Solutions

Math 243 **Sections 3.1** **Introduction to Probability Lab**

Overview

- Why Study Probability?
- Outcomes, Events, Sample Space, Trials
- Probabilities and Complements (not)
- Theoretical vs. Empirical Probability
- The Law of Large Numbers (LLN)
- Disjoint and Independent Events
- Sampling With and Without Replacement and the 10% Rule

Small Group Lab: These are some activities to guide you through the language and concepts of probability. Have one person read each section out loud and help each other.

Why Study Probability? Probability Forms the Foundation for Inferential Statistics

Three Examples:

a. A certain county has a population that is 50% women and 50% men. A jury is supposedly randomly selected. The jury ends up having a composition that is 40% women. Was there selection bias, or was this just due to random chance?

b. In a randomized double-blind controlled experiment, a new surgery saved lives 60% of the time, while the old surgery saved lives only 55% of the time. Is this a big enough difference to replace the old surgery with the new one? In other words, is the difference statistically significant?

c. A pack of potato chips is supposed to be manufactured to have an average weight of 10 ounces. Thirty random bags of chips are weighed, and have an average weight of 9.6 ounces. Is the manufacturer cheating?

> If the bags really have an average of 10 ounces, what is the probability we would get a sample average this low? We will learn to calculate this probability in chapter 3.

Flipping a Coin

1. If you flip a coin, there are two possible **outcomes**.

> Let event H = the outcome of heads. An **event** is a combination of outcomes.
>
> Let event T = the outcome of tails.

2. The **sample space** is the set of all possible outcomes. List the sample space for flipping one coin.

S = { H , T }

3. For each event we can state the probability or chance that it will occur. The **theoretical probability** is based on the idea of a **fair** coin where both outcomes are equally likely. A probability can be written as a fraction or as a decimal between 0 and 1.

P(H) means the probability of getting heads. P(H) = $\frac{1}{2}$ or .5 or 50%

P(T) means the probability of getting tails. P(T) = $\frac{1}{2}$ or .5 or 50%

4. Get a penny for each group member and each flip them 20 times. Record your results below:

Trial	1	2	3	4	5	6	7	8	9	10
Outcome	H	H	H	H	H	H	T	T	H	H

Trial	11	12	13	14	15	16	17	18	19	20
Outcome	H	H	H	H	T	H	H	T	H	T

5. "Heads" or event H, was the result what percentage of the time? This is an **empirical probability** because it is the result of observations or an experiment. Find 3 other people and write down their empirical probabilities.

Yours 15/20 or .75 3 Classmates .35 .40 .60

6. How do your empirical and theoretical probabilities compare?

They are quite different!

7. Is each flip of the coin **independent**? If knowing the result of one flip does not change the probability of the next flip they are independent events.

yes - one coin flip does not influence the next coin flip

8. What would happen to the percentage of "heads" if you flipped the coin 100 times? 1,000 times? 1,000,000 times?

The percentage of heads would get closer and closer to 50% or the theoretical probability

If you said your percentage of "heads" would get closer and closer to 50% you are right!

The Law of Large Numbers (LLN) states that the long-run relative frequency of repeated independent events gets closer and closer to the *true* relative frequency as the number of trials increases.

Rolling a Fair Die

Actually rolling is optional here. We're going to find theoretical probabilities, not empirical. It can be helpful to look at the die to see the probabilities, though.

1. If you roll a fair die, what are the possible outcomes? List the sample space.

S = $\{1, 2, 3, 4, 5, 6\}$

2. To find the theoretical probabilities, take the number of ways the event can occur and divide it by the total number of possibilities. You can leave the probability as a fraction.

$$P(1) = \frac{\text{The number of ways to roll a 1}}{\text{The total number of possible outcomes}} = \frac{1}{6}$$

Theoretical Probabilities

P(1) = $\frac{1}{6}$ P(2) = $\frac{1}{6}$ P(3) = $\frac{1}{6}$

P(4) = $\frac{1}{6}$ P(5) = $\frac{1}{6}$ P(6) = $\frac{1}{6}$

"OR" Events

If A and B are disjoint events, **P(A or B) = P(A) + P(B)** or → add

3. Using your theoretical probabilities, what is the probability of rolling a 2 or a 5? These are **disjoint** or **mutually exclusive** events because they cannot happen at the same time.

To calculate an "OR" probability for disjoint events, add the probabilities together.

