7 - 26 - 10

The 1 hr 50 min midterm exam Wednesday, 7-28-10 will be closed book, no notes or extra papers or electronics in use or in view (except a calculator). A normal table will be provided.

For Monday 26th we will work through some exercises (below). You will not hand in solutions. There will then be a short bonus quiz on the material (to date) (bonus applicable to midterm grade).

Rules of Probability.

1. Box 1 contains 4 R and 6 G balls.
   Box 2 contains 7 R and 3 G balls.
   A choice of box is made.
   P(box 1 is chosen) = 0.1
   P(box 2 is chosen) = 0.9

A ball is then selected with equal probability from the chosen box.

Use the rules of probability to obtain the following:

a. $P(R \mid_{IF} \text{Box } 1)$ (from the assumptions)

b. P(Box 1 and R) = P(Box 1 and R) (multiplication rule)

c. P(R) = P(Box 1 and R) + P(Box 2 and R)

d. $P(\text{Box } 1 \mid_{IF} R) = \dfrac{P(\text{Box 1 and R})}{P(R)}$

This uses the fact that for any events A, B we have

$P(B \mid_{IF} A) = P(A \text{ and } B) / P(A)$.

(a variation on the multiplication rule). Here

A = "the selected ball is red"

B = "Box 1 is selected"

e. Has the probability of Box 1 having been selected increased or decreased upon learning that a red ball has been selected from the chosen box? Does this result seem sensible? Why?

Importance. Part (d) is the revised chance we quote for Box 1 having been selected upon learning a red ball was ultimately chosen. Probability is being used here to **reason backwards** from observed results (in this case getting a red ball) back to the likelihood of causes (in this case having chosen Box 1). This remarkable and useful idea is most often credited to Rev. Thomas Bayes .

From the website of the International Society for Bayesian Analysis:

" What is Bayesian Analysis?

Scientific inquiry is an iterative process of integrating and accumulating information. Investigators assess the current state of knowledge regarding the issue of interest, gather new data to address remaining questions, and then update and refine their understanding to incorporate both new and old data. Bayesian inference provides a logical, quantitative framework for this process. It has been applied in a multitude of scientific, technological, and policy settings."

2. Suppose the following probabilities apply to the given events
   - OIL (means oil is present at a prospective site)
   - \+ (means a test comes back positive for oil)
   - \- (means a test comes back negative for oil)

   $P(OIL) = 0.2$ (the probability of oil prior to testing)
   $P(+ \mid_{IF} OIL) = 0.9$ (if oil is present a positive test is likely)
   $P(- \mid_{IF} OIL^C) = 0.7$ (if oil not present a negative test is likely)

a. $P(OIL\ +) = P(OIL)\ P(+ \mid_{IF} OIL)$ (multiplication rule)

b. $P(+) = P(OIL\ +) + P(OIL^C\ +)$

c. $P(OIL \mid_{IF} +) = \dfrac{P(OIL\ +)}{P(+)}$

Has the probability of OIL increased from its initial value 0.2 upon learning that a test for oil has come back positive? Does this seem reasonable? Why?

3. An investment produces random return X with the following probability distribution:

x	0	3	5
p(x)	0.8	0.1	0.1

a. E X

b. E X^2

c. Variance X

d. Standard deviation σ_X

e. E(total of 400 independent plays of this investment)

f. Variance(total of 400 independent plays of this investment)

g. Standard deviation of total of 400 independent plays.

h. Sketch the approximate normal distribution of the total of 400 independent plays, labeling the mean and standard deviation of this normal as recognizable elements of your sketch.

i. Standard score of a total of 353 for 400 independent plays.

$$z = \frac{353 - E(\text{total of } 400)}{\text{standard deviation of total of } 400}$$

j. Use (i) and z-table to approximate

P(total of 400 independent plays < 353).