Stat 13, Tue 4/24/12.

1. Review list.
2. Examples of studies.
3. Summary of probability rules.
4. General probability strategy.
5. Probability examples.

Hw3 is due Thur, 4/26, and Midterm 1 is Thur, 4/26.
You can use calculators, pen or pencil, and one 8.5x11 page of notes, double sided, for the exam. I rarely answer questions during the exam.

1. Review list.

a) Study design concepts

- causation
- explanatory and response variables
- treatment and control groups
- longitudinal and cross-sectional studies
- experiments and observational studies
- confounding factors
- statistical significance
- coverage
- random sampling and random assignment to treatment or control
- blinding
- importance of randomization, portacaval shunt example
- adherer bias, clofibrate example

b) analysis of one variable

- stem-leaf plot
- histogram, relative frequency histogram, area <--> probability
- describing distributions: symmetry, skew, peaks, gaps, outliers, normal
- mean, median, MAD, s, s^2, IQR, range,
- sensitivity to outliers, transformations, parameters and statistics
- interpreting the sd and the 68%/95% rule of thumb

1. Review list.

c) Probability

- $P(A \mid B)$
- disjoint events
- independent events
- or
- axioms of probability
- addition rules
- multiplication rules
- probability trees
- 1 minus trick
- mean or expected value of a random variable
- combinations

2. Studies.

Obesity and autism example.
Observational study or experiment?

Gender and mathematical toys for kids.
Explanatory and response variables? Obs. study or experiment?

Researchers in New England, attempting to study the causes of obesity, obtained a simple random sample of 2500 Americans of ages 30-60 years. Subjects were asked to fill out a questionnaire asking several questions pertaining to the subjects' diets and lifestyles. Each subject's weight was also recorded. The researchers found that the subjects who stated that they shower more than 10 times per week weighed less, on average, compared to those who shower 5-10 times per week or those who shower fewer than 5 times per week. The researchers claimed that their results suggest that Americans should shower more, and hypothesized that frequent showering ``may be a healthy way to energize one's circulatory system", and ``may also cleanse the skin of bacteria which could contribute to obesity". Comment on the researchers' conclusions. What are the explanatory and response variables? Can you think of a specific, alternative explanation for the results of their study?

3) Summary of probability rules, for 2 events.

a) if A and B are disjoint, then

a1) P(A or B) = P(A) + P(B).

a2) P(A and B) = 0.

a3) P(A|B) = 0.

b) if A and B are ind., then

b1) P(A or B) = P(A) + P(B) – P(A)P(B).

Also = 1 – {1 – P(A)}{1 – P(B)}.

b2) P(A and B) = P(A)P(B).

b3) P(A|B) = P(A).

c) in general,

c1) P(A or B) = P(A) + P(B) – P(A and B)

also = 1 – P(A^c B^c).

c2) P(A and B) = P(A) P(B | A).

c3) P(A | B) = P(A and B)/P(B).

4. General strategy for probability problems.

a) Phrase the question in terms of "OR" or "AND". Often it's useful here to break down the problem by conceptually *numbering* outcomes.

b) Decide which events are disjoint and which are independent.

c) Use the rules above.

5) Probability examples.

Roll two dice. Probability at least one is even. (or both!)

Number dice 1 and 2. We want P(first is even OR second is even).

These are independent.

So P(1st is even OR 2nd is even)

= P(1st is even) + P(2nd is even) – P(1st is even)P(2nd is even)

= 1/2 + 1/2 – (1/2)(1/2) = 3/4.

Probability that one is even, *but not both*?

See above, but subtract P(both even) which is 1/4. So, the answer is 1/2.

Flip two coins. What is P(neither is head)?

P(neither is head) = P(1st is tails AND 2nd is tails)

independent.

= P(1st is tails) x P(2nd is tails)

= $\frac{1}{2}$ x $\frac{1}{2}$ = $\frac{1}{4}$.

Probability rules for 3 or more events?

Suppose you have 4 events: A, B, C, D (e.g. rolling 4 dice, or 4 coins).

It should be obvious from the following rules what to do with 3, or 5, or some other number of events.

a) if disjoint, then

a1) P(A or B or C or D) = P(A) + P(B) + P(C) + P(D).

a2) P(A and B) = 0.

a3) P(A|B,C, and D) = 0.

b) if ind., then

b1) P(A or B or C or D) = 1 – [1 – P(A)][1 – P(B)][1 – P(C)][1 – P(D)].

b2) P(A and B and C and D) = P(A)P(B)P(C)P(D).

b3) P(A | BCD) = P(A).

c) in general,

c1) P(A or B or C or D) = 1 – P(A^c B^c C^c D^c).

c2) P(A and B and C and D) = P(A) P(B|A) P(C | AB) P(D | ABC).

c3) P(A | BCD) = P(ABCD)/P(BCD).

Roll 3 dice. P(at least one 5)?
Imagine numbering the dice.
P(at least one 5) = 1 – P(none is a 5)
= 1 – P(1^{st} isn't 5 and 2^{nd} isn't 5 and 3^{rd} isn't 5)
independent
= 1 – P(1^{st} isn't 5)P(2^{nd} isn't 5)P(3^{rd} isn't 5)
= 1 – (5/6)(5/6)(5/6)
= 1 – 125/216
= 91/216.

Draw 3 cards, without replacement. P(all 3 are ◆)?
Imagine numbering the cards.
P(all 3 are ◆) = P(1^{st} is ◆ and 2^{nd} is ◆ and 3^{rd} is ◆)
not independent!
= P(1^{st} is ◆) x P(2^{nd} is ◆ | 1^{st} is ◆) x P(3^{rd} is ◆ | 1^{st} and 2^{nd} are ◆)
= 13/52 x 12/51 x 11/50
~ 1.29%.