Spring 2005
Stat 1601 Name: ________________________
Midterm 2

For questions 1-3 use the following information. Suppose a basketball player typically makes 80 percent of her free throws.

1. If the player tries 50 free throws, what number is the biggest, but yet still reasonable, guess for the number of free throws she will make?

Ans. A binomial rv X has $E(X) = np = 50 * .8 = 40$ the sd is $sd(X) = \sqrt{np(1-p)} = 2.828$ so an upper bound would be $E(X) + 2 * sd(X) = 40 + 2(2.828) = 40 + 5.656 = 45.656$ or 46 successes. I will accept a range between 45 and 46. No credit for other answers.

2. What is the probability that the player only makes one free throw in three tosses?

Ans. A probability tree or a reverse binomial table $n = 3, p = .2, Y = 2$ look up gives solution .0960. Using a Z-score, p-hat approach is not as appropriate as an exact solution.

3. Which of the following events is most likely? Circle correct letter.

(a) Our basketball player making more than 85 percent of her free throws in 60 tosses.

(b) Observing less than 40 percent heads when tossing a coin 200 times.

(c) Winning roulette twice in two consecutive turns by betting on red in a Las Vegas casino that has two green numbers on a 38 slot wheel.

(d) Winning the Powerball lottery jackpot after buying 10 tickets with different numbers. Winning means having a ticket where all numbers match the winning numbers.

Ans. Solution is (c), chance equals $(18/38) * (18/38) = .2244$. Solution to (a),$P(Z > .85 - .8/sd(\hat{p}) = P(Z > .97) = .1660$. Solution to (b) is $P(Z < -.1/\sqrt{.25/200}) = P(Z < -2.83) = small$. Option (d)? Forget about it !

For questions 4-7, use the following information: A recent Gallup poll found that 54 percent of Americans believe the US government spends too much on foreign aid. The sample size was 1008 adults.

4. What is the margin of error for a 95 percent confidence interval using these data? Express your answer in terms of percentage points, and use the best two digits of accuracy past the decimal point.

Ans. $m = 1.96\sqrt{.54(1 - .54)/1008} = .0308$, so answer is 3.08 percentage points.

5. Compute a 95 percent confidence interval for the population proportion that believe the US government spends too much on foreign aid. Give the upper confidence limit here:

Ans. $\hat{p} + 1.96\sqrt{.54(1 - .54)/1008} = .5708$, I'll accept .5707 to .5708.

6. Do you think the proportion of adults that believe the US gives too much foreign aid could be as high as 60 percent? Why? Explain.

Ans. Nope. All reasonable values for the true proportion are below .6, so there is evidence the true proportion could not be as high as .6 - the truth must be lower. A true proportion of $p = .6$ could not have produced a sample proportion like .54 in the study.

7. In another poll next month, how many subjects would be needed to obtain a margin of error of one percentage point in a 95 percent confidence interval?

Ans. $n = (1.96/.01)^2(.54)(1 - .54) = 9542.5344$ so a sample size of 9543 is needed.

For problems **8–11** use the following information:
The table below gives the Hepatitis C status and the tattoo status of 626 subjects.

Tattoo Status	Hep. C Status		
	Hepatitis C	No Hep. C	Total
Tattoo, Parlor	17	35	52
Tattoo, Elsewhere	8	53	61
No Tattoo	22	491	513
Total	47	579	626

8. What is the probability of reaching into this sample and selecting at random, a subject with Hepatitis C?

Ans. $P(HepC) = 47/626 = .07508$.

9. Are the events Tattoo, Parlor (a tattoo at a parlor) and Hepatitis C independent?

(a) Yes

(b) No

Ans. Nope. $P(TattooP \mid HepC) \neq P(TattooP)$, $17/47 > 52/626$.

10. What is the probability someone got a tattoo (at either a parlor or elsewhere) given the subject has Hepatitis C?

Ans. $P(TattooE \mid HepC) = 25/47 = .53191$.

11. What is the probability a person has Hepatitis C given they do not have at tattoo?

Ans. $P(HepC \mid None) = 22/513 = .0429$.

12. Suppose a casino game is such that a player wins \$2 with probability .1 and \$10 with probability .05. Suppose it costs \$1 to play this game. What is the expected net gain of playing this game?

Ans. From the table below we find the expected payoff, $E(X) = .7$, so the expected net gain is $.7 - 1 = -.3$, or negative .3 dollars.

X	P(X)	X*P(X)
0	.85	0
2	.10	.20
10	.05	.50
		$E(X) = .70$

13. Suppose we play the casino game described in the previous question, 2 times. What is the probability of making a profit? A profit means that at the end of three plays, more money is made than was spent playing the game. Three plays are required.

Ans. A probability tree with two sets of three payoff branches shows a profit will occur at outcomes (0,10),(2,2),(2,10),(10,0),(10,2),(1
The probabilities of these outcomes are: .0425, .01, .005, .0425, .005, .0025. Final probability is found by summing the values, and gives solution .1075.

