

Probability Fundamentals: Solutions

By Jim Frost ©2026

This file contains the worked solutions for all the practice problems in my book, *Probability Fundamentals: An Intuitive Guide to Understanding Uncertainty*.

Tip: In your PDF viewer, use the bookmarks to jump to the specific chapter and question number for you which you need the solution.

Chapter 1

Question 1

A probability of 0.85 means that in many similar situations, the flight would arrive on time about 85 out of 100 times. It does not guarantee that a specific flight will be on time. Probability describes long-run frequency rather than predicting a single outcome with certainty.

Question 2

Note that the total number of students (20) fixes the denominator for all the following probabilities. The definition of an even determines the value in the numerator.

Total students = 20

- A. $P(\text{home}) = 9/20 = 0.45$
- B. $P(\text{library}) = 7/20 = 0.35$
- C. $P(\text{library or café}) = 11/20 = 0.55$ (Library or café = $7 + 4 = 11$)

Question 3

Having two categories does not mean the probabilities are equal. The probability depends on how many people fall into each category. In reality, far fewer people are left-handed than right-handed, so the probability of selecting a left-handed person is much less than 0.5.

Question 4

- A. **Model-based probability:** The probability is produced by a statistical model that combines multiple variables.
- B. **Empirical probability:** The probability is estimated from observed past data.
- C. **Theoretical probability:** The probability is derived using logical reasoning about numbers.

Question 5

Random processes naturally produce variability in small samples. Even when outcomes are equally likely, short sequences can appear uneven. A few early results do not determine whether the generator is fair. Only after many repetitions would the proportions begin to stabilize near the expected probabilities.

Question 6

Consider selecting a student at random from a class and recording their major.

- A. Experiment:** The experiment is the process that produces an outcome. Here, the experiment is randomly selecting a student from the class and observing their major.
- B. Trial:** A trial is a single repetition of the experiment. In this case, one random selection of a student is a single trial.
- C. Outcome:** An outcome is the specific result of a trial. For example, the selected student's major might be Statistics, Biology, Economics, Psychology, etc.
- D. Event:** An event is the outcome or set of outcomes that we are interested in. It can be based on the study's goals. For example, for a study assessing the prevalence of science majors, an event would be selecting a student who majors in a science field, which might include outcomes such as Biology, Chemistry, or Physics.

Chapter 2

Question 1

- A. Conditional
- B. Joint
- C. Marginal

Question 2

The joint probability that a person prefers Netflix and watches mostly on a TV is

$$P(\text{Netflix} \cap \text{TV}) = 54 / 150 = 0.36$$

Question 3

The marginal probability that a person watches mostly on a mobile device is

$$P(\text{Mobile}) = 72 / 150 = 0.48$$

Question 4

The conditional probability that a person watches mostly on a TV given that they prefer Netflix is

$$P(\text{TV} | \text{Netflix}) = 54 / 90 = 0.60$$

Question 5

The conditional probability that a person watches mostly on a TV given that they prefer another streaming service is

$$P(\text{TV} | \text{Other Service}) = 24 / 60 = 0.40$$

Question 6

If the variables were independent, the conditional probability of watching mostly on a TV would be the same for both groups.

$$P(\text{TV} | \text{Netflix}) = 0.60$$

$$P(\text{TV} | \text{Other Service}) = 0.40$$

Because these probabilities differ, preferred streaming service and device used do not appear to be independent. Instead, they appear to be related.

Question 7

First compute the marginal probability that a person prefers Netflix.

$$P(\text{Netflix}) = 90 / 150 = 0.60$$

Next compute the conditional probability that a person prefers Netflix given that they watch mostly on a TV.

$$P(\text{Netflix} | \text{TV}) = 54 / 78 = 0.692$$

Comparing the two probabilities:

$$P(\text{Netflix}) = 0.60$$

$$P(\text{Netflix} | \text{TV}) = 0.692$$

The conditional probability is larger in this case.

These probabilities can differ because conditional probabilities focus on a specific subgroup rather than the entire sample. The marginal probability $P(\text{Netflix})$ describes the proportion of all surveyed people who prefer Netflix. The conditional probability $P(\text{Netflix} | \text{TV})$ describes the proportion of Netflix users only among those who mostly watch on a TV.

If the preference for Netflix were unrelated to the device used for watching, the conditional probability would be about the same as the marginal probability. Because the probabilities differ here, it suggests that preferred streaming service and viewing device might be related.

Chapter 3

Question 1

- A. $P(A)$ Marginal probability. It measures the probability that event A occurs without considering any other event.
- B. $P(A \cap B)$ Joint probability. It represents the probability that both events A and B occur together.
- C. $P(A | B)$ Conditional probability. It measures the probability that event A occurs given that event B has already occurred.
- D. $P(A \cup B)$ Union probability. It represents the probability that event A occurs, event B occurs, or both occur.
- E. $P(B^c)$ Complement probability. It represents the probability that event B does not occur.

Question 2

The statement is: $P(A | B) = P(B | A)$

It is sometimes true but not always true.

Conditional probabilities depend on which event forms the restricted sample space.

$P(A | B)$ means we look only at the outcomes where B occurred and ask how often A occurred within that group.

$P(B | A)$ means we instead restrict the sample space to the outcomes where A occurred.

Because these two restricted sample spaces are usually different, the two probabilities are generally not equal.

Question 3

We are given:

$$A = \{1, 2, 4, 5, 7, 9, 11\}$$

$$B = \{2, 3, 5, 6, 9, 10\}$$

A. Find $A \cap B$

The intersection contains outcomes that appear in both sets.

$$A \cap B = \{2, 5, 9\}$$

B. Find $A \cup B$

The union contains all outcomes appearing in either set.

$$A \cup B = \{1, 2, 3, 4, 5, 6, 7, 9, 10, 11\}$$

C. How many students participate in at least one activity?

The number of outcomes in the union is:

10 students

D. How many students participate in neither activity?

The sample space contains 12 students.

$$\text{Students in neither activity} = 12 - 10 = 2$$

Question 4

Event A = product passes inspection

Total products = 20

Products that pass = 16

A. Find $P(A)$

$$P(A) = 16 / 20 = 0.80$$

B. Find $P(A^c)$

The complement represents products that fail inspection.

$$\text{Number that fail} = 20 - 16 = 4$$

$$P(A^c) = 4 / 20 = 0.20$$

C. Interpretation

Event A^c represents the probability that a randomly selected product fails inspection.

Question 5

Given:

$$A = \{2, 4, 5, 8, 11, 13\}$$

$$B = \{1, 4, 6, 8, 9, 12\}$$

A. Identify $A \cap B$

$$A \cap B = \{4, 8\}$$

B. Find $P(A \cap B)$

Total employees = 15

Employees in the intersection = 2

$$P(A \cap B) = 2 / 15 \approx 0.133$$

C. Interpretation

There is about a 13.3% joint probability that a randomly selected employee both works remotely and is in the marketing department.

Question 6

Compare the two statements. Which is more likely?

A. Linda is a bank teller

B. Linda is a bank teller and active in the feminist movement

Let

A = Linda is a bank teller

B = Linda is active in the feminist movement

Statement B represents the joint event:

$$A \cap B$$

According to the properties of intersections:

$$P(A \cap B) \leq P(A)$$

The overlap between two events can never be more likely than one of the events alone.

Therefore, statement A must be at least as likely as statement B.

Many people incorrectly choose option B. This error is known as the Conjunction Fallacy. It occurs because our brains tend to favor stories that seem more detailed or representative. The description of Linda fits many people's stereotype of someone active in the feminist movement, so the combined description feels more believable.

However, probability follows strict logical rules. Because statement B is more specific, it must represent a subset of event A, and therefore cannot be more probable.

Question 7

The student calculated:

$$P(A | B) = 1.25$$

This result is not possible.

Conditional probabilities must always lie between 0 and 1.

This follows from the conditional probability formula:

$$P(A | B) = P(A \cap B) / P(B)$$

Because the joint probability cannot exceed $P(B)$,

$$P(A \cap B) \leq P(B)$$

Therefore:

$$0 \leq P(A | B) \leq 1$$

A value greater than 1 indicates that the calculation or interpretation of the probabilities is incorrect.

Question 8

A small online store tracks purchases for 40 customers. (Recalculate)

Given:

$$P(A \cap B) = 0.15$$

$$P(B) = 0.20$$

$$P(A) = 0.70$$

A. Calculate $P(A | B)$

$$P(A | B) = P(A \cap B) / P(B)$$

$$P(A | B) = 0.15 / 0.20$$

$$P(A | B) = 0.75$$

B. Compare $P(A)$ and $P(A | B)$

$$P(A) = 0.70$$

$$P(A | B) = 0.75$$

C. Interpretation

Among birds, 75% are active during the day. This is slightly higher than the overall probability of daytime activity in the entire animal population (70%). This indicates that birds in this study are somewhat more likely to be active during the day than animals in general.

Question 9

A transportation survey records whether commuters drive to work and whether they live in the city center.

A: Calculate $P(A)$

Use the partitioning method described in the chapter tip. Event A can be divided into two non-overlapping parts:

$A \cap B$ (city-center commuters who drive)

$A \cap B^c$ (non-city-center commuters who drive)

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

$$P(A) = 0.15 + 0.35$$

$$P(A) = 0.50$$

So the probability that a randomly selected commuter drives to work is 0.50 (50%).

B: Calculate $P(A | B)$

Use the conditional probability formula:

$$P(A | B) = P(A \cap B) / P(B)$$

$$P(A | B) = 0.15 / 0.20$$

$$P(A | B) = 0.75$$

Among commuters who live in the city center, 75% drive to work.

C: Interpretation

The overall probability that a commuter drives to work is **50%**. However, among commuters who live in the city center, the probability of driving is **75%**.

Looking only at the marginal probability $P(A)$ would suggest that driving occurs about half the time. The conditional probability shows that driving is actually much more common among city-center residents.

This example illustrates how conditional probabilities reveal patterns within specific groups that may not be visible from the overall marginal probability alone.

Question 10

Boy Born on Tuesday Variant of the Two-Child Problem

We are finding:

$P(\text{both children are boys} \mid \text{at least one child is a boy born on Tuesday})$

The wording matters. The condition is **not** “one specific child is a boy born on Tuesday.” Instead, we are told that **at least one** of the two children satisfies that description. That changes the sample space.

Process:

- Each child has 14 possible sex/day combinations: 2 sexes \times 7 days.
- For two children, there are $14 \times 14 = 196$ ordered outcomes.
- The condition is: at least one child is a boy born on Tuesday.
- There are 27 outcomes that satisfy that condition.
- Of those 27 outcomes, 13 have two boys.

