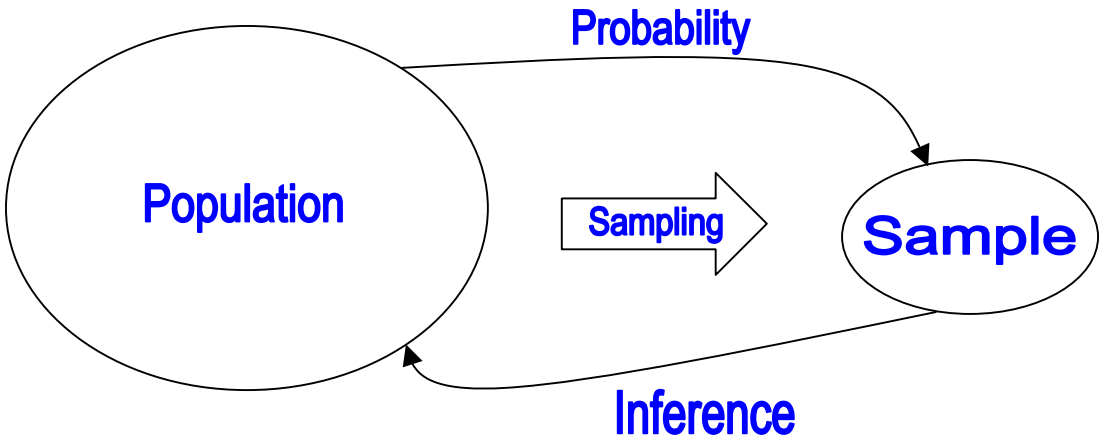
6. Probability

We have looked at methods for summarizing our data. We would like to use our data to say something about the population from which it was obtained. That is, we would like to infer characteristics of the population by using our sample data. First, we need some tools from the field of probability.



Roughly, probability asks questions like: Given a population with known characteristics, what can I say about subjects selected from the population? Statistics asks: given that I have selected a subset of subjects from the population (the sample), what can I say about the population?

We'll discuss probability in a somewhat "loose" fashion. It is beyond the scope of this class to pursue the technical details of probability. Our goal is to establish some basic notions and fundamental operations of probability that will help us later.

6.1. Operations on Events and Probability

First, some terminology:

- We will consider an experiment as the process by which we observe a data point from objects (subject, experimental unit). The result of one application of the experiment will be called an outcome.

E.g. In a study to test the effectiveness of a drug for treating headaches, you administer the drug or placebo (e.g. Sugar pill) to a patient (subject, experimental unit, unit) and, after some specified period of time you ask the patient if they experienced headache relief. Say you simply record the patient's response (the outcome) as "Yes" or "No".

- The sample space (denoted S) is the set of all possible outcomes of an experiment/measurement/observation.

E.g. Ask a patient if they smoke: S = {all possible outcomes of this experiment} = {smoke, not smoke}.

E.g. Measure the amount of zinc (μg/dl) in a patient's bloodstream S=[0,∞)

E.g. Toss a coin and observe the outcome: S={H, T}. Toss a coin twice: S={HH, TT, HT, TH}

- Event-An event is any possible collection of outcomes of an experiment; that is, any subset of the sample space S. We may use a verbal description for an event and will use

capital letters to denote events. We say an event "has occurred" if the outcome of an experiment is contained in the event.

E.g. A={observe at least 1 tail when tossing a coin twice}={HT, TT}

E.g. Test for breast cancer in women patients from ages 35 to 40 years. Event A={a woman tests positive for breast cancer}. Event B={a woman does not test positive for breast cancer}. Note that B is simply the complement of A, or B is "not A".

Three basic operations on events (set operations):

- Complement of an event A is denoted A^C (read "A complement" or "not A"); or sometimes $\overline{A}$ (read "A-bar").
- The union of event A with event B is denoted $A \cup B$. Includes outcomes in A or in B or in both. Note $A \cup A^C = S$ (always for any event A).
- The intersection of two events A and B is denoted $A \cap B$. Includes only outcomes in both A and B

E.g. Tossing a die: A={observe 3 dots or greater}={3,4,5,6} and B={observe an even number of dots} = {2,4,6}. $A \cup B$ = $\{3,4,5,6\} \cup \{2,4,6\} = \{2,3,4,5,6\}$.

E.g. Die toss continued: $A \cap B = \{3,4,5,6\} \cap \{2,4,6\} = \{4,6\}$

- The null event or impossible event is denoted $\varnothing$. Doesn't include any outcomes--it's empty. Note that $A \cap A^C = \varnothing$ for any event A.

