MATH-11X0 (DUPRÉ) SPRING 2017 TEST 3 ANSWERS

1. PRINT YOUR LAST NAME IN THE UPPER RIGHT CORNER IN LARGE CAPITAL LETTERS.

2. PRINT YOUR FIRST NAME UNDERNEATH YOUR LAST NAME IN THE UPPER RIGHT CORNER IN CAPITAL LETTERS.

3. PRINT YOUR LAB DAY AND LAB START TIME UNDERNEATH YOUR FIRST NAME IN THE UPPER RIGHT CORNER.

4. WRITE YOUR SPRING 2017 MATH COURSE AND SECTION NUMBER UNDERNEATH YOUR LAB DAY IN THE UPPER RIGHT CORNER.

The remaining problems all use the information that follows. Suppose that Sam has a pond containing EXACTLY 1000 fish, and he knows that the fish in his pond have an average weight of 90 grams with a standard deviation of 16 grams. Joe weighs SIXTYFOUR fish from Sam's pond, one after another, WITH REPLACEMENT (or, catch and release, in fisherman terminology) but Sam does not see which fish Joe weighs. Joe does not know anything in advance about the weights of the fish in Sam's pond. Of the 1000 fish in Sam's pond, SAM KNOWS 300 are redfish and 700 are bluefish.

5. What should Sam EXPECT is the AVERAGE weight of the SIXTYFOUR fish Joe weighed?

ANSWER: 90

FINAL ANSWER: 90

6. What should Sam EXPECT is the squared error in his expected AVERAGE weight of the SIXTYFOUR fish Joe weighed?

ANSWER: The standard deviation of $\bar{X}_n$ is $\sigma_X/\sqrt{n}$, so here the standard deviation is $(16)/\sqrt{64} = (16/8 =$2. The expected squared error is the variance which is the square of this standard deviation, $\sigma_X^2/n = (16)^2/(64) = (16)^2/8^2 = ([16]/8)^2 = 2^2 = 4.$

FINAL ANSWER: $(16)^2/(64) = (16)^2/8^2 = ([16]/8)^2 = 2^2 =$**4.**

7. If Sam tells Joe that his favorite fish weighs between 70 and 100 grams, but nothing else, so Joe has no idea what fish Sam refers to, then what should Joe think is the distribution in weight of THAT FISH, that is what is the NAME of the distribution?

ANSWER: Whenever you only know the minimum and maximum possible values for an unknown, the distribution is UNIFORM.

FINAL ANSWER: UNIFORM

8. If Sam tells Joe that his favorite fish weighs between 70 and 100 grams, but nothing else, what is the expected weight of that fish for Joe?

ANSWER: For any uniform distribution, the mean is the average of the minimum and maximum possible values, which is here (100+70)/2=85.

FINAL ANSWER: (100+70)/2=85.

9. If Sam tells Joe that his favorite fish weighs between 70 and 100 grams, but nothing else, what should Joe think is the probability Sam's favorite fish weighs between 70 and 80 grams?

ANSWER: With X denoting the fish weight of Sam's favorite fish, as X is uniformly distributed between 70 and 100,

$$P(70 \leq X \leq 80) = (80 - 70)/(100 - 70) = (10)/(30) = 1/3.$$

FINAL ANSWER: $(80 - 70)/(100 - 70) = (10)/(30) = 1/3.$

10. If Sam decides to catch 100 fish from his pond, how many of these fish does Sam expect will be redfish?

ANSWER: As Sam knows that the fraction of redfish in his pond is

$$p = (300)/(1000) = 0.3,$$

in any fair sample of size n, he expects to find np redfish. Here, $n = 100$ and $p = .3$, so he expects to find $np = (100)(.3) = 30$ redfish.

FINAL ANSWER: (100)(.3)=30.

11. If Sam decides to catch 100 fish from his pond using catch and release, what is the standard deviation in the number of redfish he will get?

ANSWER: When Sam uses catch and release, then the observations are all independent, so the refish count distribution is BINOMIAL. For a sample of size n with success rate p, the standard deviation for the binomial distribution is $\sigma = \sqrt{np(1-p)}$, so here $\sigma = \sqrt{(100)(.3)(.7)} = \sqrt{21} = 4.582575695$.

FINAL ANSWER: $\sqrt{(100)(.3)(.7)} = \sqrt{21} = 4.582575695.$

12. If Sam decides to catch 100 fish from his pond WITH REPLACEMENT, what is the probability that exactly 25 of the fish turn out to be redfish?