P(2 or 5) = P(2) + P(5) = $\frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$

4. What is the probability that you would roll a 3, 5, or 6?

$P(3, 5 \text{ or } 6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$

5. What is the probability that you would roll an even number?

$P(\text{even}) = \frac{1}{2}$

6. What is the probability of rolling at least a 5?

$P(5 \text{ or } 6) = \frac{2}{6} = \frac{1}{3}$

7. What is the probability of rolling a 4 at most?

$P(1, 2, 3 \text{ or } 4) = \frac{4}{6} = \frac{2}{3}$

Complements

8. What is the probability that you would **not** roll a 3? (This is the **complement** of rolling a 3, denoted by 3^C.)

$P(3^C)$ or P(not 3) = $P(1, 2, 4, 5 \text{ or } 6) = \frac{5}{6}$

9. A probability and its complement add up to what number?

$P(3) + P(3^C) = \frac{1}{6} + \frac{5}{6} = 1$

Another way to say this is $P(3^C) = 1 - P(3)$

$1 - \frac{1}{6} = \frac{5}{6}$

Successive Trials and "AND" Events

Two events are called **independent** if knowing that one occurs does not change the probability that the other occurs. Are successive rolls of the dice independent? If they are then we can multiply the probabilities together.

For independent events, P(A and B) = P(A)·P(B) and → multiply

10. For successive rolls, what is the probability of rolling a 3 and then a 6?

$P(3 \text{ and } 6) = P(3) \cdot P(6) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

11. What is the probability of rolling a 1 and then a 5?

$P(1 \text{ and then } 5) = P(1) \cdot P(5) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

12. What is the probability of rolling a 1, 5 times in a row? Would this be the same as the probability of rolling 5 dice and getting all 1's?

$P(1 \text{ and } 1 \text{ and } 1 \text{ and } 1 \text{ and } 1) = P(1) \cdot P(1) \cdot P(1) \cdot P(1) \cdot P(1)$

or

$= P(1)^5$

$= \left(\frac{1}{6}\right)^5 = \frac{1}{7776}$ or .0001286

Yes, we could roll 1 die 5 times or 5 dice 1 time.

Drawing Chips

1. Imagine you have the following chips in a bag: 3 black, 4 red, 2 white, 1 green.

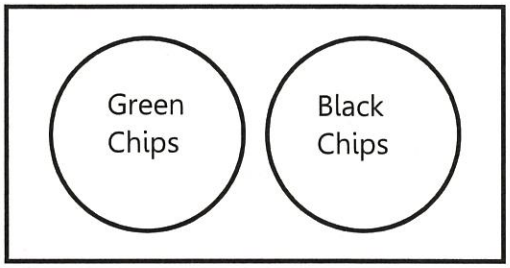
Let B be the event of drawing a black chip
Let R be the event of drawing a red chip
Let W be the event of drawing a white chip
Let G be the event of drawing a green chip

2. Write the theoretical probability of each event.

P(B) = $\frac{3}{10}$ P(R) = $\frac{4}{10} = \frac{2}{5}$ P(W) = $\frac{2}{10} = \frac{1}{5}$ P(G) = $\frac{1}{10}$

"OR" Probabilites

Disjoint Events cannot occur at the same time or share no common outcomes (a chip cannot be green and black at the same time). They are **mutually exclusive**.



If A and B are disjoint events, add the probabilities.

P(A or B) = P(A) + P(B)

3. What is the probability of drawing a black or green chip?

P(B or G) = P(B) + P(G) = $\frac{3}{10} + \frac{1}{10} = \frac{4}{10} = \frac{2}{5}$

4. What is the probability of drawing a black or red chip?

$P(B \text{ or } R) = \frac{3}{10} + \frac{4}{10} = \frac{7}{10}$

5. What is the probability of not drawing a white chip?

$P(W^c) = 1 - P(W)$

$= 1 - \frac{1}{5} = \frac{5}{5} - \frac{1}{5} = \frac{4}{5}$

6. What is the probability of not drawing a green or black chip?

$P(\text{Red or white}) = \frac{4}{10} + \frac{2}{10} = \frac{6}{10} = \frac{3}{5}$

or $1 - P(G \text{ or } B)$

$1 - (\frac{3}{10} + \frac{1}{10})$

$1 - \frac{4}{10}$

$= \frac{6}{10}$ or $\frac{3}{5}$

7. What is the probability of not drawing a red, white, or green chip?

$P(B) = \frac{3}{10}$

Successive Trials

Drawing with replacement.

Drawing **with replacement** means you would put the chip back each time before you draw another. Each draw is an independent event because you will be starting with the same chips each time.

Two events are called **independent** if knowing that one occurs does not change the probability that the other occurs. When you put the chips back each time the draws are independent.

For independent events, P(A and B) = P(A)·P(B)

8. For a series of 3 draws <u>with replacement</u>, what is the probability of getting

a. Three red chips?

P(3 red chips) = P(R)·P(R)·P(R) = $\frac{2}{5} \cdot \frac{2}{5} \cdot \frac{2}{5} = \frac{8}{125} = .064$

b. Three black chips?

