Probability Models

They are wrong, but can be useful.
• **Discrete**: A variable that is ‘countable.’ For example, number of credits taken at school, or the value of a die roll.

• **Continuous**: The variable can be placed on a number line with infinite possibilities of values.
A random variable assumes a value based on the outcome of a random event.

- We use a capital letter, like \( X \), to denote a random variable.
- A particular value of a random variable will be denoted with the corresponding lower case letter, in this case \( x \).

For example: \( P(X=x) = \frac{1}{6} \) looks confusing, but is only missing context.
- \( P(\text{rolling a die}=4) = \frac{1}{6} \) makes more sense.
A probability model for a random variable consists of:

- The collection of all possible values of a random variable, and
- the probabilities that the values occur.
If your Data looks like these:

<table>
<thead>
<tr>
<th>Cluster Size</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Frequency</td>
<td>25</td>
<td>12.5</td>
<td>6.25</td>
<td>3.13</td>
<td>1.56</td>
<td>0.78</td>
<td>0.39</td>
<td>0.19</td>
<td>0.10</td>
<td>0.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome</th>
<th>+$5000</th>
<th>-$10000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.80</td>
<td>0.20</td>
</tr>
</tbody>
</table>

You draw one card from a standard deck of playing cards. If you pick a heart, you will win $10. If you pick a face card, which is not a heart, you win $8. If you pick any other card, you lose $6. Do you want to play? Explain.
• Then, you can use the **EXPECTED VALUE** model.

• We can find the mean and the standard deviation of the data.
The population mean is equal to Expected Value which is the SUM of The values times the probabilities.

\[ E(X) = \mu_x = \sum x_i \cdot p_i \]
\[ \text{Var}(X) = \sigma_x^2 = \sum (x_i - \mu_x)^2 \cdot p_i \]

or

\[ \sigma = \sqrt{\sum (x_i - \mu_x)^2 \cdot p_i} \]
An insurance company charges $5,000 per year in premiums and will pay out $10,000 if there is a common type of accident. There is a 20% chance of the accident occurring.
Find the Mean & S.D.

<table>
<thead>
<tr>
<th>Cluster Size</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Frequency</td>
<td>25</td>
<td>12.5</td>
<td>6.25</td>
<td>3.13</td>
<td>1.56</td>
<td>0.78</td>
<td>0.39</td>
<td>0.19</td>
<td>0.10</td>
<td>0.05</td>
</tr>
</tbody>
</table>
You draw one card from a standard deck of playing cards. If you pick a heart, you will win $10. If you pick a face card, which is not a heart, you win $8. If you pick any other card, you lose $6. Do you want to play? Explain.
25. **Random variables.** Given independent random variables with means and standard deviations as shown, find the mean and standard deviation of:

a) $3X$

b) $Y + 6$

c) $X + Y$

d) $X - Y$

e) $X_1 + X_2$

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X$</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>$Y$</td>
<td>20</td>
<td>5</td>
</tr>
</tbody>
</table>
30. **Garden.** A company selling vegetable seeds in packets of 20 estimates that the mean number of seeds that will actually grow is 18, with a standard deviation of 1.2 seeds. You buy 5 different seed packets.

a) How many bad seeds do you expect to get?
b) What’s the standard deviation?
c) What assumptions did you make about the seeds? Do you think that assumption is warranted? Explain.
37. **Cereal.** The amount of cereal that can be poured into a small bowl varies with a mean of 1.5 ounces and a standard deviation of 0.3 ounces. A large bowl holds a mean of 2.5 ounces with a standard deviation of 0.4 ounces. You open a new box of cereal and pour one large and one small bowl.

a) How much more cereal do you expect to be in the large bowl?

b) What’s the standard deviation of this difference?

c) If the difference follows a Normal model, what’s the probability the small bowl contains more cereal than the large one?
d) What are the mean and standard deviation of the total amount of cereal in the two bowls?

e) If the total follows a Normal model, what's the probability you poured out more than 4.5 ounces of cereal in the two bowls together?

f) The amount of cereal the manufacturer puts in the boxes is a random variable with a mean of 16.3 ounces and a standard deviation of 0.2 ounces. Find the expected amount of cereal left in the box and the standard deviation.
The homework is problems 39, 43, and 45. Pay close attention to the words & the descriptions. Problem 25 we did in class is used in each.
YOU READY, DUDE?
YUP.

TIGHT GROUP
UNITED WE STUDY,
DIVIDED WE FAIL.
• A radio announcer crew postulated on “DeflateGate” that the balls could have had a very small hole in them. The radio announcers then asked, what is the probability that 11 out of 12 balls had holes.

