Statistics and Forensic Science

Jeff Holt

2021-12-20

Contents

Chapter 1

Prerequisites

This is a sample book that is currently a work in progress by the *University of Virginia*. This is not a finished product.

Chapter 2

Introduction

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter 2. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter 5.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 2.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 2.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the **bookdown** package (Xie, 2021) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015).

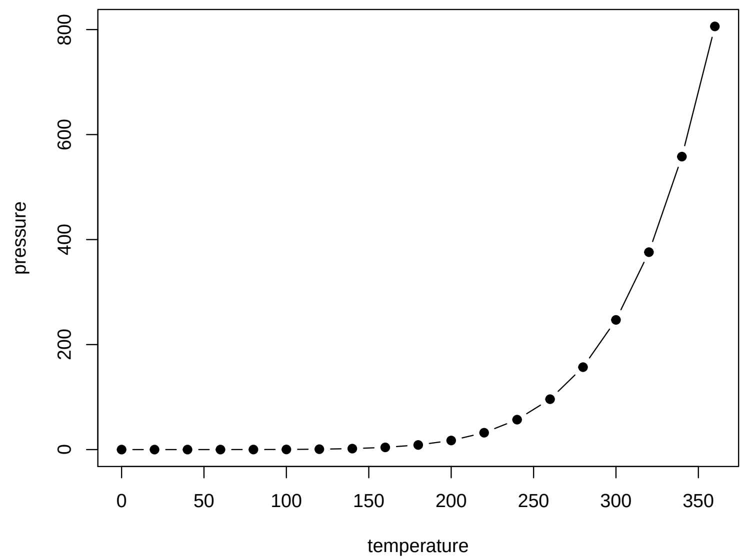

Figure 2.1: Here is a nice figure!

Table 2.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Chapter 3

Literature

Here is a review of existing methods.

Chapter 4

Probability

Probability is an important and complex field of study. Fortunately, only a few basic issues in probability theory are essential for understanding statistics at the level covered in this book. These basic issues are covered in this chapter.

4.1 Introduction

TODO: Edit the introductory section below.

"The introductory section discusses the definitions of probability. This is not as simple as it may seem. The section on basic concepts covers how to compute probabilities in a variety of simple situations. The Gambler's Fallacy Simulation provides an opportunity to explore this fallacy by simulation. The Birthday Demonstration illustrates the probability of finding two or more people with the same birthday. The Binomial Demonstration shows the binomial distribution for different parameters. The section on base rates discusses an important but often-ignored factor in determining probabilities. It also presents Bayes' Theorem. The Bayes' Theorem Demonstration shows how a tree diagram and Bayes' Theorem result in the same answer. Finally, the Monty Hall Demonstration lets you play a game with a very counterintuitive result."

4.2 Another Subsection

In what follows we introduce the basics of probability, which quantifies the chance that something occurs. Let's start with a specific situation to serve as an example. Suppose that we have a population of 1000 people, each classified by gender (Female or Male) and blood type (A, AB, B, or O). The number of people of each combination of gender and blood type is shown in the table below.

	F	M
A	175	235
AB	16	24
B	37	63
O	202	248

Gender and blood type counts

Suppose that we select one of the people at random, and note that person's gender and blood type. There are various possibilities for outcomes, such as the person selected is female with Type A blood. The chance that this outcome occurs is called the **probability** of this outcome. As 175 out of the 1000 people are female of blood type A, the probability of selecting such a person is

$$\text{P(Female and Type A)} = \frac{175}{1000} = 0.175$$

Here P(Female and Type A) stands for the probability that the person selected is Female **and** has Type A blood. A sample of other probabilities that come directly from the table are

$$\begin{aligned}
\text{P(Female and Type O)} &= \frac{202}{1000} &= 0.202 \\
\text{P(Male and Type AB)} &= \frac{24}{1000} &= 0.024 \\
\text{P(Male and Type A)} &= \frac{235}{1000} &= 0.235
\end{aligned}$$