Therefore:

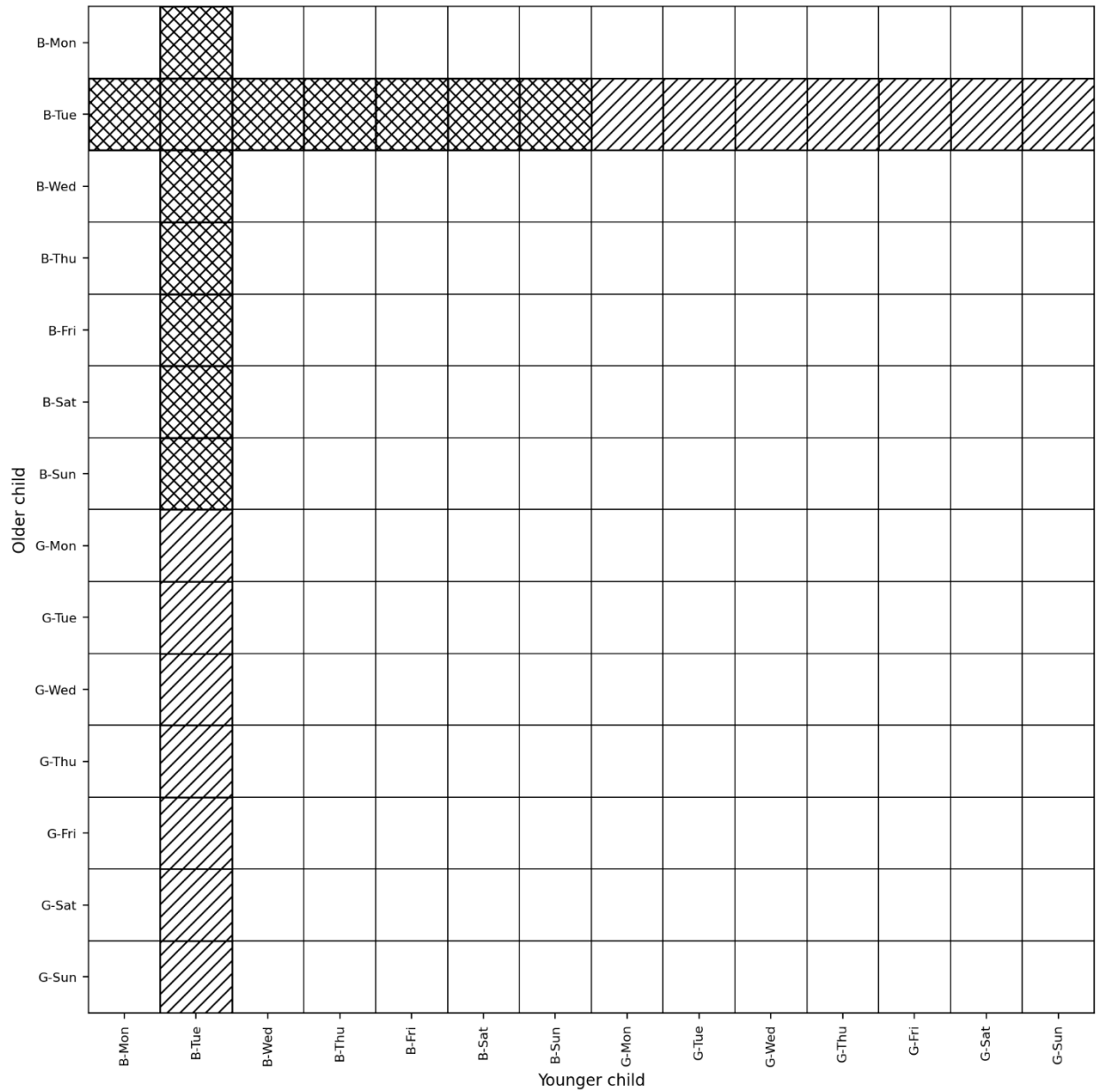
$P(\text{both children are boys} \mid \text{at least one child is a boy born on Tuesday}) = 13 / 27$



So the probability is:

$13 / 27 \approx 0.481$

The answer is not $1/2$ and not $1/3$. The extra wording about Tuesday changes which outcomes remain in the restricted sample space.

Sample Space for the Tuesday-Boy Variant
 196 ordered outcomes; 27 match the condition; 13 are favorable



 At least one child is B-Tue
 Both children are boys and at least one is B-Tue

Chapter 4

Question 1

A university survey found the following probabilities among students:

We are given:

- $P(A) = 0.55$
- $P(B) = 0.40$
- $P(A \cap B) = 0.25$

We want the probability that a student prefers online classes **or** morning classes, which is the union: $P(A \cup B)$.

Use the additive rule:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Substitute the values:

$$P(A \cup B) = 0.55 + 0.40 - 0.25 = 0.70$$

Answer: 0.70

Question 2

A fair coin is flipped twice. We want the probability of getting heads on both flips.

Let:

- H_1 = heads on the first flip
- H_2 = heads on the second flip

Because coin flips are independent:

$$P(H_1 \cap H_2) = P(H_1) \times P(H_2)$$

Each flip has probability $1/2$ of landing heads, so:

$$P(H_1 \cap H_2) = 1/2 \times 1/2 = 1/4$$

Answer: $1/4$ or 0.25.

Question 3

A jar contains 5 red marbles and 7 blue marbles, for a total of 12 marbles. Two marbles are drawn without replacement. We want the probability that both are red.

Let:

- R1 = first marble is red
- R2 = second marble is red

Because the first marble is not replaced, the events are dependent. Use the general multiplication rule:

$$P(R1 \cap R2) = P(R1) \times P(R2 | R1)$$

The probability that the first marble is red is:

$$P(R1) = 5/12$$

If the first marble is red, then 4 red marbles remain out of 11 total marbles:

$$P(R2 | R1) = 4/11$$

Now multiply:

$$P(R1 \cap R2) = 5/12 \times 4/11 = 20/132 = 5/33$$

Answer: 5/33 or about 0.152.

Question 4

What is the joint probability that a randomly selected employee completes the training, reports improved productivity, **and** recommends the training to others?

We are given:

- 70% of employees complete the training.
- Among those who complete the training, 80% report improved productivity.
- Among employees who report improved productivity, 90% recommend the training to others.

Because this is an “and” question, we must find the joint probability: P(training \cap productivity \cap recommend)

Let:

- A = completes the training
- B = reports improved productivity
- C = recommends the training

Notice that the wording “among those who” indicates conditional probabilities. For example, “among those who complete the training” means we are evaluating productivity only within that subgroup, which corresponds to the conditional probability $P(B | A)$.

Use the multiplication rule for three events:

$$P(A \cap B \cap C) = P(A) \times P(B | A) \times P(C | A \cap B)$$

Substitute the values:

$$P(A \cap B \cap C) = 0.70 \times 0.80 \times 0.90 = 0.504$$

Answer: 0.504

This means that 50.4% of employees are expected to complete the training, report improved productivity, and recommend it to others.

Question 5

A study examines two events:

A = A person owns a bicycle

B = A person drinks coffee daily

We are given:

- $P(A) = 0.35$
- $P(B) = 0.60$
- $P(A \cap B) = 0.21$

We want to determine whether A and B are independent or dependent.

For independent events:

$$P(A \cap B) = P(A) \times P(B)$$

Calculate:

$$P(A) \times P(B) = 0.35 \times 0.60 = 0.21$$

This matches the given joint probability:

$$P(A \cap B) = 0.21$$

Because the actual joint probability equals the product of the marginal probabilities, the events are independent.

Answer: Independent. Learning whether a person owns a bicycle does not change the probability that they drink coffee daily.

Question 6

Two events A and B have the following probabilities:

- $P(A) = 0.30$
- $P(B) = 0.40$
- $P(A \cap B) = 0$

We want to determine whether the events are independent, mutually exclusive, both, or neither.

Because: $P(A \cap B) = 0$

the events cannot occur together. That means they are mutually exclusive.

Now check whether they are independent. If A and B were independent, then:

$$P(A \cap B) = P(A) \times P(B)$$

$$\text{Calculate: } P(A) \times P(B) = 0.30 \times 0.40 = 0.12$$

But the given joint probability is:

$$P(A \cap B) = 0$$

Because:

$$0 \neq 0.12$$

the events are **not independent**.

Answer: Mutually exclusive and dependent.

These events are mutually exclusive because they cannot happen at the same time. They are not independent because the joint probability would have to equal the product of their individual probabilities.

Question 7

Repeated Independent Events

Each email has a 2% probability of being a phishing email, so each email has a 98% probability of not being a phishing email.

We want:

P(at least one phishing email)

Use the complement rule:

$$P(\text{at least one phishing email}) = 1 - P(\text{no phishing emails})$$

Because the emails are processed independently, use the multiplication rule for independent events:

$$P(\text{no phishing emails}) = 0.98^{50}$$

Therefore:

$$\begin{aligned} P(\text{at least one phishing email}) &= 1 - 0.98^{50} \\ &= 1 - 0.3642 \\ &= 0.6358 \end{aligned}$$

Answer: 0.6358

So, the probability that at least one phishing email is detected among 50 emails is about 63.6%.

Question 8

Repeated Dependent Events

The shipment contains 6 hardcover books and 10 paperback books, for a total of 16 books. Four books are selected without replacement.

We want:

P(all 4 selected books are hardcover)

Because the books are selected without replacement, the probabilities change after each selection. Use the general multiplication rule:

$$\begin{aligned} &P(H_1 \cap H_2 \cap H_3 \cap H_4) \\ &= P(H_1) \times P(H_2 \mid H_1) \times P(H_3 \mid H_1 \cap H_2) \times P(H_4 \mid H_1 \cap H_2 \cap H_3) \end{aligned}$$

Now substitute the probabilities:

$$\begin{aligned} &P(H_1 \cap H_2 \cap H_3 \cap H_4) \\ &= (6/16) \times (5/15) \times (4/14) \times (3/13) \\ &= 360/43680 \\ &= 3/364 \\ &\approx 0.0082 \end{aligned}$$

Answer: $3/364$, or about 0.0082

So, the probability that all 4 selected books are hardcover is about 0.82%.

Chapter 5

Question 1

Converting Between Probability and Odds

A manufacturing defect occurs in 8% of products.

A. $p = 0.08$

$$\text{odds} = p / (1 - p) = 0.08 / 0.92 = 0.087$$

B. The odds are approximately 0.087, meaning about 1 defective product for every $1 / 0.087 = 11.5$ non-defective products.

C. $p = \text{odds} / (1 + \text{odds}) = 0.25 / 1.25 = 0.20$

The probability is 0.20 (20%).

Question 2

Interpreting Odds

A sports team has odds of winning equal to 2.

A. Odds = 2 means 2 wins for every 1 loss.

B. The team is more likely to win because the odds are greater than 1.

C. $p = 2 / (1 + 2) = 2 / 3 = 0.667$

The probability of winning is **66.7%**.

Question 3

Contingency Table and Risk Measures

A study examines whether exposure to a chemical is associated with a respiratory illness.

A. Absolute risks

Exposed:

$$AR_1 = 40 / (40 + 160) = 40 / 200 = 0.20$$

Not exposed:

$$AR_2 = 20 / (20 + 180) = 20 / 200 = 0.10$$

B. Relative risk

$$RR = AR_1 / AR_2 = 0.20 / 0.10 = 2.0$$

C. Odds

Exposed:

$$\text{odds}_1 = 40 / 160 = 0.25$$

Not exposed:

$$\text{odds}_2 = 20 / 180 = 0.111$$

D. Odds ratio

$$\text{OR} = 0.25 / 0.111 = 2.25$$

E. Interpretation

The exposed group has twice the probability of illness (RR = 2.0).

The odds of illness are about 2.25 times higher in the exposed group.

