

Probability and Conditional Probability

Bret Hanlon and Bret Larget

Department of Statistics
University of Wisconsin—Madison

September 23, 2010

Case Study

Example

Example 9.3 beginning on page 213 of the text describes an experiment in which fish are placed in a large tank for a period of time and some are eaten by large birds of prey. The fish are categorized by their level of parasitic infection, either uninfected, lightly infected, or highly infected. It is to the parasites advantage to be in a fish that is eaten, as this provides an opportunity to infect the bird in the parasites next stage of life. The observed proportions of fish eaten are quite different among the categories.

	Uninfected	Lightly Infected	Highly Infected	Total
Eaten	1	10	37	48
Not eaten	49	35	9	93
Total	50	45	46	141

The proportions of eaten fish are, respectively, $1/50 = 0.02$, $10/45 = 0.222$, and $37/46 = 0.804$.

Questions

- There are three *conditional probabilities* of interest, each the probability of being eaten by a bird given a particular infection level.
- How do we test if these are the same?
- How do we estimate differences between the probability of being eaten in different groups?
- Is there a relationship between infection level in the fish and bird predation?

A Second Case Study

Example

Example 9.4 on page 220 describes an experiment. In Costa Rica, the vampire bat *Desmodus rotundus* feeds on the blood of domestic cattle. If the bats respond to a hormonal signal, cows in estrous (in heat) may be bitten with a different probability than cows not in estrous. (The researcher could tell the difference by harnessing painted sponges to the undersides of bulls who would leave their mark during the night.)

	In estrous	Not in estrous	Total
Bitten by a bat	15	6	21
Not bitten by a bat	7	322	329
Total	22	328	350

The proportion of bitten cows among those in estrous is $15/22 = 0.682$ while the proportion of bitten cows among those not in estrous is $6/328 = 0.018$.

Questions

- Are the probabilities of being bitten different for cows in estrous or not?
- How do we estimate the difference in probabilities of being bitten?
- How do we estimate the odds ratio?
- Here, the odds of a cow in estrous being bitten are roughly 2 to 1, while the odds of a cow not in estrous being bitten are roughly 2 to 100, so the odds ratio is about 100 times larger to be bitten for cows in estrous compared to those not.
- How do we quantify uncertainty in this estimate?

The Big Picture

- When comparing two categorical variables, it is useful to summarize the data in tables.
- Data in the tables can be used to calculate observed proportions sampled from different populations.
- We may have interest in estimating differences between population probabilities.
- We may wish to test if population proportions are different.
- We may wish to test if two categorical variables are independent.

More Probability

- To understand the methods for comparing probabilities in different populations and analyzing categorical data, we need to develop notions of:
 - ▶ *conditional probability*; and
 - ▶ *independence*;
- We will also more formally introduce some probability ideas we have been using informally.

Example

- We will introduce formal notions of probability in the context of the following trivial example.

Bucket 1 contains colored balls in the following proportions:

- 10% red;
- 60% white; and
- 30% blue.

Bucket 2 has colored balls in different proportions:

- 10% red;
- 40% white; and
- 50% blue.

A bucket is selected at random with equal probabilities and a single ball is selected at random from that bucket.

- Think of the buckets as two biological populations and the colors as traits.

Outcome Space

Definition

- A *random experiment* is a setting where something happens by chance.
 - An *outcome space* is the set of all possible elementary outcomes.
 - An *elementary outcome* is a complete description of a single result from the random experiment (which, in fact, might be rather complicated, but is called elementary because it cannot be divided any further).
-
- In the example, one elementary outcome is $(1, W)$ meaning Bucket 1 is selected and a white ball is drawn.
 - The outcome space is the set of six possible elementary outcomes:

$$\Omega = \{(1, R), (1, W), (1, B), (2, R), (2, W), (2, B)\}$$

Probability

- Each elementary outcome has an associated *probability*.
- All probabilities are between 0 and 1.
- The sum of all probabilities in the outcome space is 1.

$$\begin{array}{lll} P((1, R)) = 0.05 & P((1, W)) = 0.30 & P((1, B)) = 0.15 \\ P((2, R)) = 0.05 & P((2, W)) = 0.20 & P((1, B)) = 0.25 \end{array}$$

Events

Definition

- An *event* is a subset (possibly empty, possibly complete) of elementary outcomes from the outcome space.
- The probability of an event is the sum of probabilities of the outcomes it contains.

Examples:

- $P(\text{Bucket 1}) = P(\{(1, R), (1, W), (1, B)\}) = 0.05 + 0.30 + 0.15 = 0.5.$
- $P(\text{Red Ball}) = P(\{(1, R), (2, R)\}) = 0.05 + 0.05 = 0.1.$
- $P(\text{Bucket 1 and Red Ball}) = P(\{(1, R)\}) = 0.05.$
- $P(\Omega) = 1.$

Combining Events

Definition

- The *union* (or) of two events is the set of all outcomes in either or both events.
- The *intersection* (and) of two events is the set of all outcomes in both events.
- The *complement* (not) of an event is the set of everything in Ω not in the event.

Definition

Events are *mutually exclusive* if they have no outcomes in common. This is the same as saying their intersection is empty.

The Addition Rule

Addition Rule for Mutually Exclusive Events

If events A and B are mutually exclusive, then $P(A \text{ or } B) = P(A) + P(B)$.

Example:

- The probability of a red ball is 0.1 because

$$P\left((1, R) \text{ or } (2, R)\right) = P\left((1, R)\right) + P\left((2, R)\right) = 0.05 + 0.05$$

The General Addition Rule

General Addition Rule

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B).$$

Example:

- The probability that either Bucket 1 is chosen or the ball is red is

$$\begin{aligned} P(\text{Bucket 1 or Red}) &= P(\text{Bucket 1}) + P(\text{Red}) - P(\text{Bucket 1 and Red}) \\ &= 0.5 + 0.1 - 0.05 = 0.55 \end{aligned}$$

Random Variables

Definition

A *random variable* is a rule that attaches a number to each elementary outcome.

