Basic Principle of Statistics: Rare Event Rule

If, under a given assumption, the probability of a particular observed event is exceptionally small, we conclude that the assumption is probably not correct.

Example: if you flip a coin 10 times and observe 10 Heads and 0 Tails, would you believe that it is a normal (balanced) coin? Or would you rather have a doubt and have the coin checked out?

The need to compute probabilities

To use the rare even rule, we need to be able to compute probabilities under given assumptions.

Basic concepts of Probabilities

- **Event**

  any *collection of results or outcomes* of a procedure, or an experiment, or a game

- **Simple Event**

  an outcome or an event that cannot be further broken down into simpler components

- **Sample Space**

  (for a procedure) the sample space consists of all possible simple events; that is, the sample space consists of all outcomes that cannot be broken down any further

Notation for Probabilities

P **- denotes a probability.**

A*, *B*, and *C **- denote specific events.**

***P*(*A*)** **- denotes the probability of event *A* occurring.**

Classical Approach to Probability
(Requires Equally Likely Outcomes)

Assume that a given procedure has n different simple events and that each of those simple events has an equal chance of occurring. If event A can occur in s of these n ways, then

$$P(A) = \frac{s}{n} = \frac{\text{number of ways } A \text{ can occur}}{\text{total number of different simple events}}$$

Law of Large Numbers

As a procedure is repeated again and again, the relative frequency of an event tends to approach its actual probability.

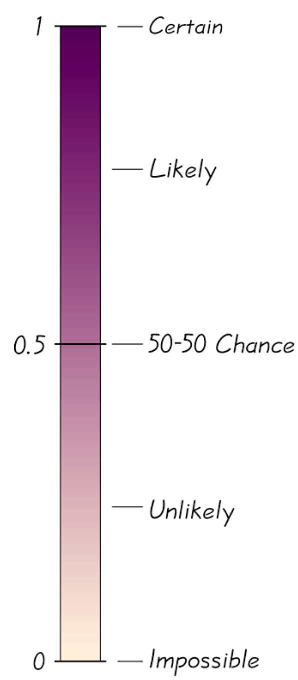
$$\text{Relative frequency} = \frac{\text{\# of times } A \text{ occurred}}{\text{\# of times procedure was repeated}}$$

Probability Limits

Always express a probability as a fraction or decimal number between 0 and 1.

- **The probability of an impossible event is 0.**
- **The probability of an event that is certain to occur is 1.**
- **For any event *A*, the probability of *A* is between 0 and 1 inclusive.**
  That is, $0 \le P(A) \le 1$.

Possible Values for Probabilities



Rounding Off Probabilities

When expressing the value of a probability, either give the **exact** fraction or decimal or **round off** final decimal results to **three** significant digits.

(*Suggestion*: When a probability is not a simple fraction such as 2/3 or 5/9, express it as a decimal so that the number can be better understood.)

Compound Event

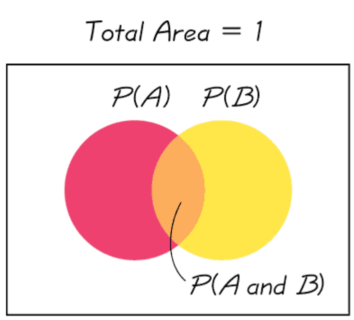
$P(A \text{ or } B)$ = P (event A occurs or event B occurs or they both occur)

Formal Addition Rule

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

where $P(A \text{ and } B)$ denotes the probability that A and B both occur at the same time

Illustration (Venn Diagram)



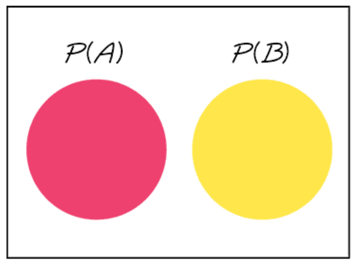
A **is the red disk,** ***B*** **is the yellow disk**

A* or *B **is the total area covered by both disks.**

Disjoint Events

Events *A* and *B* are disjoint (or mutually exclusive) if they cannot occur at the same time. (That is, disjoint events do not overlap.)

Total Area = 1



Venn Diagram for Disjoint Events

Formal Addition Rule for disjoint events

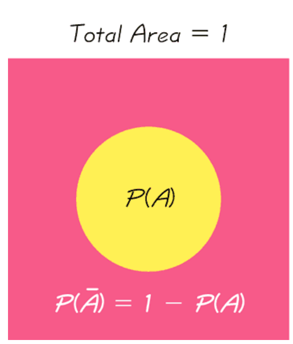
$$P(A \text{ or } B) = P(A) + P(B)$$

(only if *A* and *B* are disjoint)

Complementary Events

The complement of event A, denoted by $\bar{A}$, consists of all outcomes in which the event A does not occur.

Venn Diagram for the Complement of Event *A*



A is yellow, $\bar{A}$ is pink

Rules for Complementary Events

$$P(A) + P(\bar{A}) = 1$$

$$P(\bar{A}) = 1 - P(A)$$

$$P(A) = 1 - P(\bar{A})$$

Fundamental Counting Rule (multiplication rule)

For a sequence of two events in which the first event can occur m ways and the second event can occur n ways, the events together can occur a total of $m \cdot n$ ways.