Exposure is associated with a higher risk of illness.

Question 4

Interpreting Reported Measures

A study reports that a treatment has:

- Relative risk = 0.75
- Odds ratio = 0.70

A. RR = 0.75 means the treatment group has 75% of the risk of the control group, or a 25% reduction in probability.

B. OR = 0.70 means the odds of the outcome are 30% lower in the treatment group.

C. Yes. Both indicate a protective effect. The values are similar, suggesting the event may be relatively rare.

Question 5

Case-Control Logic

Researchers conduct a case-control study of a disease and prior exposure to a pollutant.

They compare:

- the odds of exposure among cases
- the odds of exposure among controls

A. Because the number of cases and controls is fixed by design, the sample does not reflect the true proportion of diseased individuals in the population. Therefore, absolute risks and relative risk cannot be estimated.

B. The odds ratio compares the odds of prior exposure among cases to the odds of exposure among controls. It measures the association between exposure and disease.

Question 6

Same Odds Ratio, Different Risk

Two studies examine the effect of the same exposure on a disease. Both studies report an odds ratio of 2.0.

A. Probability of disease in the exposed group

Study 1 (Rare outcome)

Unexposed group:

$$p_0 = 1 / 100 = 0.01$$

$$\text{odds}_0 = p_0 / (1 - p_0) = 0.01 / 0.99 = 0.0101$$

Given OR = 2:

$$\text{odds}_1 = 2 \times 0.0101 = 0.0202$$

Convert to probability:

$$p_1 = \text{odds}_1 / (1 + \text{odds}_1) = 0.0202 / 1.0202 = 0.0198$$

Study 2 (Common outcome)

Unexposed group:

$$p_0 = 40 / 100 = 0.40$$

$$\text{odds}_0 = 0.40 / 0.60 = 0.667$$

Given OR = 2:

$$\text{odds}_1 = 2 \times 0.667 = 1.333$$

Convert to probability:

$$p_1 = 1.333 / (1 + 1.333) = 1.333 / 2.333 = 0.571$$

B. Relative risk for each study

Study 1

$$RR = p_1 / p_0 = 0.0198 / 0.01 = 1.98$$

Study 2

$$RR = p_1 / p_0 = 0.571 / 0.40 = 1.43$$

C. Comparison of relative risks

- Study 1: RR = 1.98
- Study 2: RR = 1.43

Even though both studies have the same odds ratio (2.0), the relative risks differ substantially.

D. Explanation

The odds ratio is based on odds, which change nonlinearly with probability. When the outcome is rare, odds and probabilities are very similar, so the odds ratio closely approximates the relative risk.

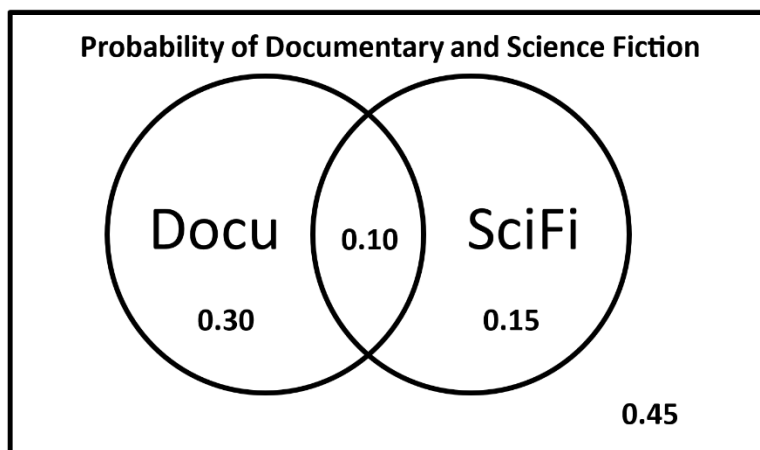
When the outcome is common, odds and probabilities differ more, causing the odds ratio to diverge from the relative risk.

As a result, the same odds ratio can correspond to different relative risks depending on the baseline probability in the unexposed group.

Chapter 6

Question 1

A streaming service tracks whether subscribers watch documentaries and whether they watch science fiction.



Overlap

$$\begin{aligned}P(D \cap S) &= P(D) \times P(S) \\ &= 0.40 \times 0.25 \\ &= 0.10\end{aligned}$$

Documentaries only

$$\begin{aligned}P(D \cap S^c) &= P(D) - P(D \cap S) \\ &= 0.40 - 0.10 \\ &= 0.30\end{aligned}$$

Science fiction only

$$\begin{aligned}P(S \cap D^c) &= P(S) - P(D \cap S) \\ &= 0.25 - 0.10 \\ &= 0.15\end{aligned}$$

Neither

$$\begin{aligned}P(D \cup S)^c &= 1 - [P(D) + P(S) - P(D \cap S)] \\ &= 1 - [0.40 + 0.25 - 0.10] \\ &= 1 - 0.55 \\ &= 0.45\end{aligned}$$

Final Venn Diagram Values

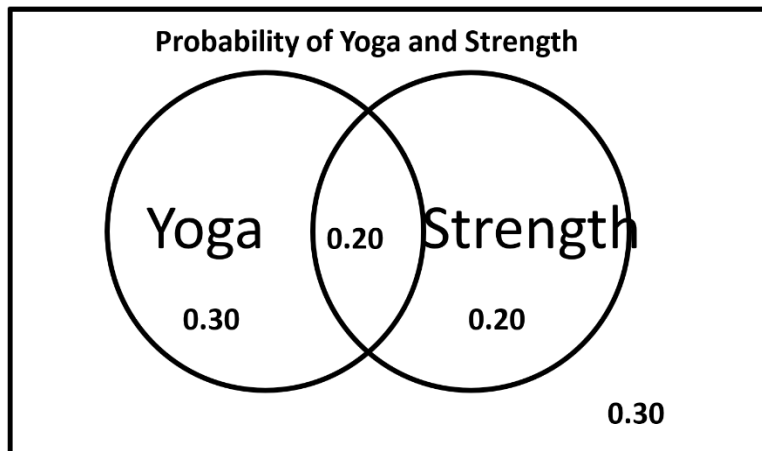
- $P(D \cap S) = 0.10$
- $P(D \cap S^c) = 0.30$
- $P(S \cap D^c) = 0.15$
- $P(D \cup S)^c = 0.45$

Check

$$0.10 + 0.30 + 0.15 + 0.45 = 1.00$$

Question 2

A fitness center tracks whether members attend yoga classes and whether they attend strength training classes.



When a problem gives you the marginal probabilities and the probability of the union, start by using the additive rule to find the joint probability.

$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$

After you have the overlap, fill in the remaining regions by subtracting the overlap from each marginal probability. Finally, find the probability of neither event by taking the complement of the union.

Overlap

$$\begin{aligned} P(Y \cap S) &= P(Y) + P(S) - P(Y \cup S) \\ &= 0.50 + 0.40 - 0.70 \\ &= 0.20 \end{aligned}$$

Yoga only

$$\begin{aligned} P(Y \cap S^c) &= P(Y) - P(Y \cap S) \\ &= 0.50 - 0.20 \\ &= 0.30 \end{aligned}$$

Strength only

$$\begin{aligned} P(S \cap Y^c) &= P(S) - P(Y \cap S) \\ &= 0.40 - 0.20 \\ &= 0.20 \end{aligned}$$

Neither

$$\begin{aligned} P((Y \cup S)^c) &= 1 - P(Y \cup S) \\ &= 1 - 0.70 \\ &= 0.30 \end{aligned}$$

Final Venn Diagram Values

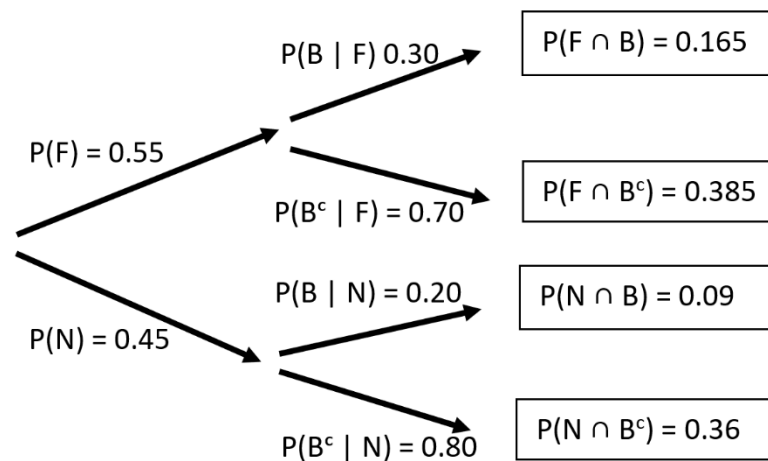
- $P(Y \cap S) = 0.20$
- $P(Y \cap S^c) = 0.30$
- $P(S \cap Y^c) = 0.20$
- $P(Y \cup S)^c = 0.30$

Check

$$0.20 + 0.30 + 0.20 + 0.30 = 1.00$$

Question 3

A bookstore tracks whether a customer buys a fiction book or a nonfiction book.



A. Independence or dependence

Because the bookmark probabilities differ by branch,

$$P(B | F) = 0.30$$

$$P(B | N) = 0.20$$

the events are dependent.

B. Probability of nonfiction and bookmark

$$P(N \cap B) = P(N) \times P(B | N)$$

$$= 0.45 \times 0.20$$

$$= 0.09$$

C. Overall probability of buying a bookmark

$$\begin{aligned}
 P(B) &= P(F \cap B) + P(N \cap B) \\
 &= 0.165 + 0.09 \\
 &= 0.255
 \end{aligned}$$

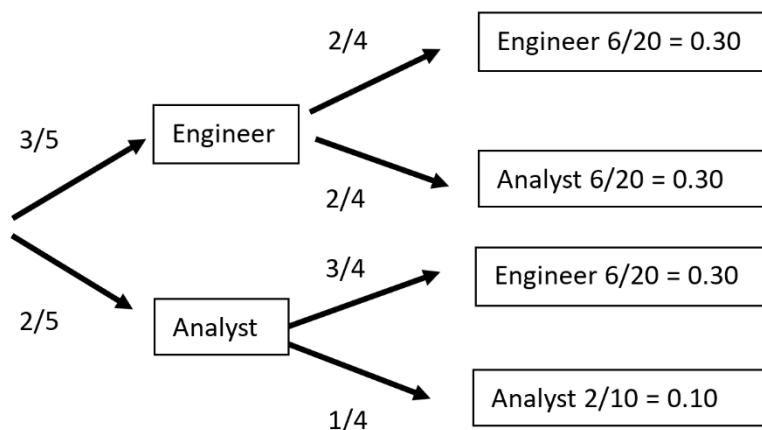
Question 4

The two employees are selected one after the other, without replacement. What is the probability that both selected employees are engineers? That both are analysts? And one of each?

To build the tree, start with the total number of employees and determine the probability for each outcome on the first selection. These probabilities are based on the initial counts. For example, engineers out of all employees for the upper branch and analysts out of all employees for the lower. For the second selection, adjust both the counts and the total because one employee (either an engineer or an analyst) has already been selected and is no longer available. This change creates conditional probabilities for the second stage.