- Note that as each elementary outcome has a probability and these probabilities sum to one, the random variable specifies how the probability in Ω should be distributed on the real line.
- This is called the *distribution* of the random variable.
- If we define W to be the number of white balls sampled, then W has possible values 0 and 1 and

$$\begin{aligned} P(W = 0) &= P\left((1, R) \text{ or } (1, B) \text{ or } (2, R) \text{ or } (2, B)\right) \\ &= 0.05 + 0.15 + 0.05 + 0.25 = 0.5 \end{aligned}$$

and $P(W = 1) = 1 - P(W = 0) = 0.5$.

Conditional Probability

Definition

- The *conditional probability* of an event given another is the probability of the event given that the other event has occurred.
- If $P(B) > 0$,

$$P(A | B) = \frac{P(A \text{ and } B)}{P(B)}$$

- Define events B_1 and B_2 to mean that Bucket 1 or 2 was selected and let events R , W , and B indicate if the color of the ball is red, white, or blue.
- By the description of the problem, $P(R | B_1) = 0.1$, for example.
- Using the formula,

$$\begin{aligned} P(R | B_1) &= \frac{P(R \text{ and } B_1)}{P(B_1)} \\ &= \frac{0.05}{0.5} = 0.1 \end{aligned}$$

Independence

Definition

- Events A and B are *independent* if information about one does not affect the other.
- That is,

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

or

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

- This is equivalent to

$$P(A \text{ and } B) = P(A)P(B)$$

General Multiplication Rule

General Multiplication Rule

$$P(A \text{ and } B) = P(A)P(B | A)$$

or

$$P(A \text{ and } B) = P(B)P(A | B)$$

- This rule follows directly from the definition of conditional probability.
- It makes sense: think of the events occurring in order. For A and B to both occur, either A occurs and then B occurs given A has already occurred, or vice versa.

Example

- Events B_1 and R are independent as $P(B_1 \text{ and } R) = 0.05$ and $P(B_1) \times P(R) = 0.5 \times 0.1 = 0.05$.
- Events B_2 and W are not independent as $P(B_2 \text{ and } W) = 0.2$, while and $P(B_2) \times P(W) = 0.5 \times 0.5 = 0.25$.

Partitions

Definition

Events B_1, B_2, \dots form a *partition of Ω* if they are mutually exclusive and Ω is the union of the B_i .

- Note that if B_1, B_2, \dots are a partition of Ω , then

$$\sum_i P(B_i) = 1$$

and

$$P(B_i \text{ and } B_j) = 0$$

provided $i \neq j$.

- In the example, B_1 and B_2 form a partition as exactly one bucket is selected.
- Also, R , W , and B form a partition as the ball must have exactly one color.

Law of Total Probability

Definition

If B_1, B_2, \dots form a partition, then

$$P(A) = \sum_i P(B_i)P(A | B_i)$$

- A tree diagram is helpful (use the board)!
- The *law of total probability* says that the unconditional $P(A)$ is a weighted average of the conditional probabilities $P(A | B_i)$ weighted by the probabilities of the conditions $P(B_i)$.
- For example, the probability of a blue ball is

$$\begin{aligned} P(B) &= P(B_1)P(B | B_1) + P(B_2)P(B | B_2) \\ &= (0.5)(0.3) + (0.5)(0.5) = 0.4 \end{aligned}$$

Bayes' Theorem

Definition

Bayes' theorem shows how to invert conditional probabilities.

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

- The denominator is usually found using the law of total probability.
- Bayes' theorem is the fundamental result in probability necessary for the *Bayesian approach to statistical inference*.
- It helps to understand Bayes' theorem with a tree diagram.

Example

- Find the probability that Bucket 1 was selected given that we see a white ball.

See chalk board for solution.

Problem 1

Problem

A male fruit fly is equally likely to carry alleles a or b on the X chromosome. Allele a is lethal for females, so if the male fruit fly carries a , all of its offspring will be male. If the male fly carries allele b , then each offspring is equally likely to be male or female.

If we observe three offspring and all are male, what is the conditional probability that the fly carries allele a ?

Problem 2

Problem

The prevalence of a certain type of cancer among women aged 50–60 is 1 in 150. A blood test will be positive 98% of the time if the cancer is present, but is also positive 7% of the time if the cancer is not present.

- 1 In a routine check-up, a 55-year-old woman receives a positive blood test. What is the probability that she has this type of cancer?
- 2 How would you interpret this calculation if the test were ordered because of the presence of other symptoms associated with the disease?

Problem 3

Problem

A species of plant was present at a site in northern Wisconsin in 1950 and appeared in ten percent of randomly located quadrats. If the plant is present in 2010, it has a ten percent chance of appearing in a quadrat independently. However, if the plant has gone extinct at this site, it has no chance of appearing. Before seeing new 2010 data, an ecologist believes that there is a 30 percent chance that the plant is extinct at this site.

- 1 Given that the plant is present at the site, what is the probability that the plant is not present in any of a random sample of five quadrats?
- 2 What is the unconditional probability that the plant is not present in any of a random sample of five quadrats?
- 3 Given that the plant is not present in any of a random sample of five quadrats, what is the probability that the plant is extinct at this site?
- 4 What if the same is true for 20 quadrats?

What you should know

You should know:

- facts about probabilities and events;
- the definition of independence;
- the definition of conditional probability;
- how to use multiplication and addition rules properly;
- how to use the law of total probability;
- how to use Bayes' theorem.