MATH-1110 (DUPRÉ) SPRING 2011 TEST 3 ANSWERS

FIRST: PRINT YOUR LAST NAME IN LARGE CAPITAL LETTERS ON THE UPPER RIGHT CORNER OF EACH SHEET TURNED IN.

SECOND: PRINT YOUR FIRST NAME IN CAPITAL LETTERS DIRECTLY UNDERNEATH YOUR LAST NAME ON EACH SHEET TURNED IN.

THIRD: WRITE YOUR CORRECT SPRING 2011 MATH-1110 SECTION NUMBER DIRECTLY UNDERNEATH YOU FIRST NAME ON EACH SHEET TURNED IN.

FOURTH: THERE ARE 11 QUESTIONS AND EACH IS WORTH 9 POINTS. WRITE ALL YOUR ANSWERS NEATLY IN THE SPACE PROVIDED UNDER EACH QUESTION. NEATNESS COUNTS. IF I CANNOT READ IT WITHOUT STRAINING MY EYES YOU GET NO CREDIT.

FIFTH: Any failure to follow any part of any of the above directions will result in whatever additional loss of credit the grader wishes.

Suppose that adult weight of African elephants is normally distributed with population standard deviation 0.374 tons.

1. What is the probability that an adult African elephant weighs EXACTLY 5.8 tons?

SOLUTION: Since the normal distribution is a continuous distribution, the probability of any exact value is zero.

FINAL ANSWER: 0

2. What is the probability that an adult African elephant weighs more than the population mean weight for adult African elephants?

SOLUTION: Since the normal distribution is symmetric about its mean, the probability of being above the population mean is always 0.5 in ANY normal population or any normal distribution.

FINAL ANSWER: 0.5

3. If the true mean weight of adult African elephants is 6 tons, then what is the probability that the **TOTAL** weight of 4 randomly chosen adult African elephants is between 24.3 tons and 25.2 tons?

SOLUTION: Since the total of an independent random sample of size n from a normal population is also normally distributed with mean μn and standard deviation $\sigma\sqrt{n}$, the probability here is simply

$$P(24.3 \le T_4 \le 25.2) = \text{normalcdf}\left(24.3, 25.2, 6 \cdot 4, .374 \cdot \sqrt{(4)}\right) = 0.2898581327 \text{ or about } 0.290,$$

to three significant digits.

FINAL ANSWER: 0.290

4. If the true mean weight of adult African elephants is 6 tons, then what is the value of x such that the probability an adult African elephant weighs **MORE** than x is 0.1?

SOLUTION: If the probability that an adult African elephant weighs more than x is 0.1, then the probability it weighs no more than x is 0.9, and therefore we can find the value of x from the inverse normal in the distribution menu of the calculator.

$$x = \text{invNorm}(.9, 6, .374) = 6.479300286 \text{ or about } 6.48,$$

to three significant digits.

FINAL ANSWER: 6.48

5. If we take an independent random sample of 7 adult African elephants and weigh them to determine the sample mean weight, what is the **MARGIN OF ERROR** in the 99 percent confidence interval for the true population mean weight of adult African elephants?

SOLUTION: The simplest thing to do here is go right to the ZInterval and enter the information except using zero for $\bar{x}$. Then the upper limit of the reported interval in the readout is the margin of error. The readout here is $(-.3641,\ .36412)$ from which we see that the margin of error is $M = 0.36412$, or about 0.364, to three significant digits. Alternately, you can use the inverse normal in the distribution menu

$$M = \left(\text{invNorm}(.995, 0, 1) \cdot .374/\sqrt{7}\right) = \text{invNorm}\left(.995, 0, .374/\sqrt{7}\right) = 0.3641159149,$$

which is obviously the same, to three signiificant digits.

FINAL ANSWER: 0.364

You should notice here, that you do not need to know the sample mean in order to calculate the margin of error, since we actually know the population standard deviation. In fact, if we do not know the population standard deviation and someone tells us the sample standard deviation, we can use the TInterval in the TEST menu to find the margin of error by enetering zero for $\bar{x}$. Now, in general, if a problem is given in which a number seems to be missing, it is either a mistake in the statement of the problem, or the number you think is needed is actually not needed. If it is actually not needed, and your method uses the number, that will mean that it does not matter what that number is, so if you simply make up a number, you can use that number in your method to get the correct answer. Obviously, the simplest number to make up is zero. But if you like, you can say use $\bar{x} = 6$ in the ZInterval, and the readout will give $(5.6359, 6.3641)$ from which we immediately see the margin of error is

$$M = 6.3641 - 6 = 0.3641,$$

which is again, 0.364, to three significant digits.

6. If an independent random sample of 7 adult African elephants has a sample mean weight of 6.3 tons, what is the **SIGNIFICANCE** of this as evidence that the true mean adult weight of adult African elephants exceeds 6 tons?