Remember, the probability must apply to the outcome the arrow points towards. So, in some cases it's the probability of drawing an engineer while in other cases it's the probability of drawing an analyst.

After all branch probabilities are in place, find the probability for each terminal node by multiplying the probabilities along its path. Each terminal node represents one complete sequence of selections, and the product of the branch probabilities gives the joint probability for that outcome.



- A. There is a 0.30 probability of two engineers.
- B. There is a 0.10 probability of two analysts.
- C. There is a 0.60 probability of one engineer and one analyst.

Chapter 7

Question 1

You are asked to find the overall probability that a randomly selected user watches at least one show per day.

Use the Law of Total Probability. The problem describes different cases (subscription plans) and gives conditional probabilities within each case. It asks for an overall probability across all users. This structure matches combining probabilities across cases, which is exactly what LOTP does.

Question 2

A wildlife researcher studies three habitats where a bird species nests.

- A. The cases are H_1 , H_2 , and H_3 (the three habitat types).
- B. The outcome of interest is S (a nest successfully produces chicks).
- C. The conditional probabilities describe the probability that a nest successfully produces chicks given each habitat: $P(S | H_1)$, $P(S | H_2)$, $P(S | H_3)$.
- D. The Law of Total Probability expression is:
$$P(S) = P(H_1) * P(S | H_1) + P(H_2) * P(S | H_2) + P(H_3) * P(S | H_3)$$

Question 3

A company screens job applicants using an automated assessment designed to identify highly qualified candidates.

Even though the test is highly accurate at identifying both qualified and unqualified applicants, the overall probability can still be lower than expected because highly qualified candidates are rare. When the base rate is low, even a small false positive rate applied to the much larger group of unqualified applicants can produce more false positives than true positives. As a result, many applicants flagged as highly qualified may not actually be highly qualified.

Question 4

A university offers students three ways to contact academic advising.

A. Completed Probability Tree

First, find the missing probabilities needed for the tree.

Because the advising method probabilities must sum to 1:

$$P(M_3) = 1 - P(M_1) - P(M_2)$$

$$P(M_3) = 1 - 0.30 - 0.45 = 0.25$$

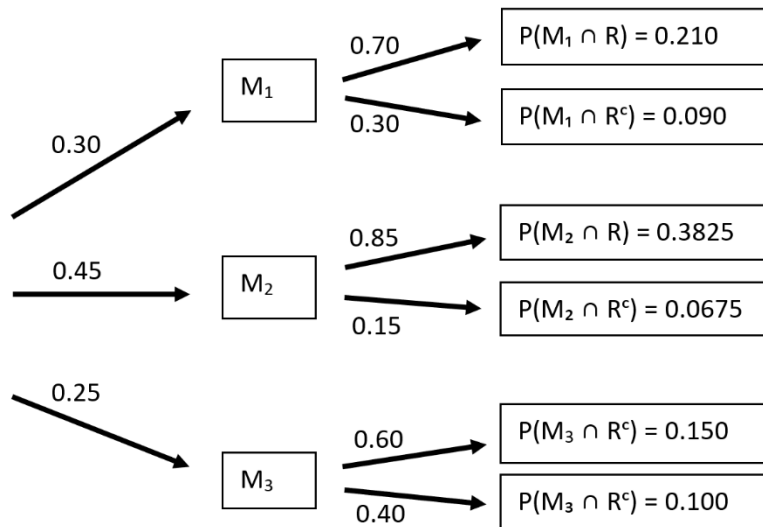
Because resolved and not resolved are complements within each advising method:

$$P(R^c \mid M_1) = 1 - 0.70 = 0.30$$

$$P(R^c \mid M_2) = 1 - 0.85 = 0.15$$

$$P(R^c \mid M_3) = 1 - 0.60 = 0.40$$

So, the completed tree uses:



B. Law of Total Probability

Use LOTP to add the contributions from all advising methods that can lead to a resolved issue.

- $P(R) = P(M_1) * P(R \mid M_1) + P(M_2) * P(R \mid M_2) + P(M_3) * P(R \mid M_3)$
- $= (0.30 * 0.70) + (0.45 * 0.85) + (0.25 * 0.60)$
- $= 0.210 + 0.3825 + 0.150$
- $= 0.7425$

So, the overall probability that a student's issue is resolved is 0.7425, or 74.25%.

C. Bayes' Theorem

Now find the probability for each advising method given that the issue was resolved. You can reuse the previous LOTP value (0.7425) in all three denominators.

Drop-in advising:

- $P(M_1 \mid R) = P(M_1 \cap R) / P(R)$
- $= P(M_1) * P(R \mid M_1) / [P(M_1) * P(R \mid M_1) + P(M_2) * P(R \mid M_2) + P(M_3) * P(R \mid M_3)]$
- $= (0.30 * 0.70) / [(0.30 * 0.70) + (0.45 * 0.85) + (0.25 * 0.60)]$
- $= 0.210 / (0.210 + 0.3825 + 0.150)$

- $= 0.210 / 0.7425$

- $= 0.283$

Scheduled appointment:

- $P(M_2 | R) = P(M_2 \cap R) / P(R)$

- $= P(M_2) * P(R | M_2) / [P(M_1) * P(R | M_1) + P(M_2) * P(R | M_2) + P(M_3) * P(R | M_3)]$

- $= (0.45 * 0.85) / [(0.30 * 0.70) + (0.45 * 0.85) + (0.25 * 0.60)]$

- $= 0.3825 / (0.210 + 0.3825 + 0.150)$

- $= 0.3825 / 0.7425$

- $= 0.515$

Email advising:

- $P(M_3 | R) = P(M_3 \cap R) / P(R)$

- $= P(M_3) * P(R | M_3) / [P(M_1) * P(R | M_1) + P(M_2) * P(R | M_2) + P(M_3) * P(R | M_3)]$

- $= (0.25 * 0.60) / [(0.30 * 0.70) + (0.45 * 0.85) + (0.25 * 0.60)]$

- $= 0.150 / (0.210 + 0.3825 + 0.150)$

- $= 0.150 / 0.7425$

- $= 0.202$

Final Answer

Given that a student's issue is resolved:

- $P(M_1 | R) = 0.283$

- $P(M_2 | R) = 0.515$

- $P(M_3 | R) = 0.202$

So, the most likely advising method is scheduled appointment (M_2) because it has the highest conditional probability.

Question 5

A security system scans files for malware.

Let:

- M = file contains malware

- F = file is flagged

Use Bayes' Theorem to find the probability that the file contains malware given that the file was flagged: $P(M | F)$.

- $P(M | F) = P(M \cap F) / P(F)$

- $= P(M) * P(F | M) / [(P(M) * P(F | M) + P(M^c) * P(F | M^c))]$
- $= (0.01 * 0.98) / [(0.01 * 0.98) + (0.99 * 0.03)]$
- $= 0.0098 / (0.0098 + 0.0297)$
- $= 0.0098 / 0.0395$
- $= 0.248$

So, the probability that a flagged file actually contains malware is 0.248, or 24.8%.

Chapter 8

Question 1

Identifying Priors, Likelihoods, and Posteriors

A cybersecurity system monitors login attempts.

A. Identify the prior probability.

The prior probability is the probability that a login attempt is malicious before observing whether it was flagged.

Prior: $P(\text{Malicious}) = 0.002$

B. Identify the likelihood(s).

The likelihoods are the probabilities of the observed evidence under each possible explanation.

Because the observed evidence is that the login attempt was flagged, the relevant likelihoods are:

$P(\text{Flagged} \mid \text{Malicious}) = 0.90$

$P(\text{Flagged} \mid \text{Legitimate}) = 0.08$

These tell us how expected a flag is under each explanation.

C. State the posterior probability in words.

The posterior probability is the updated probability that the login attempt is malicious after observing that it was flagged.

D. Explain, in words, what the likelihood tells you in this context.

The likelihood tells you how well each explanation accounts for the evidence. Here, it tells you how expected a flag is if the login attempt is malicious and how expected a flag is if the login attempt is legitimate. Because malicious attempts are flagged much more often than legitimate ones, the evidence is more consistent with the malicious-login explanation.

Question 2

What Changes After Evidence?

A company manufactures products using three suppliers. Supplier A provides most parts but has a low defect rate. Supplier C provides fewer parts but has a higher defect rate.

Before observing the defect, Supplier A would likely seem the most plausible source because it supplies the most parts. That is the prior idea: more parts from Supplier A means a randomly chosen part is more likely to have come from A.

After observing that the product is defective, the belief should shift. The defect is the new evidence, and Supplier C's higher defect rate means that this evidence is more consistent with Supplier C than with Supplier A. In Bayesian terms, Supplier A can still have a strong prior because it supplies most parts, but Supplier C has a stronger likelihood because defective parts are more expected from Supplier C.

The posterior combines both pieces of information. The conclusion depends on balancing how common each supplier is to begin with and how strongly each supplier predicts the observed defect. So, observing the defect should increase the probability that Supplier C produced the part, even if Supplier A still provides most parts overall.

Question 3

Explaining Why a Common Answer Is Wrong

A rare software bug occurs in 1 out of every 5,000 programs. A diagnostic tool correctly detects the bug 95% of the time when it is present and incorrectly flags 4% of programs that do not have the bug.

A program is flagged by the tool.

A developer says:

"The tool is highly accurate, so there's about a 95% chance this program has the bug."

A. Explain why this conclusion is incorrect.

This conclusion is incorrect because 95% is not the probability that a flagged program has the bug. Instead, 95% is the probability that the tool flags a program when the bug is actually present.

That is a forward conditional probability:

$$P(\text{Flagged} \mid \text{Bug}) = 0.95$$

But the question the developer cares about is the reverse conditional probability:

$$P(\text{Bug} \mid \text{Flagged})$$

These are not the same.

B. Identify the mistake in terms of Bayesian reasoning.

The mistake is that the developer ignored the prior probability and confused the likelihood with the posterior.

The bug is very rare, occurring in only 1 out of 5,000 programs. That tiny prior matters a great deal. Even though the tool is good at detecting real bugs, the 4% false positive rate will produce many false alarms because almost all programs do not have the bug.

This is a form of base rate neglect. The developer focused on the test's accuracy and ignored how rare the bug is to begin with.

C. Without calculating, explain whether the true probability is likely to be high, moderate, or low, and why.

The true probability is likely to be low.

The prior probability is extremely small, so the program is very unlikely to have the bug before seeing any flag. Although a real bug is likely to be flagged, there are so many bug-free programs that even a 4% false positive rate will create many flagged programs that do not actually have the bug. As a result, the posterior probability $P(\text{Bug} \mid \text{Flagged})$ is unlikely to be anywhere near 95%.

Question 4

Interpreting a Posterior Probability

A test result leads to the conclusion:

$$P(\text{Fraud} \mid \text{Flagged}) = 0.12$$

Explain what this probability means in plain language and the group it refers to.

This means that among flagged cases like this one, about 12% are actually fraudulent under the assumed conditions. It refers to the group of cases that were flagged. Among cases like this with a flag, about 12% would be expected to involve actual fraud, assuming the model and probabilities used are correct.

Question 5

A company screens job applications for potential fraud. Historically, about 1% of applications contain false information.

$$P(\text{Fraud} \mid \text{Flagged}) = 0.18$$

A. What was the prior probability in this scenario?

The prior probability is the historical probability of fraud before observing whether an application was flagged.