We also can combine table entries to compute other probabilities. For instance, there are a total of $235 + 24 + 63 + 248 = 570$ males in our population, so

$$\text{P(Male)} = \frac{570}{1000} = 0.57$$

There are a total of $16 + 24 = 40$ people with blood type AB, so

$$\text{P(Type AB)} = \frac{40}{1000} = 0.04$$

4.2.1 Sample Questions

1. What is the probability that the person selected is female?

Answer: $P(\text{Female}) = 1 - P(\text{Male}) = 1 - 0.57 = 0.43$ This is equivalent to $(175 + 16 + 37 + 202)/1000$.

2. What is the probability that the person selected is male and has type A blood and type O blood?

Answer: $P(\text{Male and A and O}) = 0$. The probability of this event occurring in the table is zero, as no men have two different types of blood.

4.3 And vs Or

Above we saw that P(Female and Type A) = 0.175. Suppose we we instead would like to know P(Female or Type A)? That is, we want to know the probability that a randomly selected person is either female or has type A blood. (Note that this group includes those who are both female **and** have type A blood.) From our table we see that the total number in this group is $175 + 16 + 37 + 202 + 235 = 665$ so that

$$\text{P(Female or Type A)} = \frac{665}{1000} = 0.665$$

Other combinations are also possible. For instance, there are $175 + 235 = 310$ people with type A blood and $37 + 63 = 100$ people with type B blood, so there are a total of $310 + 100 = 410$ people that have either type A or type B blood. Therefore, we have

$$\text{P(Type A or Type B)} = \frac{410}{1000} = 0.41$$

4.3.1 Sample Questions

1. What is the probability of selecting a female or male with type AB blood?

Answer: $P((\text{Male or Female}) \text{ and AB}) = P((\text{Male and AB}) \text{ or } (\text{Female and AB})) = (16 + 24)/1000 = 0.04$

4.4 Conditional Probability

Now suppose that we just focus on the females in the population, which forms a subset of the population. The probability that a randomly selected female has blood type O is an example of a **conditional probability**.

There are 430 females in the population, and 202 of those have type O blood. Hence the probability that female chosen at random has type O blood is

$$\text{P(Type O} \mid \text{Female)} = \frac{202}{430} = 0.470$$

The notation P(Type O | Female) is read out loud as *The probability of blood type O, given the individual is female.* Conditional probabilities arise all the time when evaluating forensic evidence. Other examples of conditional probabilities:

The probability of Type AB, given that the person is male:

$$\Pr(\text{Type AB} \mid \text{Male}) = \frac{24}{570} = 0.042$$

The probability the person is female, given Type B blood:

$$\Pr(\text{Female} \mid \text{Type B}) = \frac{37}{100} = 0.37$$

The probability the person is male, given Type A blood:

$$\Pr(\text{Male} \mid \text{Type A}) = \frac{235}{410} = 0.573$$

4.4.1 Sample Questions

1. What is the probability of blood type AB or B given the person selected is female?

Answer: $P(\text{AB or B} \mid \text{Female}) = (16 + 37)/430 = 0.123$

Chapter 5

Methods

We describe our methods in this chapter.

Math can be added in body using usual syntax like this

5.1 math example

p is unknown but expected to be around 1/3. Standard error will be approximated

$$SE = \sqrt(\frac{p(1-p)}{n}) \approx \sqrt{\frac{1/3(1-1/3)}{300}} = 0.027$$

You can also use math in footnotes like this[1].

We will approximate standard error to 0.027[2]

[1] where we mention $p = \frac{a}{b}$

[2] p is unknown but expected to be around 1/3. Standard error will be approximated

$$SE = \sqrt(\frac{p(1-p)}{n}) \approx \sqrt{\frac{1/3(1-1/3)}{300}} = 0.027$$

Chapter 6

Applications

Some *significant* applications are demonstrated in this chapter.

6.1 Example one

6.2 Example two

Chapter 7

Final Words

We have finished a nice book.

Bibliography

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2021). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.24.