SOLUTION: Since the word significance appears in the problem statement, this by itself should tip us off that the question is likely a hypothesis test, and careful reading shows in fact it is, since we are asked for the significance of data as evidence for something. Since we know the population standard deviation, we use the ZTest. In entering the information, we need to be careful to choose the correct alternate hypothesis. Remember, the alternate hypothesis is what we are asking the data to try and prove. Here we are asking the data to prove the mean adult weight of adult African elephants exceeds 6 tons which is $\mu > 6$. Now when we enter the information, we use $\mu_0 = 6$ as the value of the true mean under the null hypothesis, so the alternate hypothesis is here $\mu > \mu_0$, sos you highlight $> \mu_0$ in the ZTest dialogue, to choose the correct alternate hypothesis. In the readout, the significance is reported as p, since the significance is by definition a probability, namely the conditional probability of a sample of 7 elephants having a sample mean more than our result of 6.3 given that the true mean is $\mu_0 = 6$. Keep in mind that for this reason, the significance of the data is also called the P-Value of the data. Here, the result for the P-Value is 0.0169078781, or, to three significant digits, 0.0169, a reasonbly significant result since the number is near zero.

FINAL ANSWER: 0.0169

Suppose that Uncle Scrooge has filled a swimming pool with diamonds and rubies. He allows Minnie Mouse to scoop up a cup full of these jewels.

7. If exactly 40 percent of the jewels in Scrooge's swimming pool are diamonds, and if Minnie's cup contains 100 jewels, what is the probability that her cup contains no more than 36 diamonds?

SOLUTION: Notice that each jewel is either a diamond or not a diamond, a "yes-no" situation, and we are counting the diamonds. Also, we are given the exact number of jewels we are looking at, namely $n = 100$, which is a sample size, since the jewels in Minnie's cup form a sample of jewels from the swimming pool. Therefore, the distribution for the success count, if 40 percent of the jewels in the swimming pool are diamonds is the binomial distribution with $n = 100$ and $p = 0.4$. Therefore, if we let T denote the success count, then

$$P(T \leq 36) = \text{binomcdf}(100, .4, 36) = 0.238610717, \text{ or about } 0.239,$$

to three significant digits.

FINAL ANSWER: 0.239

8. How large must an independent random sample of jewels from Scrooge's swimming pool be in order that the **MARGIN OF ERROR** in a 95 percent confidence interval for the true proportion of diamonds in Scrooge's swimming pool will be no more than 0.01?

SOLUTION: We know that for any random variable X, the margin of error M in a confidence interval for μ_X is given by

$$M = \frac{z\sigma_X}{\sqrt{n}}, \; z = \text{invNorm}\left(\frac{1+C}{2}, 0, 1\right),$$

where n is the sample size and C is the level of confidence. Solving this equation for n gives

$$n = \left[\frac{z\sigma_X}{M}\right]^2 = \left[\frac{\sigma_X}{M} \cdot \text{invNorm}\left(\frac{1+C}{2}, 0, 1\right)\right]^2 = \left[\text{invNorm}\left(\frac{1+C}{2}, 0, \frac{\sigma_X}{M}\right)\right]^2.$$

In a situation where we do not know σ_X, we must find a number B which we know is at least as big as σ_X, that is for which $\sigma_X \leq B$. In the case at hand, the yes count, the basic variable we are sampling is simply the indicator that a jewel is a diamond, so taking X to be an indicator, we know $\sigma \leq 0.5$. Therefore, for the case of the jewels in the sqimming pool, to estimate the true proportion of diamonds with 95 percent confidence and have a margin of error no more than 0.01, we replace σ_X by 0.5 and replace M by 0.01. Notice that $0.5/(0.01) = 50$, so

$$n \geq [\text{invNorm}(.975, 0, 50)]^2 = 9603.647066.$$

As n must be a whole number, this means $n \geq 9604$, so the sample size must be at least 9604.

FINAL ANSWER: 9604

9. Suppose that Scrooge claims that at least 40 percent of the jewels in his swimming pool are diamonds and that Minnie takes an independent random sample of 100 jewels from Scrooge's swimming pool and only finds 32 diamonds in her sample. What is the **SIGNIFICANCE** of Minnie's sample data as evidence **AGAINST** Scrooge's claim?

SOLUTION: Scrooge is claiming that the true proportion p of diamonds in the swimming pool satisfies $p \geq .4$. Since Minnie is trying to disprove Scrooge's claim with her sample data, she is trying to get the significance of her data for proving $p < .4$. Thus, the alternate hypothesis is $p < .4$. This also means the null hypothesis is Scrooge's claim, that is $p \geq .4$. Now, this in turn means we will assume that $p_0 = .4$ is the true value of the proportion of diamonds in the swimming pool when we calculate the significance of the data. The sample size is the number of jewels Minnie examines which is $n = 100$. In her sample she found a total $T = 32$ diamonds which certainly seems lower than it should be if there really are at least 40 percent diamonds in the swimming pool. Thus, the significance of her data is

$$P(T \leq 32 | T \text{ is binomial, } n = 100,\ p_0 = .4) = \text{ binomcdf}(100, .4, 32) = 0.0615039101, \text{ or about } 0.0615,$$

to three significant digits.