Prior: $P(\text{Fraud}) = 0.01$

B. What is the posterior probability?

The posterior probability is the updated probability of fraud after observing that the application was flagged.

Posterior: $P(\text{Fraud} \mid \text{Flagged}) = 0.18$

C. Describe how the belief changed from prior to posterior.

The belief changed from 0.01 to 0.18.

Before seeing the flag, the prior probability of fraud was only 1%. After seeing the flag, the updated or posterior probability rises to 18%. The application is still more likely to be legitimate than fraudulent, but the evidence has made fraud much more plausible than it was before. The posterior is the prior after being updated by that evidence.

D. In words, explain what caused this change. Your answer should refer to both the prior and the likelihood.

The change occurred because the flag is evidence that is more expected when fraud is present than when fraud is absent. In Bayesian terms, the prior started very low because fraud is rare, but the likelihood of seeing a flag under the fraud hypothesis was strong enough to raise the probability.

The posterior reflects both facts: fraud was rare to begin with, but the flag provided meaningful evidence in its favor.

Chapter 9

Question 1

- A. Permutations with repetition
- B. Combinations without repetition
- C. Combinations with repetition
- D. Permutations without repetition

Question 2

A streaming service generates a 3-letter recommendation code using the letters A–Z. Letters can repeat. We want to find the probability that a randomly generated code contains no repeated letters.

This problem involves permutations because the order of the letters matters. For example, ABC is a different code from BAC. It uses different types of repetition for the denominator and numerator:

- For the denominator, repetition is allowed because letters can repeat in any code.
- For the numerator, repetition is not allowed because we want codes with no repeated letters.

Denominator: Total number of possible codes

The denominator counts all possible 3-letter codes.

Because:

- order matters, and
- repetition is allowed,

we use permutations with repetition.

The formula is:

$$n^r$$

where:

- n = number of possible outcomes for each position
- r = number of positions in the code

For this problem:

- $n = 26$ because there are 26 letters

- $r = 3$ because the code has 3 positions

Substitute those values into the formula:

$$\begin{aligned}\text{Total number of possible codes} &= 26^3 \\ &= 26 \times 26 \times 26 \\ &= 17,576\end{aligned}$$

That value, 17,576, is the denominator in the probability ratio.

Numerator: Number of codes with no repeated letters

The numerator counts only the codes that contain no repeated letters.

Because:

- order still matters, and
- repetition is not allowed,

we use **permutations without repetition**.

The formula is:

$$nPr = n! / (n - r)!$$

where:

- n = number of available letters
- r = number of positions to fill

For this problem:

- $n = 26$
- $r = 3$

Substitute those values into the formula:

$$\begin{aligned}{}_{26}P_3 &= 26! / (26 - 3)! \\ &= 26! / 23! \\ &= 26 \times 25 \times 24 \\ &= 15,600\end{aligned}$$

That value, 15,600, is the numerator in the probability ratio.

Probability

Now divide the number of favorable codes by the total number of possible codes:

$$P(\text{no repeated letters}) = 15,600 / 17,576 \\ = 0.8873$$

So, the probability that a randomly generated 3-letter code has no repeated letters is about 0.8873, or 88.73%.

Question 3

A museum has 9 different paintings and randomly selects 4 of them for inclusion in a special exhibit.

We want to find the probability that the selected set of paintings includes both a specific Monet painting and a specific Van Gogh painting.

This problem involves combinations without repetition because:

- the question asks which paintings are selected for the exhibit
- each painting can be selected only once

We will count:

- the total number of possible 4-painting selections for the denominator
- the number of selections that include both specific paintings for the numerator

Denominator: Total number of possible selections

The denominator counts all possible ways to select 4 of the 9 paintings.

Because:

- only the selected group matters, and
- repetition is not allowed,

we use combinations without repetition.

The formula is:

$${}^n C_r = n! / [r!(n - r)!]$$

where:

- n = number of available items
- r = number of items selected

For this problem:

- $n = 9$ because there are 9 different paintings
- $r = 4$ because 4 paintings are selected

Substitute those values into the formula:

$$\begin{aligned} {}_9C_4 &= 9! / [4!(9 - 4)!] \\ &= 9! / (4!5!) \\ &= (9 \times 8 \times 7 \times 6) / (4 \times 3 \times 2 \times 1) \\ &= 126 \end{aligned}$$

That value, 126, is the denominator in the probability ratio.

Numerator: Number of selections that include both specific paintings

Now count the number of 4-painting selections that include both the specific Monet and the specific Van Gogh.

If those two paintings must be included, then 2 of the 4 exhibit spots are already taken. That means we need to choose the remaining 2 paintings from the 7 paintings that are left.

So, for the numerator:

- $n = 7$ because 7 other paintings remain after including the Monet and Van Gogh
- $r = 2$ because we still need to choose 2 more paintings

Substitute those values into the combinations formula:

$$\begin{aligned} {}_7C_2 &= 7! / [2!(7 - 2)!] \\ &= 7! / (2!5!) \\ &= (7 \times 6) / (2 \times 1) \\ &= 21 \end{aligned}$$

That value, 21, is the numerator in the probability ratio.

Probability

Now divide the number of favorable selections by the total number of possible selections:

$$\begin{aligned} P(\text{selected set includes both specific paintings}) \\ &= 21 / 126 \\ &= 1 / 6 \\ &= 0.1667 \end{aligned}$$

So, the probability is $1/6$, or about 16.67%.

Question 4

A smoothie shop offers 7 different fruits. A customer selects 3 fruits to include in a smoothie. The same fruit can be chosen more than once.

We want to find the probability that a randomly created smoothie includes at least one banana.

This problem involves combinations with repetition because:

- the question asks which fruits are included in the smoothie
- the same fruit can be chosen more than once

We will count:

- the total number of possible 3-fruit smoothies for the denominator
- the number of smoothies that include at least one banana for the numerator

Denominator: Total number of possible smoothies

The denominator counts all possible ways to choose 3 fruits from 7 options when repetition is allowed.

Because:

- only the selected fruits matter, and
- repetition is allowed,

we use combinations with repetition.

The formula is:

$$(n + r - 1)C_r = (n + r - 1)! / [r!(n - 1)!]$$

where:

- n = number of available fruit choices
- r = number of fruits selected

For this problem:

- $n = 7$ because there are 7 fruit options
- $r = 3$ because the smoothie contains 3 fruit selections

Substitute those values into the formula:

$$\begin{aligned}(7 + 3 - 1)C_3 &= 9C_3 \\ &= 9! / [3!(9 - 3)!] \\ &= 9! / (3!6!) \\ &= (9 \times 8 \times 7) / (3 \times 2 \times 1) \\ &= 84\end{aligned}$$

That value, 84, is the denominator in the probability ratio.

Numerator: Number of smoothies that include at least one banana

Now count the number of 3-fruit smoothies that include at least one banana.

This is easiest to do using a complement. Instead of counting smoothies with at least one banana directly, we count the smoothies with no banana and subtract from the total.

Smoothies with no banana means all 3 fruit selections come from the other 6 fruits.

Because:

- only the selected fruits matter, and
- repetition is still allowed,

we again use combinations with repetition.

For the no-banana smoothies:

- $n = 6$ because banana is excluded, leaving 6 fruit options
- $r = 3$ because the smoothie still contains 3 fruit selections

Substitute those values into the formula:

$$\begin{aligned} (6 + 3 - 1)C_3 &= 8C_3 \\ &= 8! / [3!(8 - 3)!] \\ &= 8! / (3!5!) \\ &= (8 \times 7 \times 6) / (3 \times 2 \times 1) \\ &= 56 \end{aligned}$$

So, the number of smoothies with at least one banana is:

$$\begin{aligned} \text{Smoothies with at least one banana} \\ &= \text{Total smoothies} - \text{Smoothies with no banana} \\ &= 84 - 56 \\ &= 28 \end{aligned}$$

That value, 28, is the numerator in the probability ratio.

Probability

Now divide the number of favorable smoothies by the total number of possible smoothies:

$$\begin{aligned} P(\text{at least one banana}) \\ &= 28 / 84 \\ &= 1 / 3 \\ &\approx 0.3333 \end{aligned}$$

So, the probability is $1/3$, or about 33.33%.

Question 5

A music app has a curated list of 14 songs: 8 upbeat tracks and 6 slow tracks.

First, the app randomly selects 5 songs to create a featured playlist. Then, from those 5 songs, it randomly assigns two distinct positions: Opening Track and Closing Track.

We want to find the probability that the playlist contains exactly 3 upbeat songs and 2 slow songs, and that a specific slow song is chosen as the Opening Track.

This problem has two stages, so it uses two counting tools:

- Stage 1 selects the 5 songs for the playlist. This uses combinations without repetition because the question is about which songs are selected, and a song cannot be selected more than once.
- Stage 2 assigns the distinct positions of Opening Track and Closing Track within the selected playlist. This uses permutations without repetition because the positions are different and the same song cannot fill both positions.

We will count:

- the total number of possible outcomes for the denominator
- the number of favorable outcomes for the numerator

Denominator: Total number of possible outcomes

First count the number of possible 5-song playlists.

Because:

- only the selected set of songs matters at this stage, and
- repetition is not allowed,

we use combinations without repetition.

The formula is:

$$nC_r = n! / [r!(n - r)!]$$

where:

- n = number of available songs
- r = number of songs selected

For this problem:

- $n = 14$ because there are 14 songs total
- $r = 5$ because 5 songs are selected for the playlist

Substitute those values into the formula:

$$\begin{aligned}
 {}_{14}C_5 &= 14! / [5!(14 - 5)!] \\
 &= 14! / (5!9!) \\
 &= (14 \times 13 \times 12 \times 11 \times 10) / (5 \times 4 \times 3 \times 2 \times 1) \\
 &= 2,002
 \end{aligned}$$

Now count the number of ways to assign the two distinct positions of Opening Track and Closing Track within a given 5-song playlist.

Because:

- the two positions are different, and
- one song cannot fill both positions,

we use permutations without repetition.

The formula is:

$${}_nP_r = n! / (n - r)!$$

where:

- n = number of available songs in the playlist
- r = number of positions to assign

For this stage:

- $n = 5$ because there are 5 songs in the selected playlist
- $r = 2$ because we are assigning 2 positions

Substitute those values into the formula:

$$\begin{aligned}
 {}_5P_2 &= 5! / (5 - 2)! \\
 &= 5! / 3! \\
 &= 5 \times 4 \\
 &= 20
 \end{aligned}$$

For every possible 5-song playlist, there are 20 possible ways to assign the two positions, so we multiply the number of playlists by the number of position assignments.

Total outcomes

$$= ({}_{14}C_5) \times ({}_5P_2)$$

$$= 2,002 \times 20$$
$$= 40,040$$

That value, 40,040, is the denominator in the probability ratio.

Numerator: Number of favorable outcomes

Now count the outcomes that meet both conditions:

- the playlist contains exactly 3 upbeat songs and 2 slow songs
- a specific slow song is chosen as the Opening Track

Start with the playlist selection stage.

Because the playlist must contain exactly 3 upbeat songs and 2 slow songs, we count those separately and then multiply.

