\ \text{definition }\ \mathbf{E}(X)\\ &= a\sum_x xf(x);\ a \text{ is a constant}\\ &= a\,\mathbf{E}(X);\ \text{definition }\ \mathbf{E}(X)\end{aligned}$$

2. Proof of $\mathbf{E}(b) = b$.

$$\begin{aligned}\mathbf{E}(b) &= \sum_x bf(x);\ \text{definition }\ \mathbf{E}(X)\\ &= b\sum_x f(x);\ b \text{ is a constant}\\ &= b \text{ since } \sum_x f(x) = 1\end{aligned}$$

3. Proof of $\mathbf{E}(aX \pm b) = a\,\mathbf{E}(X) \pm b$.

$$\begin{aligned}\mathbf{E}(aX \pm b) &= \sum_x (ax \pm b)\, f(x);\ \text{definition }\ \mathbf{E}(X)\\ &= \sum_x (axf(x) \pm bf(x))\,;\ \text{distribute}\\ &= \sum_x axf(x) \pm \sum_x bf(x);\ \text{split into 2 sums}\\ &= \mathbf{E}(aX) \pm \mathbf{E}(b);\ \text{see parts 1 and 2}\\ &= a\,\mathbf{E}(X) \pm b;\ \text{see parts 1 and 2}\end{aligned}$$

4. Proof of $\mathbf{E}\left[g(X) \pm h(X)\right] = \mathbf{E}(g(X)) \pm \mathbf{E}(h(X))$.

$$\begin{aligned}\mathbf{E}\left[g(X) \pm h(X)\right] &= \sum_x (g(x) \pm h(x))\, f(x);\ \text{definition }\ \mathbf{E}(X)\\ &= \sum_x \left[g(x)f(x) \pm h(x)f(x)\right];\ \text{distribute}\\ &= \sum_x g(x)f(x) \pm \sum_x h(x)f(x);\ \text{split into 2 sums}\\ &= \mathbf{E}(g(X)) \pm \mathbf{E}(h(X));\ \text{definition }\ \mathbf{E}(X)\end{aligned}$$

**Section 4.3 Proofs: Theorem 2 (#1 during lecture; #2 and 3 read on own)**
Recall $\mathbf{V}(X) = \mathbf{E}(X^2) - [\mathbf{E}(X)]^2$.

1. Proof of $\mathbf{V}(aX) = a^2\,\mathbf{V}(X)$.

$$\begin{aligned}
\mathbf{V}(aX) &= \mathbf{E}\left[(aX)^2\right] - [\mathbf{E}(aX)]^2\,;\ \text{definition of variance}\\
&= \mathbf{E}\left(a^2X^2\right) - [a\,\mathbf{E}(X)]^2\,;\ \text{multiply and pull out constant}\\
&= a^2\,\mathbf{E}(X^2) - a^2\,[\mathbf{E}(X)]^2\,;\ \text{pull out constant and multiply}\\
&= a^2\left[\mathbf{E}(X^2) - [\mathbf{E}(X)]^2\right]\,;\ \text{factor out } a^2\\
&= a^2\,\mathbf{V}(X);\ \text{definition of variance}
\end{aligned}$$

2. Proof of $\mathbf{V}(X \pm b) = \mathbf{V}(X)$.

$$\begin{aligned}
\mathbf{V}(X \pm b) &= \sum_x \left[(x \pm b) - \mathbf{E}(X \pm b)\right]^2 f(x);\ \text{definition of variance}\\
&= \sum_x \left[x \pm b - \mathbf{E}(X) \mp b\right]^2 f(x);\ \text{property of expected value}\\
&= \sum_x \left[x - \mathbf{E}(X)\right]^2 f(x)\\
&= \mathbf{V}(X);\ \text{definition of variance}
\end{aligned}$$

3. Proof of $\mathbf{V}(aX \pm b) = a^2\,\mathbf{V}(X)$.

$$\begin{aligned}
\mathbf{V}(aX \pm b) &= \sum_x \left[ax \pm b - \mathbf{E}(aX \pm b)\right]^2 f(x);\ \text{definition of variance}\\
&= \sum_x \left[ax \pm b - a\,\mathbf{E}(X) \mp b\right]^2 f(x);\ \text{property of expected value}\\
&= \sum_x \left[ax - a\,\mathbf{E}(X)\right]^2 f(x)\\
&= \mathbf{V}(aX);\ \text{definition of variance}\\
&= a^2\,\mathbf{V}(X);\ \text{see part 1}
\end{aligned}$$

*Example* 11 (YOU TRY).
Let $X$ be a random variable with probability distribution

| $x$ | 0 | 1 | 3 |
|---|---|---|---|
| $f(x)$ | 1/3 | 1/2 | 1/6 |

(a) What is $\mathsf{P}(X = 2)$?

(b) What is $\mathsf{P}(X < 2)$?

(c) What is $\mathbf{E}(X)$?

(d) What is $\mathbf{V}(X)$?

(e) What is $\mathbf{E}(Y)$, where $Y = (X - 1)^2$?

Example `R` code for Expected Value and Variance from Distributions

```
values <- c(0, 1, 3)
prob <- c(1/3, 1/2, 1/6)

## E(X) Calculation
ex <- sum(values * prob)

## E(X^2) Calculation
exsq <- sum( (values^2) * prob)

## V(X) Calculation
exsq - (ex^2)
```

Section 3.2 - Cumulative Distribution Function (CDF)

*Example* 12. Suppose we have the following probability distribution for $X$:

| $x$ | 0 | 1 | 2 | 3 | 4 | 6 |
|---|---|---|---|---|---|---|
| $f(x)$ | 0.1 | 0.2 | 0.1 | 0.3 | 0.1 | 0.2 |

Draw a graph of the PMF and the CDF of $X$. Write down the CDF as a piecewise function.