$P(3 \text{ black}) = P(B) \cdot P(B) \cdot P(B)$

$= \frac{3}{10} \cdot \frac{3}{10} \cdot \frac{3}{10} = \frac{27}{1000} = .027$

Drawing Without Replacement

A probability is **conditional** if it depends on knowing what has already happened. If you draw the chips without putting them back you are drawing **without replacement.** Now successive draws are not independent because the chips in the bag are different each time.

9. For a series of 3 draws <u>without replacement</u>, what is the probability of getting

a. Three red chips

$\frac{4}{10} \cdot \frac{3}{9} \cdot \frac{2}{8} = \frac{24}{720} = .0333$

conditional on what happened before

b. Three black chips

$\frac{3}{10} \cdot \frac{2}{9} \cdot \frac{1}{8} = \frac{6}{720} = .0083$

conditional probabilities

Survey Example. According to Gallup.com in 2017, the percentage of smokers in the U.S. who smoke a pack or more of cigarettes a day is at an all-time low of 26%. If we take a survey of people in the U.S. and ask them whether they smoke a pack or more a day, would we be sampling with or without replacement? How would this change our calculation of probabilities?

We are sampling without replacement because we would not poll the same person twice. However, the population is very large so the difference is small.

The 10% Condition

When we sample less than 10% of the population, the probabilities with and without replacement are very close. So, if we sample less than 10% of the population, we can use our independence rule and multiply the probabilities.

Concept Summary

The **complement** of any event A is the event that A does not occur. It's denoted by A^C.

Two events are called **disjoint** if they have no outcomes in common and can never happen together.

Two events are called **independent** if knowing that one occurs does not change the probability that the other occurs.

A **conditional probability** is used when events are not independent. P(A|B) means the probability of A given B.

When sampling without replacement, make sure to sample **less than 10%** of the population to assume independence.

Probability Rules

1. $0 \leq P(A) \leq 1$, where P(sample space) = 1

2. $P(\text{not } A) = P(A^C) = 1 - P(A)$

3. P(A or B) = P(A) + P(B) if A and B are disjoint

4. P(A and B) = P(A) · P(B) if A and B are independent

Practice Problems

$9+4+2+1+2 = 18$ pairs

1. In your sock drawer you have 9 pairs of black socks, 4 brown, 2 red, 1 white and 2 blue.

a. If you draw at random, what is the probability of drawing a red pair of socks?

$P(\text{red}) = \frac{2}{18} = \frac{1}{9} \approx .1111$

b. If you draw at random three days in a row, what is the probability of wearing brown socks every time? w/o replacement – unless you do laundry every day

$P(3 \text{ brown}) = \frac{4}{18} \cdot \frac{3}{17} \cdot \frac{2}{16} = \frac{24}{4896} \approx .0049$

c. In four days with random draws, what is the probability of drawing three black pairs and then a blue pair? $P(3 \text{ black, then brown}) = \frac{9}{18} \cdot \frac{8}{17} \cdot \frac{7}{16} \cdot \frac{2}{15} = \frac{1008}{73440} \approx .0137$

2. Roughly 20% of undergraduates at a university are vegetarian or vegan. What is the probability that, among a random sample of 3 undergraduates, all three are vegetarian or vegan?

a. All three are vegetarian or vegan?

$P(3\ V's) = .20(.20)(.20) = .20^3 = .008$

b. None of the students sampled are vegetarian or vegan?

$P(3 \text{ not } V) = P(\text{not } V) \cdot P(\text{not } V) \cdot P(\text{not } V) = .80^3 = .512$

c. What condition allows us to calculate these probabilities?

The 10% condition – we are sampling without replacement so the trials are not independent. Since we are sampling less than 10% of the population, the difference is small.

3. Suppose that 40% of cars in your area are manufactured in the United States, 30% in Japan, 10% in Germany, and 20% in other countries. If cars are selected at random, find the probability that:

$P(US) = .40$ $P(J) = .30$ $P(G) = .10$

$P(O) = .20$

a. A car is not U.S.-made.

$P(\text{not } US) = 1 - P(US)$

$= 1 - .40 = .60$

b. A car is made in Japan or Germany.

$P(J \text{ or } G) = P(J) + P(G)$

$= .30 + .10 = .40$

c. You select two cars in a row from Japan.

$P(J \text{ and } J) = P(J) \cdot P(J)$ (10% rule)

$= (.30)(.30) = .09$

d. None of three cars came from Germany.

$P(\text{not } G \text{ and not } G \text{ and not } G)$

$= P(G^c) \cdot P(G^c) \cdot P(G^c)$

$= [1 - P(G)]^3$

$= (1 - .10)^3$

$= (.90)^3$

$= .729$

or $(.90)(.90)(.90) = .729$

Answers: 1a. ~~2/14~~ 2/18 or ~~1/7~~ 1/9, b. ~~0.0110~~ .0049, c. ~~0.0050~~ .0137
2a. 0.008, b. 0.512, c. 10% Condition
3a. 0.6, b. 0.4, c. 0.09, d. 0.729