First choose 3 upbeat songs from the 8 upbeat tracks:

$$8C3 = 8! / [3!(8 - 3)!]$$
$$= 8! / (3!5!)$$
$$= (8 \times 7 \times 6) / (3 \times 2 \times 1)$$
$$= 56$$

Next, because one specific slow song must be in the playlist and later assigned as the Opening Track, that song must already be included in the playlist. So 1 of the 2 slow-song spots is already taken by that specific slow song.

That means we need to choose only 1 additional slow song from the remaining 5 slow songs:

$$5C1 = 5$$

For the playlist selection stage, the number of favorable playlists is:

$$\text{Favorable playlists}$$
$$= (8C3) \times (5C1)$$
$$= 56 \times 5$$
$$= 280$$

Now move to the position-assignment stage.

For any one of those playlists, we require the specific slow song to be the Opening Track. That position is already filled.

This leaves only the Closing Track to assign. It must be assigned to one of the other 4 songs in the playlist.

Because there are 4 possible choices for the Closing Track, the number of valid position assignments is:

$$4P1 = 4$$

For each favorable playlist, there are 4 valid ways to assign the remaining position, so we multiply:

$$\begin{aligned} \text{Favorable outcomes} \\ &= 280 \times 4 \\ &= 1,120 \end{aligned}$$

That value, 1,120, is the numerator in the probability ratio.

Probability

Now divide the number of favorable outcomes by the total number of possible outcomes:

$$\begin{aligned} P(\text{exactly 3 upbeat and 2 slow, and specific slow song is Opening Track}) \\ &= 1,120 / 40,040 \\ &\approx 0.02797 \end{aligned}$$

So, the probability is about 0.0280, or 2.80%.

Chapter 10

Question 1

A. An expected value of 1.8 books means that if the bookstore observed many customers, the average number of books purchased per customer would tend to be about 1.8. It describes the long-run average outcome of the random process, not what must happen for any one customer.

B. Yes, 1.8 books is a valid expected value. Expected values for discrete random variables do not need to be possible individual outcomes. Instead, the expected value is a weighted average of all possible outcomes, using their probabilities as weights. Even though no single customer can buy exactly 1.8 books, the average across many customers can be 1.8 books per customer.

Question 2

A streaming service tracks the number of hours users watch content per week.

A. Group A has more consistent viewing behavior because it has the smaller standard deviation. A standard deviation of 1.5 hours indicates that most users' viewing times are relatively close to the average of 10 hours. Their viewing habits do not vary much from week to week.

B. Group B is more likely to include users with very low or very high viewing times because it has the larger standard deviation. A standard deviation of 6 hours indicates that viewing times are spread out over a much wider range. This means more users are likely to be far from the average, including both very low and very high values.

Question 3

A delivery company tracks the number of packages that arrive late on a given day. Let X be the number of late packages per day.

A. The given column, $P(X \leq x)$, is the cumulative distribution function (CDF). It gives the probability that the random variable is less than or equal to each value of x .

To find the probability for an individual value, subtract consecutive cumulative probabilities:

$$P(X = x) = P(X \leq x) - P(X \leq x - 1)$$

For the first value, there is no previous cumulative probability, so:

$$P(X = 0) = P(X \leq 0)$$

Using that process, we obtain the following probability distribution:

x	$P(X \leq x)$	$P(X = x)$
0	0.30	0.30
1	0.55	0.25
2	0.80	0.25
3	0.95	0.15
4	1.00	0.05

B. To verify that this is a valid probability distribution, check the two required conditions:

1. Every probability must be between 0 and 1.
All the probabilities in the $P(X = x)$ column satisfy that condition.
2. The probabilities must sum to 1.
 $0.30 + 0.25 + 0.25 + 0.15 + 0.05 = 1.00$

Because both conditions are satisfied, this is a valid probability distribution.

C. The probability that exactly 2 packages are late is:

$$P(X = 2) = 0.25$$

There is a 0.25 probability that exactly 2 packages are late.

D. The probability that more than 2 packages are late is:

$$\begin{aligned}P(X > 2) &= P(X = 3) + P(X = 4) \\ &= 0.15 + 0.05 \\ &= 0.20\end{aligned}$$

Alternatively, using the cumulative probabilities:

$$\begin{aligned}P(X > 2) &= 1 - P(X \leq 2) \\ &= 1 - 0.80 \\ &= 0.20\end{aligned}$$

There is a 0.20 probability that more than 2 packages are late.

E. The probability that between 1 and 3 packages are late, inclusive, is:

$$\begin{aligned}P(1 \leq X \leq 3) &= P(X \leq 3) - P(X \leq 0) \\ &= 0.95 - 0.30 \\ &= 0.65\end{aligned}$$

There is a 0.65 probability that between 1 and 3 packages are late, inclusive.

F. The probability that at least 1 package is late is:

$$\begin{aligned}P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - 0.30 \\ &= 0.70\end{aligned}$$

Alternatively, because $P(X \leq 0) = 0.30$:

$$\begin{aligned}P(X \geq 1) &= 1 - P(X \leq 0) \\ &= 1 - 0.30 \\ &= 0.70\end{aligned}$$

There is a 0.70 probability that at least 1 package is late.

Question 4

A casino offers a slot machine that costs \$1.00 per play.

A. To verify that this is a valid probability distribution, check the two required conditions:

1. Every probability must be between 0 and 1.
All the probabilities in the table satisfy that condition.
2. The probabilities must sum to 1.
 $0.489 + 0.260 + 0.100 + 0.100 + 0.040 + 0.010 + 0.001 = 1.000$

Because both conditions are satisfied, this table is a valid probability distribution.

B. The expected value of the payout is:

$$\mu = \Sigma[x \cdot P(x)]$$

$$= (0)(0.489) + (0.50)(0.260) + (1.00)(0.100) + (2.00)(0.100) + (5.00)(0.040) + (20.00)(0.010) + (100.00)(0.001)$$

$$= 0 + 0.13 + 0.10 + 0.20 + 0.20 + 0.20 + 0.10$$

$$= 0.93$$

The expected value of the payout is \$0.93 per play.

C. The machine costs \$1.00 per play, so the player's expected net gain or loss is:

Expected net = expected payout – cost to play

$$= 0.93 - 1.00$$

$$= -0.07$$

The player's expected net is -\$0.07 per play.

D. Players tend to lose money in the long run because the expected payout is less than the cost to play.

On average:

- the machine pays back \$0.93 per play
- the player pays \$1.00 per play

So, players lose an average of:

$$1.00 - 0.93 = \$0.07$$

Players lose 7 cents per play on average.

E. To calculate the standard deviation, use:

$$\sigma = \sqrt{\Sigma[(x - \mu)^2 \cdot P(x)]}$$

We already found that $\mu = 0.93$.

Now calculate the weighted squared deviations:

x	P(x)	x - μ	(x - μ) ²	(x - μ) ² · P(x)
0.00	0.489	-0.93	0.8649	0.4223361
0.50	0.260	-0.43	0.1849	0.0480740
1.00	0.100	0.07	0.0049	0.0004900
2.00	0.100	1.07	1.1449	0.1144900
5.00	0.040	4.07	16.5649	0.6625960
20.00	0.010	19.07	363.6649	3.6366490
100.00	0.001	99.07	9814.8649	9.8148649

Sum = 14.6995

Now take the square root:

$$\sigma = \sqrt{14.6995} = 3.83$$

The standard deviation of the payout distribution is about \$3.83.

F. The payout percentage is the percentage of money played that the machine returns to players in the long run.

$$\begin{aligned} \text{Payout percentage} &= (\text{expected payout} / \text{cost to play}) \times 100 \\ &= (0.93 / 1.00) \times 100 \\ &= 93\% \end{aligned}$$

The machine's payout percentage is 93%.

G. This slot machine pays back an average of \$0.93 per play, even though each play costs \$1.00. That means players lose money on average in the long run.

The standard deviation of about \$3.83 shows that payouts vary substantially from play to play. Most plays produce little or no payout, but occasional larger payouts increase the variability. So, while the average result is a small loss, the individual outcomes are spread out considerably.

Although the average loss per play is small, repeated play leads to predictable losses over time. The variability in outcomes can make short-term results feel favorable, but the long-run average favors the casino.

Chapter 11

Question 1

Identifying the distributions

- A. **Geometric distribution:** This is geometric because it asks for the probability that the first success occurs on a specific trial.

- B. **Binomial distribution:** This is binomial because it asks for the probability of exactly 6 successes in a fixed number of trials.
- C. **Negative binomial distribution:** This is negative binomial because it asks for the probability that the 6th success occurs on a specific trial.

Question 2

With and without replacement

- A. **Binomial distribution:** This is binomial because returning each item to the box keeps the probability of success constant from draw to draw, making the trials act independently.
- B. **Hypergeometric distribution:** This is hypergeometric because the items are not returned, so the composition of the box changes after each draw and the probability of success changes from one selection to the next.

Question 3

Customer support escalation

$$P(X = 4) = 0.1029$$

This is a geometric distribution problem because it asks for the probability that the first success occurs on a specific trial.

Distribution and method: Geometric distribution, exact probability

Parameter values:

- $p = 0.30$
- $x = 4$

The first escalated call occurring on the 4th call means:

- the first 3 calls are not escalated
- the 4th call is escalated

Calculation:

$$P(X = x) = (1 - p)^{x-1} p$$

$$\begin{aligned} P(X = 4) &= (1 - 0.30)^{4-1} (0.30) \\ &= (0.70)^3 (0.30) \\ &= 0.343(0.30) \\ &= 0.1029 \end{aligned}$$

Question 4

Defective items in 20 inspections

$$P(X \leq 2) = 0.9245$$

This is a binomial distribution problem because it counts the number of successes in a fixed number of independent trials with a constant probability of success.

Distribution and method: Binomial distribution, CDF

Parameter values:

- $n = 20$
- $p = 0.05$
- $x = 2$

“At most 2” means 0, 1, or 2 defective items.

Calculation:

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$P(X \leq 2) = \sum [20C_x (0.05)^x (0.95)^{20-x}], \text{ for } x = 0 \text{ to } 2$$

$$\begin{aligned} &= 20C_0(0.05)^0(0.95)^{20} + 20C_1(0.05)^1(0.95)^{19} + 20C_2(0.05)^2(0.95)^{18} \\ &= 0.3585 + 0.3774 + 0.1886 \\ &= 0.9245 \end{aligned}$$

Question 5

Fragile boxes

$$P(X = 3) = 0.2811$$

This is a hypergeometric distribution problem because it counts successes in a sample drawn from a finite population, where each draw changes the remaining population.

Distribution and method: Hypergeometric distribution, exact probability

Parameter values:

- $N = 40$
- $K = 12$
- $n = 8$
- $x = 3$

Calculation:

$$P(X = x) = [(K C x)(N - K C n - x)] / (N C n)$$

$$\begin{aligned} P(X = 3) &= [(12 C 3)(28 C 5)] / (40 C 8) \\ &= [220 \times 98,280] / 76,904,685 \\ &= 0.2811 \end{aligned}$$

Question 6

Free throws

$$P(X \geq 9) = 0.7946$$

This is a binomial distribution problem because it counts the number of successes in a fixed number of independent trials with a constant probability of success.