• If the probability of a hole is 1%:

• If the probability of a hole is 5%:

• What if we were looking for the probability the first deflated ball was the 11\textsuperscript{th} ball?

• What if we were looking for the probability the first deflated ball was the 5\textsuperscript{th} ball?
A Geometric model tells us the probability for a random variable that counts the number of successes until a fail.

One parameter defines the Geometric model:

- \( p \), the probability of success.
- We denote this \( \text{Geom}(p) \).
• **Conditions / Assumptions for Geometric Probability**

1. There are two possible outcomes (success and failure).

2. The probability of success, $p$, is constant.

3. Independence: The sample is less than 10% of the population.
Geometric probability model for Bernoulli trials: \textbf{Geom}(p)

$p = \text{probability of success}$

$1 - p = \text{probability of failure}$

$X = \text{number of trials until the first success occurs}$

$$P(X = x) = p(1 - p)^{x-1}$$
Mr. Waddell has a secret life as a semi-pro basketball player. He is currently shooting 75% from the free throw line. What is the probability that his first miss was the 4th shot?

P(making a shot) = .75
P(missing a shot) = .25

x = 4, the first ‘event’

P(missing 4th shot) = (.75)^3(.25)^1 = 0.1055
For the geometric model ONLY WHEN the conditions are met:

\[ \mu = E(X) = \frac{1}{p} \quad \text{and} \]

\[ \sigma = \sqrt{\frac{1-p}{p^2}} \]
Mr. Waddell has a secret life as a semi-pro basketball player. He is currently shooting 75% from the free throw line. What is the mean number of baskets he makes before missing? What is the standard deviation?

\[
\mu = E(X) = \frac{1}{p} = \frac{1}{.75} = 1.333
\]

Average shots made before missing

\[
\sigma = \sqrt{\frac{1-p}{p^2}} = \sqrt{\frac{.25}{.75^2}} = 0.6667
\]

Shots made
• A Binomial model tells us the probability for a random variable that counts the number of successes in a fixed number of Bernoulli trials.

• Two parameters define the Binomial model:
  • $n$, the number of trials; and, $p$, the probability of success.
  • We denote this $\text{Binom}(n, p)$. 
Binomial probability model for Bernoulli trials: 
Binom(n,p)

\[ P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \]
Mr. Waddell has a secret life as a semi-pro basketball player. He is currently shooting 75% from the free throw line. What is the probability he makes exactly 7 out of 10 free throws?

- \( n = 10 \)
- \( p = 0.75; \quad 1 - p = 0.25 \)
- \( X = 7 \)
- \( P(\text{makes 7 out of 10}) = \binom{10}{7} \cdot 0.75^7 \cdot 0.25^3 = 0.2503 \)
For the binomial model **ONLY WHEN** the conditions are met:

- \( \mu = E(X) = np \) and

- \( \sigma = \sqrt{np(1 - p)} \)
The conditions for the binomial model are:

1. There are **two possible outcomes** (success and failure) with a **FIXED number of trials**.

2. The **probability of success**, \( p \), is **constant**.

3. **Independence**: The sample is less than 10% of the population

4. **Success/Fail condition**: \( np \geq 10 \); \( n(1-p) > 10 \)
Mr. Waddell has a secret life as a semi-pro basketball player. He is currently shooting 75% from the free throw line. What is the mean number of free throws he makes if he shoots 10 free throws per game? What is the standard deviation?

- \( \mu = E(X) = np = 10(0.75) = 7.5 \) shots made

- \( \sigma = \sqrt{np(1-p)} = \sqrt{10 \cdot .75 \cdot .25} = 1.369 \) shots
A radio announcer crew postulated on “DeflateGate” that the balls could have had a very small hole in them. The radio announcers then asked, what is the probability that 11 out of 12 balls had holes.

- If the probability of a hole is 1%:
- If the probability of a hole is 5%:
- What if we were looking for the probability the first deflated ball was the 11\textsuperscript{th} ball?
- What if we were looking for the probability the first deflated ball was the 5\textsuperscript{th} ball?
Uh, oh, it got more complicated - stupid English language

- Less than
- Greater than
- Exactly
- Some
- None
3 types of probability models.

- Expected Value model
  - When we’re given probabilities & values (a table can be made)
- Geometric model
  - When we’re interested in the number of Bernoulli trials until the next success.
- Binomial model
  - When we’re interested in the number of successes in a certain number of Bernoulli trials.