- Subset. An event A is a subset of event B, denoted $A \subseteq B$, if all outcomes that are in A are also in B. Note that $\varnothing \subseteq A$ by definition for any event A. You can also read $\subseteq$ as "is contained in.")

E.g. Breast cancer continued. Also, ask each patient if they smoke. Let A={woman tests positive for breast cancer} and let B={woman smokes}. Then $A \cap B$ is {woman smokes and tests positive for breast cancer} and is depicted using Venn diagrams as

- Frequentist definition of probability: If an experiment is repeated n times under identical conditions (i.e., on the same population of subjects or units), and if the event A occurs m times, then as n grows large, the ratio m/n approaches a fixed limit that is the probability of A.
  - P(A) = m/n
  - Interpretation: the relative frequency (proportion) of occurrence of an event in a large number of experiments or observations.
  - Intuitively, we see that probabilities should be between 0 and 1 (inclusive).
  - We can think of probability as the chance of seeing an event occur ("subjective" probability interpretation).

E.g. Tossing a fair coin.
http://www2.SPSU.edu/math/deng/m2260/stat/coin/TossCoin.html

Tossing a die.
http://www2.SPSU.edu/math/deng/m2260/stat/roll/RollDice.html

- Some basic rules of probability:
  - $0 \le P(A) \le 1$ for any event A
  - P(all events) = P(union of all possible events) = P(S) = 1
  - P(no events) = $P(\varnothing)$
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - If $(A \cap B) = \varnothing$, then $P(A \cap B) = P(\varnothing) = 0$, and we say A and B are disjoint or mutually exclusive events.
  - In the case of mutually exclusive events (or, more specifically, if $P(A \cap B) = 0$), we have the additive rule of probability:
    - $P(A \cup B) = P(A) + P(B)$
  - $P(A \cup A^C) = P(A) + P(A^C)$

Think about the above rules using Venn diagrams:
(See applet:
http://www2.SPSU.edu/math/deng/m2260/stat/venn2/prob.html)

6.2. Conditional Probability

E.g. Breast cancer continued: Say we know the event B={woman smokes} has occurred. Given that a woman smokes (i.e. given event B), what's the probability of the event A={woman tests positive for breast cancer}?
Use a Venn diagram to think about this.

- The conditional probability of an event A given an event B has occurred is denoted/defined as $P(A|B)=P(A\cap B)/P(B)$ (assuming P(B) is not zero!)
- The multiplicative rule of probability says $P(A\cap B)=P(A)P(B|A)=P(B)P(A|B)$.
- Note that we can get the conditional probability using the multiplicative rule (as long as the denominator is not zero!).

E.g. Assuming the life table (Table 5.1) on page 98 of POB describes the entire US population in 1992, we see that 92562 out of 100000 people live to be at least 50 years of age. Let A={a person lives to at least 50 years}. Let B={a person lives to at least 70 years}. Then, according to the frequentist definition of probability we have P(A) = 92562/100000 = 0.92562 and P(B) = 72063/100000 = 0.72063.

What's the probability of A∩B?

Given that a person lived to at least 50 years, what's the probability that the person will live to at least 70?

Check out this simple applet demonstrating conditional probability:
http://www2.SPSU.edu/math/deng/m2260/stat/cond/cond.html

- When one event has no effect on the probability of another event (and vice versa), we say the events are independent. We define independence as $P(A \cap B) = P(A)P(B)$. This means $P(A|B) = P(A)$ (and $P(B|A) = P(B)$). That is, the probability of A is the same whether we know B or not, and vice versa. In other words, the fact that an event A does or does not occur has no bearing on the chance of seeing event B occur.

E.g. Toss coin.

Some questions:

If events are mutually exclusive are they necessarily independent?

If events are independent, are they necessarily mutually exclusive?

E.g. Natality: Mothers Age distribution for Births in 1992

Age	Probability
<15	0.003
15-19	0.124
20-24	0.263
25-29	0.290
30-34	0.220
35-39	0.085
40-44	0.014
45-50	0.001

What's the probability that a randomly selected woman who gave birth in 1992 was age 19 or younger?

What's the probability that she was 45 or older? 50 or older?

Given that a mother was 34 years or younger, what's the probability of her having been a teenager (or younger!)?

If you randomly select two mothers, what's the probability that they were both 20-24 years of age?