ANSWER: With replacement of course means catch and release, so again, the successive observations are all independent and the success count is BINOMIAL. Here again, $n = 100$ is the number of trials, $p = .3$ is the success rate, so the binomial distribution formula gives for the probability of exactly $k = 25$ successes, letting T be the success count as an unknown,

$$P(T = k) = C(n, k)p^k(1 - p)^{n-k},$$

so here,

$$P(T = 25) = C(100, 25)(.3)^{25}(.7)^{75} = .0495599228.$$

FINAL ANSWER: $C(100, 25)(.3)^{25}(.7)^{75}$ =**.0495599228.**

13. If Joe found that the sample mean for his SIXTY FOUR fish is 90 grams, and if the sample standard deviation for his sample is 20 grams, then what should Joe think is the STANDARD ERROR he must use to compute the margin of error for his 95 percent confidence interval for the true mean weight of all the 1000 fish in the pond, if Joe assumes that the distribution of fish weight in the pond is normal?

ANSWER: Let X **denote fish weight. Then** $\bar{X}_n$ **is the sample mean unknown for a sample of size** n. **With independent random sampling (IRS), we know that**

$$\sigma_{\bar{X}_n} = \frac{\sigma_X}{\sqrt{n}}.$$

BUT, Joe does not know σ_X, **so he must use the sample standard deviation** $s = 20$ **in its place. So for Joe, his standard error is** $s/\sqrt{n} = (20)/\sqrt{64} = (20)/8 = (10)/4 = 5/2 = 2.5.$

FINAL ANSWER: $20)/\sqrt{64} = (20)/8 = (10)/4 = 5$**/2=2.5.**

14. If Joe found that the sample mean for the first FOUR fish of his SIXTYFOUR fish sample is 90 grams, and if the sample standard deviation for the first FOUR fish in the sample is 19 grams, then what should Joe think is the TRUE NUMBER OF DEGREES OF FREEDOM that governs the statistic used to compute the margin of error for his 95 percent confidence interval for the true mean weight of all the 1000 fish in the pond based only on the FIRST FOUR FISH, if Joe assumes that the distribution of fish weight in the pond is normal?

ANSWER: With X **denoting the weight of a fish, with** U **denoting the sample mean of size** n, **and with** S **denoting the sample standard deviation, all as unknowns, then the statistic Joe is using for standardizing his estimation is** T_{n-1} **given by**

$$T_{n-1} = \frac{U - \mu_X}{S/\sqrt{n}},$$

which has the student-t **distribution for** $n-1$ **degrees of freedom. Here** $n = 4$, **so the number of degrees of freedom is 3.**

FINAL ANSWER: 3.

15. If Joe found that the sample mean for his SIXTYFOUR fish is 90 grams, and if the sample standard deviation for his sample is 20 grams, then what should Joe think is the STANDARD ERROR he must use to compute the margin of error for his 95 percent confidence interval for the true mean weight of all the 1000 fish in the pond, if Joe assumes that the distribution of fish weight in the pond is normal AND that the population standard deviation is 16 grams because Sam told him so?

ANSWER: Let X denote fish weight. Then $U = \bar{X}_n$ is the sample mean unknown for a sample of size n. With independent random sampling (IRS), we know that

$$\sigma_{\bar{X}_n} = \frac{\sigma_X}{\sqrt{n}}.$$

Notice that Joe now knows the population standard deviation is $\sigma_X = 16$, so that over rules the sample standard deviation $s = 20$, which means the sample standard deviation is ignored now and the standard error is

$$\sigma_U = \frac{\sigma_X}{\sqrt{n}} = (16)/\sqrt{64} = (16)/8 = 2.$$

FINAL ANSWER: $(16)/\sqrt{64}$ **=(16)/8=2.**

16. If 16 of the SIXTYFOUR fish Joe caught are redfish, what should Joe guess is the true proportion of redfish in Sam's pond, ignoring any margin of error?

ANSWER: For Joe, the only thing he knows is that 16 out of 64 fish are redfish, so he should guess that the true proportion of redfish in Sam's pond is (16)/(64)=2/8=1/4 ignoring any margin of error.

FINAL ANSWER: (16)/(64)=2/8=1/4.

17. If 16 of the SIXTYFOUR fish Joe caught are redfish, what is the STANDARD ERROR he should use to compute the 95 percent confidence interval for the true proportion of redfish in Sam's pond based on Joe's sample?

ANSWER: In dealing with the redfish count, Joe is simply performing an independent random sample of the indicator $X = I_R$ of whether or not a fish caught is a redfish. If p denotes the true proportion of redfish in the pond, then $\mu_X = E(X) = E(I_R) = P(R) = p$ is the success rate for catching redfish. Also, we know that for any indicator, the variance is simply the success rate multiplied by the failure rate, so if we set $q = 1-p$, then $\sigma_X = \sqrt{pq}$. Unfortunately, Joe does not know what p is, so he does not know σ_X. In this situation, for very large sample, it has become acceptable practice to use the sample proportion observed, denoted $\hat{p}$ in place of the unknown true proportion. Obviously, this can be attacked on logical grounds of leading to somewhat circular reasoning, so the conservative approach which is often used is to make use of the fact that we can prove that ALWAYS $\sqrt{p(1-p)} = \sqrt{pq} \leq 1/2$. Since the standard error of the mean $\bar{X} = \hat{p}$ is $\sigma_X/\sqrt{n}$, it follows that here would would have a standard error of $\sqrt{(.25)(.75)}/\sqrt{64} = (\sqrt{(3)}/4)/8 = \sqrt{3}/(32)$ with the very large sample approach or simply $(1/2)/\sqrt{64} = 1/(16) = 2/(32)$ using the more conservative approach.