Distribution and method: Binomial distribution, complement of the CDF

Parameter values:

- $n = 12$
- $p = 0.80$
- $x = 9$

“At least 9” means 9, 10, 11, or 12 made free throws. It is easier to use the complement.

Calculation:

$$P(X \geq 9) = 1 - P(X \leq 8)$$

$$\begin{aligned} P(X \geq 9) &= 1 - \sum [12C_x (0.80)^x (0.20)^{12-x}], \text{ for } x = 0 \text{ to } 8 \\ &= 1 - 0.2054 \\ &= 0.7946 \end{aligned}$$

Question 7

Third success on the 10th trial

$$P(X = 10) = 0.0604$$

This is a negative binomial distribution problem because it asks for the probability that a specified success number occurs on a specific trial.

Distribution and method: Negative binomial distribution, exact probability

Parameter values:

- $r = 3$
- $p = 0.20$

- $x = 10$

This asks for the probability that the 3rd success occurs on the 10th trial.

Calculation:

$$P(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

$$\begin{aligned} P(X = 10) &= \binom{9}{2} (0.20)^3 (0.80)^7 \\ &= 36(0.008)(0.2097152) \\ &= 0.0604 \end{aligned}$$

Question 8

Faulty components

$$P(1 \leq X \leq 3) = 0.8628$$

This is a hypergeometric distribution problem because it counts successes in a sample drawn from a finite population, where each draw changes the remaining population.

Distribution and method: Hypergeometric distribution, probability for a range

Parameter values:

- $N = 25$
- $K = 7$
- $n = 6$
- lower bound = 1
- upper bound = 3

“Between 1 and 3, inclusive” means summing the probabilities for 1, 2, and 3 faulty components.

Calculation:

$$P(1 \leq X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3)$$

$$\begin{aligned} P(1 \leq X \leq 3) \\ &= \frac{[(7 C 1)(18 C 5)]}{(25 C 6)} \end{aligned}$$

- $\frac{[(7 C 2)(18 C 4)]}{(25 C 6)}$
- $\frac{[(7 C 3)(18 C 3)]}{(25 C 6)}$

$$= \frac{(7 \times 8,568)}{177,100}$$

- $\frac{(21 \times 3,060)}{177,100}$

- $(35 \times 816) / 177,100$

$$= 0.3388 + 0.3628 + 0.1612$$

$$= 0.8628$$

Chapter 12

Question 1

Choosing the Appropriate Distribution

- A. Binomial
- B. Poisson
- C. Binomial
- D. Poisson

Question 2

When Is the Poisson Distribution Appropriate?

- A. Appropriate
- B. Not appropriate (rate changes over time)
- C. Not appropriate (events are not independent; clustering)
- D. Appropriate

Question 3

Call center interpretation

- A. The average number of complaints per hour.
- B. No.
- C. No; variability around the average is expected.

Question 4

What is the probability that the hospital has no needle-stick injuries in a given week?

Let X represent the number of needle-stick injuries in one week.

Parameter value:

$$\lambda = 0.8$$

The question asks for the probability of no events occurring, so we need:

$$P(X = 0)$$

Use the Poisson probability formula:

$$P(X = x) = (e^{-\lambda} \cdot \lambda^x) / x!$$

Substitute the values:

$$P(X = 0) = (e^{-0.8} \cdot 0.8^0) / 0!$$

Simplify:

$$0.8^0 = 1$$

$$0! = 1$$

So:

$$P(X = 0) = e^{-0.8}$$

$$P(X = 0) \approx 0.4493$$

There is about a 0.4493 probability of observing no needle-stick injuries in a given week.

Question 5

What is the probability that there are at least 3 accidents in a given month?

Let X represent the number of accidents in one month.

Parameter value:

$$\lambda = 2.6$$

The phrase at least 3 means 3 or more, so the easiest method is to use the complement:

$$P(X \geq 3) = 1 - P(X \leq 2)$$

Now calculate the individual probabilities and add them:

$$P(X = 0) = (e^{-2.6} \cdot 2.6^0) / 0! = e^{-2.6} \approx 0.0743$$

$$P(X = 1) = (e^{-2.6} \cdot 2.6^1) / 1! \approx 0.1931$$

$$P(X = 2) = (e^{-2.6} \cdot 2.6^2) / 2! \approx 0.2510$$

Add them:

$$P(X \leq 2) = 0.0743 + 0.1931 + 0.2510 = 0.5184$$

Now take the complement:

$$P(X \geq 3) = 1 - 0.5184 = 0.4816$$

There is about a 0.4816 probability of observing at least 3 accidents in a given month.

Chapter 13

Question 1

Probability of an Exact Value

Correct answer: D. 0

For a continuous distribution, the probability of any single value is zero. That includes the mean.

So:

$$P(X = 50) = 0$$

Even though 50 is the center of the distribution and the most common value in terms of density, the probability of observing that exact value is zero. Probabilities for continuous variables are only defined over ranges of values.

Question 2

Interpreting the CDF

Correct answer: C. The probability that X is less than or equal to x

The cumulative distribution function gives the accumulated probability up to a value.

So:

$$\text{CDF at } x = P(X \leq x)$$

Question 3

Using Symmetry and Complements

Correct answer: D. 0.84

We are given:

$$P(X < \mu - a) = 0.16$$

Because the normal distribution is symmetric:

$$P(X > \mu + a) = 0.16$$

Now use the complement:

$$P(X < \mu + a) = 1 - 0.16 = 0.84$$

So the answer is:

$$P(X < \mu + a) = 0.84$$

Question 4

Comparing Relative Standing with z-scores

Correct answer: A. Student A

Calculate the z-scores.

For Student A:

$$z = (x - \mu) / \sigma = (85 - 80) / 5 = 1.0$$

For Student B:

$$z = (x - \mu) / \sigma = (90 - 85) / 10 = 0.5$$

Student A has the higher z-score, so Student A performed better relative to the other students in that course.

Question 5

Equality in Continuous Distributions

Correct answer: B. $P(X \leq x) = P(X < x)$

For continuous distributions, the probability of any exact value is zero:

$$P(X = x) = 0$$

Because of this:

$$P(X \leq x) = P(X < x)$$

Including or excluding a single point does not change the probability.

Question 6

Interpreting the Peak of a Normal Distribution

Correct answer: B. The value with the highest probability density

The peak of the normal distribution indicates where values are most concentrated. This corresponds to the highest probability density.

However, for continuous distributions, the probability at any single value is zero. So even at the peak:

$$P(X = \mu) = 0$$

The peak does not represent the highest probability. It represents the highest density.

Question 7

Empirical Rule

Mean = 70

Standard deviation = 5

The interval 65 to 75 is:

$70 - 5$ to $70 + 5 \rightarrow$ within 1 standard deviation of the mean

Using the empirical rule:

About 68% of values fall within ± 1 standard deviation.

Answer: 0.68

Question 8

Z-table Between Two Values

Mean $\mu = 800$

Standard deviation $\sigma = 100$

Convert to z-scores:

For 750:

$$z = (750 - 800) / 100 = -0.50$$

For 920:

$$z = (920 - 800) / 100 = 1.20$$

From the z-table:

$$P(Z < 1.20) = 0.88493$$

$$P(Z < -0.50) = 0.30854$$

Now subtract:

$$P(750 < X < 920) = 0.88493 - 0.30854 = 0.57639$$

Answer: 0.57639

Question 9

Z-table More Extreme Values

Mean $\mu = 70$

Standard deviation $\sigma = 8$

“More than 10 away from the mean” means:

$$X < 60 \text{ or } X > 80$$

Convert to z-scores:

$$z = (60 - 70) / 8 = -1.25$$

$$z = (80 - 70) / 8 = 1.25$$

By symmetry:

$$P(X < 60 \text{ or } X > 80) = 2 \times P(Z > 1.25)$$

From the z-table:

$$P(Z < 1.25) = 0.89435$$

So:

$$P(Z > 1.25) = 1 - 0.89435 = 0.10565$$

Now multiply:

$$2 \times 0.10565 = 0.21130$$

Answer: 0.21130

Question 10

Z-table Percentile to Data Value

Mean $\mu = 500$

Standard deviation $\sigma = 100$

The 90th percentile corresponds to:

$$P(Z < z) = 0.90$$

From the z-table:

$$z = 1.28$$

Convert back to x:

$$x = \mu + z\sigma$$

$$x = 500 + (1.28)(100)$$

$$x = 500 + 128 = 628$$

Answer: 628

Chapter 14

Question 1

Sampling Distribution vs. Data Distribution

This graph represents a sampling distribution.

It plots the sample means from 1,000 random samples of 20 strawberries each. It does not plot the sugar content values for individual strawberries.

Question 2

Interpretation of Standard Error

The standard error describes how much the sample mean varies from sample to sample.

The population mean and standard deviation stay fixed. The random samples change, so the sample means change. In this context, the standard error measures the typical variation among those sample means.

Question 3

Effect of Sample Size

Both sampling distributions center on $\mu = 60$.

The sampling distribution with $n = 64$ has less variability than the one with $n = 16$ because the standard error is σ / \sqrt{n} .

For $n = 16$:

$$\text{Standard error} = 12 / \sqrt{16} = 12 / 4 = 3$$

For $n = 64$:

$$\text{Standard error} = 12 / \sqrt{64} = 12 / 8 = 1.5$$

So, increasing the sample size from 16 to 64 cuts the standard error in half.

Question 4

Interpreting Probabilities: Sample Mean

$P(\bar{x} > 75) = 0.04$ means that 4% of random samples of 40 students would have a sample mean test score greater than 75.

It does not mean that 4% of individual students score above 75.

Question 5

Interpreting Probabilities: Sample Proportion

$P(\hat{p} \geq 0.25) = 0.03$ means that 3% of random samples of 150 orders would have a sample return proportion of at least 0.25.

In other words, even though the true return rate is 20%, about 3% of samples would show a return rate of 25% or higher due to random sampling variability.

Question 6

Central Limit Theorem & Normal Approximation

A. A population is strongly right-skewed. Sample size $n = 10$.

Unreasonable. The sample size is too small for the central limit theorem to overcome the strong skewness in the population.

B. A population is moderately skewed. Sample size $n = 35$.

Reasonable. The sample size is large enough for the central limit theorem to produce an approximately normal sampling distribution.

C. A population is normally distributed. Sample size $n = 5$.

Reasonable. When the population is normal, the sampling distribution of the mean is normal for any sample size.

D. A proportion problem with $p = 0.02$ and $n = 100$.

Unreasonable. The expected number of successes is too small ($np = 2$), so the normal approximation is not appropriate.

E. A proportion problem with $p = 0.40$ and $n = 50$.

Reasonable. Both $np = 20$ and $n(1 - p) = 30$ are sufficiently large, so the normal approximation is appropriate.

F. A population is uniform. Sample size $n = 6$.

Reasonable. A uniform distribution is symmetric, so the sampling distribution of the mean can become approximately normal with a relatively small sample size.

Question 7

Standard Deviation vs. Standard Error

The standard error equals the population standard deviation when $n = 1$.

Standard error = σ / \sqrt{n}

Set it equal to σ :

$$\sigma / \sqrt{n} = \sigma$$

That requires $\sqrt{n} = 1$, so $n = 1$.