FINAL ANSWER: 0.0615

You may have used the 1-PropZTest here from the calculator's TEST menu, but it is not really accurate enough to get three significant digits for such a small sample size. To use the 1-PropZTest here, you enter $p_0 = .4$ and $x = 32$ and $n = 100$. The readout will tell you $\hat{p} = x/n = 0.32$ and the value of p in the readout is the P-Value or significance of the data which is here 0.0512352053, and you see that the answers do not agree to any significant digits although they are both around the 0.5-0.6 region which for such a non-lethal question as the proportion of diamonds in a swimming pool may be reasonably significant. For comparison, if we change the sample size to $n = 100000$ and if Minnie's sample had say only 39300 diamonds, then she would have been 700 short of what is to be expected if there are at least 40 percent diamonds. Here. using the binomial distribution we find

$$P(T \leq 39300) = \text{binomcdf}(100000, .4, 39300) = 0.000003099167969,$$

whereas using the 1-PropZTest we get 0.0000031171824, that is, to three significant digits, the binomial distribution gives 0.00000310 and the 1PropZTest gives 0.00000312. We are still not in agreement in the last significant digit, but we are close together. If there had been 39800 diamonds in the sample, the two numbers are repsectively 0.988887215 and 0.0983528664. Notice that we have an enormous sample size and still only get two significant digits of agreement. In any case, I will not penalize you for using the 1-PropZTest in these situations, but you should be aware of the fact that it is not accurate. On the other hand, we sould need to know more about the way the calculator is actually calculating the binomial distribution for such large samples to judge whether it is actually accurate for these large samples. Obviously the calculator is not calculating the actual binomial distribution when you put in $n = 100000$. In fact, if n is over 10000, it is surely using some kind of approximation scheme. This does not mean it cannot be checked using the calculator, but I have not done it. To do so one would actually use the calculator's numerical integration capability for finding the area under the normal curve and then use the normal approximation to the binomial distribution, directly with the math menu. If any one of you are interested enough to do this and compare the output of the binomial distribution for large n with the numerical integration of the normal density which approximates it, let me know the result. The bottom line here is buyer beware, if great accuracy is necessary.

Suppose that Donald Duck needs to ride the trolley to go shopping. Assume that the number of trolleys arriving at his trolley stop during disjoint time intervals are independent. The Duckburg Trolley Company claims that the rate that trolleys arrive at his stop is on average, at least 6 per hour.

SOLUTION NOTE: Notice that we are counting trolleys, but the sample size is an amount of time. We are therefore dealing with a Poisson distribution. If we let T be the number of trolleys that arrive in a given specific hour of time, then we have $\mu = E(T)$. Moreover, this means that if we chop the hour into two half hour pieces, we have no reason to suspect apriori that either of those half hour intervals of time should have more trolleys than the other, so we would expect $\mu/2$ for each. In general, if we consider any interval of time of length t hours, then we expect μt trolleys during that interval of time.

10. If Donald assumes that trolleys arrive at the rate of 6 per hour, then how many does he expect to arrive during a 30 minute time interval?

SOLUTION: Since 30 minutes is exactly a half hour, if we expect 6 trolleys in a full hour, then we only expect 3 trolleys in 30 minutes.

FINAL ANSWER: 3

11. Suppose Donald is infuriated because he has to wait 30 minutes for a trolley. What is the **SIGNIFICANCE** of his having to wait so long as evidence that the trolleys arrive less frequently than claimed by the Duckburg Trolley Company?

SOLUTION: Since the Duckburg Trolley Company claims that trolleys arrive at the rate of at least 6 per hour, Donald can assume that trolleys arrive about every 10 minutes on average, so obviously, having to wait a half hour for a trolley seems too long. We let W denote the time to wait for a trolley. We expect that we only have to wait 10 minutes for a trolley if the Trolley Company is correct about its claim that trolleys arrive at least every 10 minutes on average. In hours, this means that $\mu_W = E(W) \leq 1/6$. Now, Donald wants his long wait to be data proving that $\mu_W > 1/6$. That is we have a hypothesis test and the alternate hypothesis is $\mu_W > 1/6$. To evaluate the significance of this data as evidence for this alternate hypothesis, we calculate its P-Value under the assumption $\mu = 1/6$, getting

$$\text{P-Value } = P(W \geq 1/2) = P(W > 1/2) = \text{poissoncdf}(3, 0).$$

You should notice here that as W itself is a continuous random variable we have

$$P(W \geq 1/2) = P(W > 1/2),$$

whereas the latter expression converts to the Poisson distribution since waiting for more than a half hour is the same as seeing none during the half hour you wait. Since

$$\text{P-Value} = \text{poisooncdf}(3, 0) = 0.0497870684,$$

to three significant digits, the significance of the half hour wait is 0.0498, as evidence that the trolleys arrive less often than claimed by the Duckburg Trolley Company.

FINAL ANSWER: 0.0498