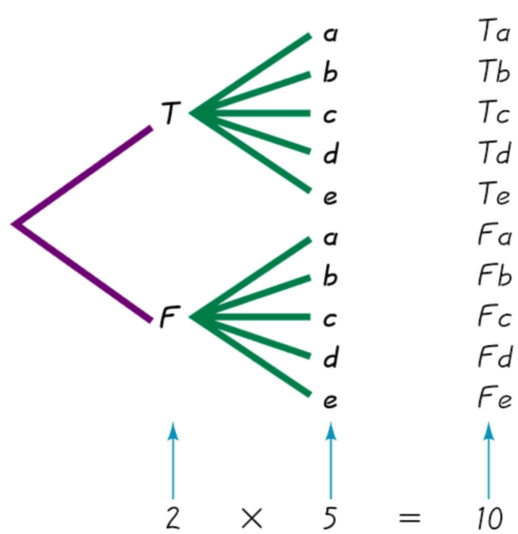
Example: a test consisting of a true/false question followed by a multiple choice question, where the choices are a,b,c,d,e.

Then there are 2 x 5 = 10 possible combinations.

Tree Diagrams

This figure summarizes the possible outcomes for a true/false question followed by a multiple choice question.

Total: 10 possible combinations.



Key Concept

The basic multiplication rule is used for finding *P*(*A* and *B*), the probability that event *A* occurs in a first trial and event *B* occurs in a second trial.

Notation

***P*(*A* and *B*) =**

***P*(event *A* occurs in a first trial and event *B* occurs in a second trial)**

Dependent and Independent

Two events *A* and *B* are independent if the occurrence of one does not affect the *probability* of the occurrence of the other.

If *A* and *B* are not independent, they are said to be dependent.

Formal Multiplication Rule for Independent Events

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

Only if *A* and *B* are independent

Dependent Events

Two events are dependent if the occurrence of one of them affects the *probability* of the occurrence of the other.

(But this does not necessarily mean that one of the events is a *cause* of the other.)

Key Concept

If the outcome of the first event *A* somehow affects the probability of the second event *B*, it is important to adjust the probability of *B* to reflect the occurrence of event *A*.

Conditional Probability Important Principle

The probability for the second event *B* should take into account the fact that the first event *A* has already occurred.

Notation for Conditional Probability

***P*(*B*|*A*) represents the probability of event *B* occurring after it is assumed that event *A* <u>has already occurred</u>**

(read *B*|*A* as "*B* given *A*")

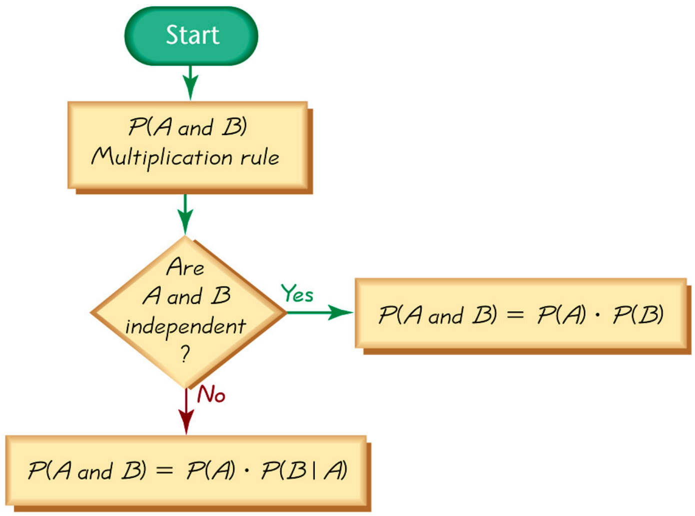
Formal Multiplication Rule for Dependent Events

- $P(A \text{ and } B) = P(A) \cdot P(B|A)$
- Note that if A and B are independent events, P(B|A) is really the same as P(B).

Intuitive Multiplication Rule

When finding the probability that event *A* occurs in one trial and event *B* occurs in the next trial, multiply the probability of event *A* by the probability of event *B*, but be sure that the probability of event *B* takes into account the previous occurrence of event *A*.

Applying the Multiplication Rule



Multiplication Rule for Several Independent Events

In general, the probability of any sequence of independent events is simply the product of their corresponding probabilities.

Selections with replacement and without replacement

Selections with replacement are always independent.

Selections without replacement are always dependent.

Treating Dependent Events as Independent

Some calculations are awkward, but they can be made manageable by using the common practice of treating events as independent when small samples are drawn from large populations. In such cases, it is rare to select the same item twice.

The 5% Guideline for Cumbersome Calculations

If a sample size is no more than 5% of the size of the population, treat the selections as being independent (even if the selections are made without replacement, so they are technically dependent).

Principle of Redundancy

One design feature contributing to reliability is the use of redundancy, whereby critical components are duplicated so that if one fails, the other will work.
For example, single-engine aircraft now have two independent electrical systems so that if one electrical system fails, the other can continue to work so that the engine does not fail.

Computing the probability of "at least" one event

Find the probability that among several trials, we get at least one of some specified event.

"At least one" is equivalent to "one or more"

The complement of getting at least one event of a particular type is that you get no events of that type

(*either* none *or* at least one)

To find the probability of at least one of something, calculate the probability of none, then subtract that result from 1. That is,

$$P(\text{at least one}) = 1 - P(\text{none})$$