Midterm 2 Solutions

Stats M12, Spring 2001

1.A lottery "scratch-off" ticket costs $1. You scratch off part of the ticket to discover your prize. The probability that you win nothing at all is 80%. However, there is a 15% chance that you win your dollar back, and a 5% chance that you win 10$.

Let X represent the amount you win on a single ticket.

a) Give the pdf of X in a table. In the first column of the table, list the possible outcomes of X, and in the second table give their probabilities.

x	**p(x)**
-1	.8
0	.15
9	.05

b) Sketch the pdf of X. The x-axis should show the values of x, clearly labelled. The y-axis should show the probabilities.

Can't do this on the web. Graph should have spikes at -1 (height .8), 0 (height .15) and 9 (height .05). Some people connected the spikes, and this is wrong because it suggests that the probability of getting a value other than -1,0 or 9 is possible. In fact, these are the only three values that are possible, and all other values have probability 0.

c) Find the expected value, E(X), of X.
E(X) = -1 * .8 + 0 * .15 + 9* .05 = -$ 0.35.

d) Find the standard deviation of X.
SD(X) = sqrt($(-1 - -.35)^2$ * .8 + $(0 - -.35)^2$ * .15 + $(9 - -.35)^2$ * .05) = $2.174

e) Let Y represent the amount of money you win on two tickets. (That is, you purchase two tickets and Y is the total winnings.) What's the probability that you will lose $2? (Recall that you lose one dollar with a probability of 0.80.)

P(lose two of two)= P(lose first AND lose second) = P(lose first) * P(lose second) (because they are independent events) = .80 * .80 = .064.

f) What are the expected value and SD of Y?
Let X1 be the amount you win on first play, and X2 on the second.
E(Y) = E(X1 + X2) = E(X1) + E(X2) = 2 * -.35 = -$0.70.
SD(Y) = sqrt(2) * SD(X1) = $3.0745

g) Suppose you buy 400 tickets. Let Y represent the total winnings. What's the expected value **and** SD of Y?

E(Y) = 400*E(X) = -$140.00
SD(Y) = sqrt(400)*SD(X) = $43.48

i) Sketch the approximate pdf of Y. (Hint: use the central limit theorem.) Indicate on your sketch where the mean minus one SD is, and also where the mean plus one SD is.

Sketch should look like a normal curve, centered at -140.

j) Using the central limit theorem, find the approximate probability that you do not lose money on 400 tickets.

P(Y >= 0) = P(Z >= (0 - -140)/43.48) = P(Z > 3.22) = 1 - P(Z < 3.22) = 1 - .9987 = .0013.

2. In a census district that includes Beverly Hills, the U.S. 2000 census found that 37% of the residents have at least one year of college. Suppose we take a random sample of 100 residents from this census district.

a) What percentage of our sample do we expect to have at least one year of college?

.37 or 37%.

b) What's the standard deviation for the percent of people in our sample that have at least one year of college?
SD(Phat) = sqrt (p * (1-p)/n) = sqrt(.37 * .63/100) = .0482 or 4.82%.

c) Assuming that the percentage of people in our sample with at least a year of college follows a normal distribution, what's the probability that more than 42% of our sample will have had at least one year of college.

P(phat >= 0.42) = P(Z >= (.42 - .37)/.0482) = P(Z >= 1.037) = 1 - P(Z < 1.04) = 1 - .8508 = .1492.

3. The fictional town of Lake Wobegone claims that "all of tour children are smarter than average." If this is true, the residents should be better educated than the residents of Beverly Hills. A survey was conducted to see if this is the case. 100 residents of Lake Wobegone were selected at random. It turned out that 48% of the *sample* had at least one year of college. Carry out a hypothesis test by answering the questions below. Recall that in Beverly Hills, 37% of the residents have at least one year of college.

a) State the null and alternative hypotheses.
H0: p = .37

Ha: $p > .37$

b) If the null hypothesis is true, in a random sample of 100 residents of Lake Wobegone, what percent do we expect to have had at least one year of college education? What's the SD for this percent?
E(phat) = .37
SD(phat) = sqrt(.37 * .63/100) = 0.0482.

c) Assuming that the sample proportion follows a normal distribution, find the p-value for this study.

P(phat > .48) = P(Z > .48 - .37/.0482) = P(Z > 2.28) = 1 - P(Z < 2.28) = 1 - .9987 = 0.0013.

d) Using a significance level of 10%, do you reject the null hypothesis?

Yes. p-value is much less than 0.10.

4. A roullette wheel is suspected of being off-balance. If it were balanced, it would land on red with probability 18/38, land on black with probability 18/38, and land on green with probability 2/38. The wheel is spun 1000 times, and the table gives the number of times each of these colors came up. (Do not round off.)

Color	Observed Number of occurances	Expected Number of occurences	Difference
Red	490	1000*(18/38) = 473.684	16.316
Black	462	473.684	-11.684
Green	48	1000*(2/38) = 52.636	-4.636

a) Fill in the missing columns of the table.

b) Find the observed value of the chi-squared test statistic for these data.
$X^2 = (16.316)^2/473.684 + 11.668^2/473.684 + 4.6316^2/52.636 = 1.25696$

c) The test statistic has a pdf that is approximately chi-square with 2 degrees of freedom. Based on these data, would you conclude the roullette wheel is fair? Explain.

There's insufficient evidence to conclude that the wheel is off-balance. From the table, for a chi-square distribution with 2 degrees of freedom, the probability to the right of 5.99 is 5%. Since our observed value is less than this, our pvalue, $P(X^2 > 1.26)$ is more than 5%, and we would not reject using a 5% significance level.