14. A randomized response survey was conducted so that a subject rolls a die and if the roll is a 5 the subject rolls the die again and if the roll is a 2 the subject answers yes. If the subsequent roll is not a 2 the subject answers no. If the original roll was 1,2,3, 4 or 6 the subject answers a sensitive question about sexual behavior. The survey found 45 yes answers out of 130 subjects. What your estimate of the proportion of the population that have the sexual behavior? Show your work.

Ans. Tree solution, first branches 5 and not five. For the five branch the next branches are 2 and not 2. For the first not 5 path, branches yes and no answers with chance p and 1-p. Solution matches the prob of yes from the tree, $(5/6)p + 1/36 = 45/130$ implies $p = (6/5) * [(45/130) - (1/36)] = .38205$.

15. A college professor hears that roughly 44 percent of college students in the US engage in binge drinking. A study of 244 students at her college finds 96 admit to binge drinking. Is this a surprising result? Is her college unusual? Why or why not?

Ans. The observed proportion is $\hat{p} = 96/244 = .3934$, $sd(\hat{p}) = .0318$ if the USA proportion holds in this case. Our observed proportion is $Z = -1.465$ standard deviations below what would be expected if the USA proportion of .44 holds for this university. A sample proportion like this, or smaller could occur with chance approx .0722. This is not a very unusual result, just a bit unusual. The proportion in her school is a bit less that would be expected based on the national average. It is not something so unusual it should be reported to anyone.

This problem could also be solved by doing a confidence interval for the proportion binge drinking in just this school and then making an argument how this differs from the national proportion. The CI for the school proportion is (.332, .455), this is not suggesting a different underlying proportion than .44. This means the observed result does not suggest a different underlying proportion than the US prop of .44, and so the observe proportion should not be considered unusual.

This problem could also be solved as a binomial problem where the number of students that binge drink from a sample of 244 is the random variable. If the US average is true, $X = 96$ should be compared with the average of a binomial of $E(X) = np = 244 * .44 = 107.36$ would be expected, with typical variation of 7.75. Our observed X=96 is not even 2 sds below what is expected from a school with the US proportion of .44. This is not a very unusual outcome.

16. Suppose a study of body temperatures for healthy people is conducted with $n = 52$ measurements and a sample average of 98.285 degrees is obtained with a standard deviation of .6824 degrees. From these data is it reasonable to assume that the average body temperature for a healthy person is 98.6 degrees? Why or why not?

Ans. A confidence interval for mu gives: $98.285 \pm 1.96 * (.6824/\sqrt{52})$ or $98.285 \pm .1855$, or (98.099, 98.470). This means that a true mean value of 98.6 degrees is not a reasonable statement for the population mean body temp. The true mean value is certainly less than 98.6 degrees.

17. If we presume that the true average body temperature for healthy people is 98.6 degrees, and the standard deviation is .6824 degrees, what body temperature should we use as an indicator of an abnormally high body temperature? Suggest a value, and give your justification for why you chose that value.

Ans. I think two standard deviations above 98.6 should be used as a cut-off. This would be 99.965 based on $98.6 + 2 * .6824$.

18. Let's presume that 50 percent of all Reese's Pieces have orange color. What is the probability of observing a mini-bag of 15 candies with more than 10 orange candies?

Ans. A binomial rv with $n = 15$ and $p = .5$ gives $P(X > 10) = .0417 + .0139 + .0032 + .0005 = .0593$.

19. Suppose a population has mean $\mu = 30$ and standard deviation $\sigma = 10$. Could an average of 34 reasonably occur from a sample of 144 observations? Provide your reasoning.

Ans. You are provided the population information and are being asked about the behavior of a sample average. Sample average behavior is measured by the sampling distribution of x-bar, or standardizing that distribution to a standard normal Z. An average of 34 is $Z = \frac{34-30}{10/\sqrt{144}} = 4.8$ standard deviations above the mean of 30. This is a very unreasonable outcome that has almost no probability of occurring by chance.

20. What is a statistic? Give a definition.

Ans. A statistic is a number calculated from sample values.

21. If I wished to know the proportion of UMM students that spend more than 2 hours per day on the internet, and a random sample of student hours spent was available from a sample of students, what statistical method should be used to address the research question. You should provide the name of the statistical approach you will use.

Ans. This situation needs a confidence interval for the population proportion p of students who spend more than two hours daily on the internet. You are given a sample and asked about the underlying population proportion p. This calls for a CI for p.

22. A recent Gallup poll found 26 percent of Chinese adults smoke regularly from a study of about 1000 adults. It reported the margin of error was 3 percentage points with 95 percent confidence. Could the proportion of Chinese adults that smoke regularly be 20 percent? Give your reasoning.

Ans. The lower confidence interval boundary would be .26 - .03 or .23 or 23 percent. A true proportion in the population of 20 percent is not indicated by the confidence interval, and so could not reasonably be .20. All reasonable values for this truth are above 20 percent.

23. A recent Gallup poll found 26 percent of Chinese adults smoke regularly. It reported the margin of error was 3 percentage points with 95 percent confidence. What does 95 percent confidence mean?

Ans. 95 percent confidence means that the confidence interval method that was used to construct the poll CI has the property that the method will produce intervals that contain the true population proportion in approximately 95 percent of all intervals constructed. This is a property of the method used to get the interval, not a statement about this interval in particular.