4.4 Conditional Probability

Dr. Lauren Perry

Goals

1. Use contingency tables to describe relationships between two variables.
2. Find and interpret joint and marginal probabilities.
3. Find and interpret conditional probabilities.
4. Use the multiplication rule to find the probability of two events occurring together.

Contingency Tables

A **contingency table** is a way to summarize **bivariate data**, or data from two variables.

Smallpox in Boston (1726)

Inoculated

yes

no

total

Result

lived

238

5136

5374

died

6

844

Contingency Tables

- These are basically two-variable frequency distributions.
- We can convert to proportions by dividing each count by the total number of observations.

Inoculated

yes

no

total

Result

lived

0.0382

0.8252

0.8634

died

0.0010

0.1356

0.1366

Inoculated

yes

no

total

Result

lived

0.0382

0.8252

0.8634

died

0.0010

0.1356

0.1366

What can we learn about the result of smallpox if we already know something about inoculation status?

- For example, given that a person is inoculated, what is the probability of death?
- To figure this out, we restrict our attention to the 244 inoculated cases. Of these, 6 died. So the probability is $6/244$.

Conditional Probability

Conditional probability: the probability of some event A if we know that event B occurred (or is true):

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

where the symbol $|$ is read as "given".

- For death given inoculation,

$$P(\text{death}|\text{inoculation}) = \frac{P(\text{death and inoculation})}{P(\text{inoculation})} = \frac{0.0010}{0.0392} = 0.0255$$

- We could also write this as

$$P(\text{death}|\text{inoculation}) = \frac{P(\text{death and inoculation})}{P(\text{inoculation})} = \frac{6/6224}{244/6224} = \frac{6}{244}$$

Independent Events

If knowing whether event B occurs tells us nothing about event A, the events are **independent**. For example, if we know that the first flip of a (fair) coin came up heads, that doesn't tell us anything about what will happen next time we flip that coin.

We can test for independence by checking if $P(A|B) = P(A)$.

Multiplication Rule for Independent Processes

If A and B are independent events, then

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

- We can extend this to more than two events:

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

- Note that if $P(A \text{ and } B) \neq P(A)P(B)$, then A and B are *not* independent.

Example

Find the probability of rolling a 6 on your first roll of a die and a 6 on your second roll.

Let $A =$ (rolling a 6 on first roll) and $B =$ (rolling a 6 on second roll). For each roll, the probabiltiy of getting a 6 is $1/6$, so $P(A) = \frac{1}{6}$ and $P(B) = \frac{1}{6}$.

Then, because each roll is independent of any other rolls,

$$P(A \text{ and } B) = P(A)P(B) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

General Multiplication Rule

If A and B are any two events, then

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

- This is just the conditional probability formula, rewritten in terms of $P(A \text{ and } B)$!

Checkpoint

Suppose we know that 38.4% of US households have dogs and that *among those with dogs*, 23.1% have cats. Find the probability that a US household has both dogs and cats.