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 the events that a student, randomly selected from this group, is male, female, studying computer science, and studying engineering respectively.

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 events (or outcomes) for one variable are listed as row headings, all possible events for a second variable are listed as column headings, and the value entered in each cell of the table is the frequency of each joint occurrence.

	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.

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

Remark:

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

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

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.

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

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.
- From the first year intake example, the marginal probabilities are $P(C)$, $P(E)$, $P(M)$ and $P(F)$ respectively.

Conditional Probabilities : Example 1

Recall the definition of conditional probability:

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

Using this 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.

Refer back to the contingency table to appraise your results.

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.

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

Conditional Probabilities : Example 1

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 $P(F \cap E)$.