MA/CS 109 Lecture 21

Review: Day 2
(Midterm Wednesday)
Template for Doing Mathematics

Problem

Model--------------Repeat-------------------| Modify

Examples/Conjectures Model

Proof---Did we answer the question?---No

Yes—Fame $$$
Probability

Probability is a subject within mathematics—with all the advantages and difficulties that implies....it is the precise subject modeling the quantification of uncertainty.

Recall the vocabulary: Outcome, sample space, event.

The probability function assigns a number between 0 and 1 to each event.
Axioms of Probability

The probability of every event is between 0 and 1.

The probability of the entire sample space is 1.

If A and B are disjoint events (share no outcomes) then the probability of A or B is probability of A plus the probability of B.
Models

There are lots of ways to define a probability function that follow these axioms. You must provide the definition of the probability function—it is part of your model of the part of the world that you are considering.

Common model “Equally likely outcomes model”—each event made up of 1 outcome has the same probability=1/(total number of outcomes).
Simple game

Suppose you play a game where you roll a six sided die. You win if 6 faces up and lose otherwise.

Q: How many outcomes are there?
Independence

Two events A and B are independent if
\[ P(A \text{ and } B) = P(A) \times P(B). \]
Assume each roll of the die is independent of the next.

**Q:** What is the probability of winning twice in a row?

**Q:** What is the probability of losing twice in a row?

**Q:** What is the probability of winning at least once if you roll twice?
Power Ball

The current power ball game involves choosing 5 balls from a drum containing 69 balls numbered 1 to 69 and one ball (red) from a drum containing 26 balls numbered 1 to 26.

Q: How many possible outcomes are there?
Q: Using the Equally likely outcomes model, if you buy one Power Ball ticket, what is your probability of winning?

Q: If you buy 10 tickets, what is your probability of winning?

Q: What is your probability of winning the second place prize of $1,000,000 by getting all the white balls correct?
Independence

Two events A and B are independent if
\[ P(A \text{ and } B) = P(A) \times P(B). \]

Assume the event of winning (or losing) the power ball lottery this week is independent of winning next week.

**Q:** What is the probability of winning first prize both weeks?

**Q:** What is the probability of not winning first prize both weeks?

**Q:** What is the probability of winning at least once if you play two weeks in a row?
Inferential Statistics

Basic question: Can we infer (or predict) information about a population from information from a sample (subset) of the population?

Answer: Sure, but that isn’t very useful if we have no idea how much to trust our prediction
Better Question: Can we infer or predict information about a population from information from a sample AND quantify how likely our prediction is to be correct?

Amazingly…..Yes.

Key tool: Central Limit Theorem.
Q: A poll of 900 people is taken before an election to choose candidate A or B. 550 of the 900 say they will vote for A and 350 say they will vote for B. What can you conclude from this information? In particular, how sure are you that Candidate A will win (if at all)?

Q: To make the conclusions you made in the previous question, you needed to make some more assumptions about the poll. What are those assumptions?
Q: Suppose you roll a 6 sided die 100 times and sides 1 and 3 face upwards 41 times. Are you 95% confident that this die is loaded (the sides do not appear with probability 1/6)? Explain.

Q: Suppose you roll a 6 sided die 1000 times and sides 1 and 3 face upwards 410 times. Are you 95% confident that this die is “loaded”. Explain?
Q: Suppose a poll is conducted and 44% of the sample says it will vote “Yes”, but the margin of error is 9%. If we want to increase the sample size so that the margin of error is only 3%, how much larger will the sample size need to be?
Association:

Using statistics, we can show, at least up to a certain degree of confidence, that two properties of a population are associated with each other. That is, that knowing one of the properties helps you predict the other.
Suppose we study the relationship between having red hair and green eyes. We take a large sample of Americans (chosen with the Equally likely outcomes model) and get

<table>
<thead>
<tr>
<th></th>
<th>Green Eyes</th>
<th>Not Green Eyes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Hair</td>
<td>893</td>
<td>2421</td>
</tr>
<tr>
<td>Not Red Hair</td>
<td>635</td>
<td>4721</td>
</tr>
</tbody>
</table>
We “do the math” (i.e., do the arithmetic...ok have the computer do the arithmetic) and get that the log-odds-ratio is 1.01

The margin of error at the 95% confidence level is 0.115

**Q:** Say precisely what this information allows us to conclude?

**Q:** Are we confident that red hair causes green eyes? If not, what are we confident of (if anything?)
Question I would like to ask...

Discuss the following from the point of view of what you know about statistics...

http://abcnews.go.com/Politics/donald-trump-leads-expectations-shows-strength-attributes-poll/story?id=34599211