Introduction to Statistics and Probability

Probability : Contingency Tables

Kevin O'Brien

Spring 2014

This Presentation

1. Contingency Tables
2. Conditional Probability: Worked Examples
3. Joint Probability Tables
4. The Multiplication Rule
5. Law of Total Probability

Contingency Tables

Suppose there are 100 students in a first year college intake.

- 44 are male and are studying computer science,
- 18 are male and studying engineering
- 16 are female and studying computer science,
- 22 are female and studying engineering.

We assign the names M, F, C and E to these events that a student, randomly selected from this group is:

- **M** Male
- **F** Female
- **C** Studying Computer Science
- **E** Studying Engineering

Contingency Tables

- The most effective way to handle this data is to draw up a table. We call this a ***contingency table***.
- A contingency table is a table in which all possible outcomes for one variable are listed as row headings and all possible outcomes for a second variable are listed as column headings.
- The value entered in each cell of the table is the frequency of each joint occurrence.

For the Student Intake example

	C	E	Total
M	44	18	62
F	16	22	38
Total	60	40	100

Contingency Tables

It is now easy to deduce the probabilities of the respective events, by looking at the totals for each row and column.

We call these probabilities the ***marginal probabilities***.

- P(C) = 60/100 = 0.60
- P(E) = 40/100 = 0.40
- P(M) = 62/100 = 0.62
- P(F) = 38/100 = 0.38

Marginal Probabilities

- In the context of joint probability tables, a ***marginal probability*** is so named because it is a marginal total of a row or a column.
- Whereas the probability values in the cells of the table are probabilities of joint occurrence, the marginal probabilities are the simple (i.e. unconditional) probabilities of particular events.

Remark:
The information we were originally given can also be expressed as:

- $P(C \cap M) = 44/100 = 0.44$
- $P(C \cap F) = 16/100 = 0.16$
- $P(E \cap M) = 18/100 = 0.18$
- $P(E \cap F) = 22/100 = 0.22$

We can call these probabilities the ***joint probabilities***.

Joint Probability Tables

- A ***joint probability table*** is similar to a contingency table, but for that the value entered in each cell of the table is the probability of each joint occurrence.
- Often, the probabilities in such a table are based on observed frequencies of occurrence for the various joint events.

Joint Probability Tables

	C	E	Total
M	0.44	0.18	0.62
F	0.16	0.22	0.38
Total	0.60	0.40	1.00

Conditional Probabilities : Example 1

Recall the definition of conditional probability:

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

- $P(A|B)$: probability of A **given** B,
- $P(A \cup B)$: joint probability of A and B,
- $P(B)$: probability of B.

Conditional Probabilities : Example 1

Using the conditional probability formula, compute the following:

(1) $P(C|M)$: Probability that a student is a computer science student, given that he is male.

(2) $P(E|M)$: Probability that a student studies engineering, given that he is male.

(3) $P(F|E)$: Probability that a student is female, given that she studies engineering.

(4) $P(E|F)$: Probability that a student studies engineering, given that she is female.

Conditional Probabilities : Example 1

Part 1) Probability that a student is a computer science student, given that he is male.

$$P(C|M) = \frac{P(C \cap M)}{P(M)} = \frac{0.44}{0.62} = 0.71$$

Part 2) Probability that a student studies engineering, given that he is male.

$$P(E|M) = \frac{P(E \cap M)}{P(M)} = \frac{0.18}{0.62} = 0.29$$

Part 3) Probability that a student is female, given that she studies engineering.

$$P(F|E) = \frac{P(F \cap E)}{P(E)} = \frac{0.22}{0.40} = 0.55$$

Part 4) Probability that a student studies engineering, given that she is female.

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{0.22}{0.38} = 0.58$$

Remark: $P(E \cap F)$ is the same as $P(F \cap E)$.

Multiplication Rule

- This useful multiplication rule follows from the definition of conditional probability.
- First we algebraically re-arrange the conditional probability equation.

$$P(A \cap B) \; = \; P(A|B) \times P(B).$$

- Equivalently $P(A \cap B) \; = \; P(B|A) \times P(A)$.
- Therefore we can say:

$$P(A|B) \times P(B) \; = \; P(B|A) \times P(A).$$

Multiplication Rule

As an aside, for **independent events**, (events which have no influence on one another), the multiplication rule simplifies to:

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

Multiplication Rule

Going back to our example:

From the first year intake example, check that

$$P(E|F) \times P(F) = P(F|E) \times P(E)$$

- $P(E|F) \times P(F) = 0.58 \times 0.38 = 0.22$
- $P(F|E) \times P(E) = 0.55 \times 0.40 = 0.22$

Law of Total Probability

- The law of total probability is a fundamental rule relating marginal probabilities to conditional probabilities.
- The result is often written as follows:

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

- Here $P(A \cap B^c)$ is joint probability that event A occurs and B does not.

Law of Total Probability

Using the multiplication rule, this can be expressed as

$$P(A) = [P(A|B) \times P(B)] + [P(A|B^c) \times P(B^c)]$$

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

From the first year intake example , check that

$$P(E) = P(E \cap M) + P(E \cap F)$$

with $P(E) = 0.40$, $P(E \cap M) = 0.18$ and $P(E \cap F) = 0.22$

$$0.40 = 0.18 + 0.22$$

Remark: M and F are complement events.