When $n = 1$, the sample mean is just one individual observation, so its variability matches the population standard deviation.

Question 8

Total and SD of Total Scaling

As sample size increases, the mean of the sample total increases because:

$$\text{Mean of total} = n\mu$$

So if $\mu = 5$, the mean total is $5n$.

The standard deviation of the sample total also increases, but more slowly:

$$\text{SD of total} = \sigma\sqrt{n}$$

So if $\sigma = 2$, the standard deviation of the total is $2\sqrt{n}$.

The total grows directly with n , but its standard deviation grows with \sqrt{n} .

Question 9

Interpretation of Probability

Correct answer: **B**. 2% of samples have an average below 48.

Because the probability statement uses \bar{x} , it refers to sample means, not individual observations.

Question 10

Sample Total

Use the sampling distribution of the total because the question asks about the combined amount in 40 bottles.

Given:

$$\mu = 250 \text{ mL}$$

$$\sigma = 6 \text{ mL}$$

$$n = 40$$

Mean of the total:

$$n\mu = 40(250) = 10,000 \text{ mL}$$

Standard deviation of the total:

$$\sigma\sqrt{n} = 6\sqrt{40} = 6(6.325) = 37.95 \text{ mL}$$

So:

$$T \sim N(10,000, 37.95)$$

We want:

$$P(T < 9,960)$$

Calculate the z-score:

$$z = (T - n\mu) / (\sigma\sqrt{n})$$

$$z = (9,960 - 10,000) / 37.95$$

$$z = -40 / 37.95$$

$$z = -1.05$$

Using the cumulative z-table:

$$P(Z < -1.05) = 0.14686$$

Therefore:

$$P(T < 9,960) = 0.14686$$

The probability that the combined amount in 40 bottles is less than 9,960 mL is about **0.147**, or **14.7%**.

Question 11

Sample Proportion

Use the sampling distribution of the sample proportion because the question asks whether the sample proportion of failed doses meets a cutoff.

Given:

$$p = 0.03$$

$$n = 400$$

$$\text{Cutoff} = 0.05$$

We want:

$$P(\hat{p} \leq 0.05)$$

First, check the normal approximation:

$$np = 400(0.03) = 12$$

$$n(1 - p) = 400(0.97) = 388$$

Both values are at least 5, so the normal approximation is reasonable.

Calculate the standard error:

$$\sqrt{p(1 - p) / n} = \sqrt{0.03(0.97) / 400}$$

$$= \sqrt{0.0291 / 400}$$

$$= \sqrt{0.00007275}$$

$$= 0.00853$$

So:

$$\hat{p} \sim N(0.03, 0.00853)$$

Calculate the z-score:

$$z = (\hat{p} - p) / \sqrt{p(1 - p) / n}$$

$$z = (0.05 - 0.03) / 0.00853$$

$$z = 0.02 / 0.00853$$

$$z = 2.35$$

Using the cumulative z-table:

$$P(Z \leq 2.35) = 0.99061$$

Therefore:

$$P(\hat{p} \leq 0.05) = 0.99061$$

The probability that the sample meets the regulatory requirement is about **0.991**, or **99.1%**.

Question 12

Sample Mean

Use the sampling distribution of the mean because the question asks about the average step count for a sample of 49 users.

Given:

$$\mu = 7,800$$

$$\sigma = 2,100$$

$$n = 49$$

Calculate the standard error:

$$\sigma / \sqrt{n} = 2,100 / \sqrt{49}$$

$$= 2,100 / 7$$

$$= 300$$

So:

$$\bar{x} \sim N(7,800, 300)$$

We want:

$$P(\bar{x} > 8,300)$$

Calculate the z-score:

$$z = (\bar{x} - \mu) / (\sigma / \sqrt{n})$$

$$z = (8,300 - 7,800) / 300$$

$$z = 500 / 300$$

$$z = 1.67$$

Using the cumulative z-table:

$$P(Z \leq 1.67) = 0.95254$$

Because we want the probability greater than 8,300:

$$P(Z > 1.67) = 1 - 0.95254$$

$$= 0.04746$$

Therefore:

$$P(\bar{x} > 8,300) = 0.04746$$

The probability that the sample mean exceeds 8,300 steps is about **0.047**, or **4.7%**.

Question 13

Sample Proportion

Use the sampling distribution of the sample proportion because the question asks for the sample proportion of complaints marking the lowest 10%.

Given:

$$p = 0.18$$

$$n = 250$$

First, check the normal approximation:

$$np = 250(0.18) = 45$$

$$n(1 - p) = 250(0.82) = 205$$

Both values are at least 5, so the normal approximation is reasonable.

Calculate the standard error:

$$\sqrt{p(1 - p) / n} = \sqrt{0.18(0.82) / 250}$$

$$= \sqrt{0.1476 / 250}$$

$$= \sqrt{0.0005904}$$

$$= 0.02430$$

So:

$$\hat{p} \sim N(0.18, 0.02430)$$

We want the sample proportion at the lowest 10%, so:

$$P(\hat{p} \leq ?) = 0.10$$

From the cumulative z-table, the z-score closest to 0.10 is:

$$z = -1.28$$

Convert the z-score back to a sample proportion:

$$\hat{p} = p + z\sqrt{p(1 - p) / n}$$

$$\hat{p} = 0.18 + (-1.28)(0.02430)$$

$$\hat{p} = 0.18 - 0.03110$$

$$\hat{p} = 0.14890$$

So the lowest 10% cutoff is approximately:

$$\hat{p} = 0.149$$

A sample complaint proportion of about **0.149**, or **14.9%**, marks the lowest 10% of the sampling distribution.

However, this result alone does not prove that performance improved. If the true complaint rate is still 18%, about 10% of random samples would have complaint rates this low or lower just by chance. The cutoff can flag an unusually low sample result, but the manager would need additional evidence before concluding that the complaint rate has truly decreased.

In other words, you expect to see 10% below this cutoff when the overall rate remains constant. Seeing rates this low in particular samples is not necessarily a sign of improvement.

Chapter 15

Question 1

Interpreting a p-value

Answer: C

C is correct because the p-value is the probability of observing results at least as extreme as the sample result, assuming the null hypothesis is true.

Question 2

Statistical significance vs practical importance

Answer: C

C is correct because statistical significance indicates that the observed result is unlikely to be due to random sampling variability alone.

Question 3

Failing to reject the null hypothesis

Answer: C

C is correct because a large p-value means there is not enough evidence to reject the null hypothesis.

Question 4

Confidence interval and hypothesis testing

Answer: Not statistically significant

The results are **not statistically significant at $\alpha = 0.01$** . The A 99% confidence interval (for $\alpha = 0.01$) includes null hypothesis value of 200, indicating it is a plausible value. Therefore, there is insufficient evidence to reject the null hypothesis at the 0.01 level.

Question 5

Effect of sample size

Answer: C

C is correct because increasing the sample size reduces the standard error.

Question 6

Meaning of margin of error

Answer: B

B is correct because the margin of error reflects how much a sample estimate typically differs from the population parameter due to random sampling variability.

Question 7

Role of probability in inference

Answer: B

B is correct because probability models the variation in sample results due to random sampling.

Question 8

Significance level interpretation ($\alpha = 0.01$)

Answer: B

B is correct because the significance level represents the risk of rejecting a true null hypothesis due to random sampling variability.

Question 9

Effect of changing the significance level

Answer: B

B is correct because lowering α requires stronger evidence to reject the null hypothesis and reduces the risk of false positives.

Question 10

Relative probability in a normal distribution

Answer: C

C is correct because both values are the same distance from the mean.

A. 38 is 12 units below the mean ($50 - 38 = 12$)

B. 62 is 12 units above the mean ($62 - 50 = 12$)

In a normal distribution, values that are equally far from the mean have the same probability because the distribution is symmetric.

Question 11

Margin of error for a survey

A. Calculate the margin of error.

Answer: As a percentage, the margin of error is approximately **62% ± 2.75%**, or **59.25% to 64.75%**.

Use the margin of error formula for a proportion:

$$\text{MOE} = z \times \sqrt{\hat{p}(1 - \hat{p}) / n}$$

For a 95% confidence level, $z = 1.96$.

$$\hat{p} = 0.62$$

$$n = 1200$$

$$\text{MOE} = 1.96 \times \sqrt{0.62(1 - 0.62) / 1200}$$

$$\text{MOE} = 1.96 \times \sqrt{0.62(0.38) / 1200}$$

$$\text{MOE} = 1.96 \times \sqrt{0.2356 / 1200}$$

$$\text{MOE} = 1.96 \times \sqrt{0.0001963}$$

$$\text{MOE} = 1.96 \times 0.0140$$

$$\text{MOE} = 0.0275$$

As a percentage, the margin of error is approximately **2.75%**.

B. Interpret the result.

The survey estimate is **62% ± 2.75%**.

We can be 95% confident that the population percentage of city residents who support adding more protected bike lanes is approximately **59.25% to 64.75%**.

The margin of error accounts for random sampling variability. It does not account for other possible problems, such as biased wording or a non-representative sample.

Question 12

Confidence interval and hypothesis test for a mean

A. Calculate the 95% confidence interval.

Answer: The 95% CI is **[49.16, 64.84]**.

Use the confidence interval formula for a mean:

$$\bar{x} \pm z \times SE$$

First, calculate the standard error:

$$SE = s / \sqrt{n}$$

$$SE = 20 / \sqrt{25}$$

$$SE = 20 / 5$$

$$SE = 4$$

For a 95% confidence level, $z = 1.96$.

Margin of error:

$$MOE = 1.96 \times 4 = 7.84$$

Confidence interval:

$$57 - 7.84 = 49.16$$

$$57 + 7.84 = 64.84$$

The 95% confidence interval is:

[49.16, 64.84]

B. Calculate the two-sided p-value.

Answer: The two-sided p-value is **0.0802**.

The null hypothesis value is:

$$\mu_0 = 50$$

Calculate the z-test statistic:

$$z = (\bar{x} - \mu_0) / SE$$

$$z = (57 - 50) / 4$$

$$z = 7 / 4$$

$$z = 1.75$$

For a two-sided test, find the probability of a z-score at least as extreme as ± 1.75 .

Using the cumulative z-table:

$$P(Z \leq -1.75) = 0.0401$$

Double this value for the two-sided p-value:

$$p = 2 \times 0.0401$$

$$p = 0.0802$$

The two-sided p-value is **0.0802**.

C. Interpret both results.

The p-value is **0.0802**, which is greater than $\alpha = 0.05$. Therefore, the result is **not statistically significant**. We fail to reject the null hypothesis.

The 95% confidence interval is **[49.16, 64.84]**. This interval barely includes the null value of 50, so 50 remains a plausible population mean.

Although the sample mean of 57 looks higher than the benchmark of 50, the evidence is not strong enough to conclude that the alert system improves the population mean score. The confidence interval and hypothesis test help prevent us from mistaking an apparent effect for a real population effect when random sampling variability could explain the result.

D. Do the confidence interval and hypothesis test agree?

Yes. The confidence interval includes the null value of 50, and the p-value is greater than 0.05. Both methods indicate that the result is not statistically significant at the 0.05 level.