**Definition 10** (Cumulative Distribution Function (CDF))**.** The cumulative distribution function, or CDF, of a discrete random variable $X$ is defined as

$$F(x) = \mathsf{P}(X \leq x) = \sum_{t \leq x} f(t), \quad for - \infty < x < \infty.$$

- $F(x)$ is defined for all real numbers.
- $F(x)$ is a step function for discrete random variables with steps at each point of non-zero probability.
- The height of a step at each point of non-zero probability is the probability of that point.
- $\lim_{x \to -\infty} F(x) = 0$
- $\lim_{x \to +\infty} F(x) = 1$
- $\lim_{x \downarrow x_0} F(x) = F(x_0)$ ($F$ is right continuous)
- $\mathsf{P}(a < X \leq b) = F(b) - F(a)$ for $a < b$

*Example* 13. Suppose we have the following probability distribution for $X$:

| $x$ | 0 | 1 | 2 | 3 | 4 | 6 |
|---|---|---|---|---|---|---|
| $f(x)$ | 0.1 | 0.2 | 0.1 | 0.3 | 0.1 | 0.2 |

(a) Find $\mathsf{P}(X = 2)$.

(b) Find $\mathsf{P}(X \leq 2)$.

(c) Find $\mathsf{P}(X < 2)$.

(d) Find $\mathsf{P}(X > 3)$.

(e) Find $\mathsf{P}(2 \leq X \leq 4)$.

*Example* 14. Suppose $f(x) = cx$ for $x = 1, 2, 3, 4$ is a pmf for a random variable $X$.

(a) Find $c$.

(b) Find the probability that $X$ is even.

(c) Determine the CDF, $F(x)$. Write your answer as a piecewise function.

(d) Graph the CDF, $F(x)$.

*Example* 15. An insurance company offers its policyholders a number of different premium payment options. For a randomly selected policyholder, let $X$ be the number of months between successive payments. The CDF of $X$ is

$$F(x) = \begin{cases} 0, & \text{if } x < 1, \\ 0.4, & \text{if } 1 \leq x < 3, \\ 0.6, & \text{if } 3 \leq x < 5, \\ 0.8, & \text{if } 5 \leq x < 7, \\ 1.0, & \text{if } x \geq 7. \end{cases}$$

Note: When $F(x)$ is NOT based on functions of $x$ (like above), it tells you that $X$ is a discrete random variable.

(a) What is the PMF of $X$?

(b) Compute $\mathsf{P}(4 < X \leq 7)$.

(c) Compute $\mathsf{P}(4 \leq X \leq 7)$.

(d) Compute $\mathsf{P}(3 \leq X \leq 7)$.

*Example* 16 (YOU TRY). Suppose $S = \{1, 2, 6, 9, 10\}$ for the random variable $X$. How would you find the following? What if the probabilities were 0.1, 0.2, 0.2, 0.1, 0.4, respectively?

(a) $\mathsf{P}(X \leq 2)$

(b) $\mathsf{P}(X < 2)$

(c) $\mathsf{P}(X \geq 6)$

(d) $\mathsf{P}(X > 6)$

(e) $\mathsf{P}(6 \leq X \leq 10)$

(f) $\mathsf{P}(6 < X < 10)$

(g) $\mathsf{P}(6 < X \leq 10)$

(h) $\mathsf{P}(6 \leq X < 10)$

*Example* 17 (YOU TRY). Suppose you are given the following CDF. Find the PMF. Write your answer in table format.

$$F(x) = \begin{cases} 0, & x < 3 \\ 0.4, & 3 \le x < 6 \\ 0.45, & 6 \le x < 9 \\ 0.6, & 9 \le x < 10 \\ 0.98, & 10 \le x < 14 \\ 1, & 14 \le x \end{cases}$$

*Example* 18 (YOU TRY). A company is considering an investment that will earn \$10,000 with a probability of 0.65, lose \$5,000 with a probability of 0.30, and earn \$100,000 with probability 0.05. Compute the expected value and standard deviation of the amount earned by the investment.

*Example* 19 (YOU TRY). An overseas shipment of 5 foreign automobiles contains 2 that have slight paint blemishes. If an agency receives 3 of these automobiles at random, list the elements of the sample space $S$, using the letters $B$ and $N$ for blemished and non-blemished, respectively, then to each sample point, assign a value $x$ of the random variable $X$ representing the number of automobiles with paint blemishes purchased by the agency.

*Example* 20 (YOU TRY). Determine the value of $c$ so that $f(x) = cx^2$ for $x = 1, 2, 3$ is a probability distribution of the discrete random variable $X$.

*Example* 21 (YOU TRY). Suppose that the probabilities are 0.4, 0.3, 0.2, and 0.1, respectively, that 0, 1, 2, or 3 power failures will strike a certain subdivision in any given year. Find the mean and variance of the random variable $X$ representing the number of power failures striking this subdivision.

*Example* 22 (YOU TRY - HARD). A salesperson has 2 appointments on a given day. At the first appointment, the salesperson believes that he/she has a 70% chance to make the deal, earning a $1000 commission. At the second appointment, the salesperson believes that he/she has a 40% chance to make the deal, earning a $1500 commission. Assume the appointments are independent. What is the expected commission?