FINAL ANSWER: $\sqrt{(.25)(.75)}/\sqrt{64} = (\sqrt{(3)}/4)/8 = \sqrt{3}/(32)$ **or alternately, simply** $(1/2)/\sqrt{64}$**=1/ (16)=2/(32).**

18. If Joe found that the sample mean for his SIXTYFOUR fish is 90 grams, and if the sample standard deviation for his sample is 20 grams, then what should Joe think is the value of the TEST STATISTIC he will compute in seeking to PROVE the true mean weight of all the 1000 fish in the pond, is over 80 grams, if Joe assumes that the distribution of fish weight in the pond is normal?

ANSWER: With X denoting the weight of a fish, with U denoting the sample mean of size n, and with S denoting the sample standard deviation, all as unknowns, then the statistic Joe is using for standardizing his estimation is T_{n-1} given by

$$T_{n-1} = \frac{U - \mu_X}{S/\sqrt{n}},$$

which has the student-t distribution for $n-1$ degrees of freedom. So here, the degrees of freedom is $n - 1 = 64 - 1 = 63$. Under the null hypothesis for proving the true mean $\mu_X > 80$, Joe would use $\mu_0 = 80$ in place of the unknown the unknown true mean μ_X which he does not know. Thus Joe uses the null hypothesis as an assumption. But then the value of the test statistic is

$$T_{n-1} = \frac{U - \mu_0}{S/\sqrt{n}} = \frac{90 - 80}{(20)/\sqrt{64}} = (10)/([20]/8) = 5/4 = 8/2 = 4 = t_{data},$$

As an aside, notice that here the significance of Joe's data as proof that $\mu_X > 80$ is

$$\alpha_{data} = P(T_{n-1} \geq t_{data}) = P(T_{63} \geq 4).$$

Moreover, here , it is acceptable to use the approximation $T_{63} = Z$, the standard normal, as the degrees of freedom is so high, and in these terms, we see the data is very significant at any reasonable level of significance.

FINAL ANSWER: $\frac{90-80}{(20)/\sqrt{64}}$ =(10)/([20]/8)=8/2=4.

19. If Joe found that the sample mean for his SIXTYFOUR fish is 90 grams, and if the sample standard deviation for his sample is 20 grams, then what should Joe think is the value of the TEST STATISTIC he will compute in seeking to PROVE the true mean weight of all the 1000 fish in the pond, is over 80 grams, if Joe assumes that the distribution of fish weight in the pond is normal AND that the population standard deviation is 16 grams because Sam told him so?

ANSWER: Here Joe knows the population standard deviation, so the test statistic as an unknown before looking at the data is the standard normal Z, so

$$Z = \frac{U - \mu_X}{\sigma_X/\sqrt{n}},$$

and again, in trying to prove $\mu_X > 80$, Joe will replace the μ_X which he does not know with value 80 which the null hypothesis gives, so the value of the test statistic with Joe's data is

$$z_{data} = \frac{(90 - 80)}{(16)/\sqrt{64}} = (10)/([16]/8) = (10)/2 = 5.$$

FINAL ANSWER: $\frac{(90-80)}{(16)/\sqrt{64}} = (10)/([16]/8) = (10)/2 = 5.$

20. Suppose that ANOTHER SAMPLE of 40 redfish in the pond had a mean weight of 100 grams with sample standard deviation 20 and ANOTHER SAMPLE of 50 bluefish had a sample mean weight of 70 grams with a standard deviation of 25, then what would be the STANDARD ERROR used to calculate the margin of error in a 95 percent confidence interval for the amount that true mean redfish weight exceeds the true mean bluefish weight for the fish in Sam's pond?

ANSWER: Let X be the weight of a redfish and Y be the weight of a bluefish. Then

$$U = \bar{X}_m - \bar{Y}_n$$

is an unbiased estimator of $E(U) = \mu_X - \mu_Y$, where m is the sample size of an independent random sample (IRS) of redfish and n is the size of an IRS of bluefish. Then, U being the difference of the two independent unknowns $\bar{X}_m$ and $\bar{Y}_n$, it follows that the VARIANCE of U is the sum of the variances

$$Var(U) = Var(\bar{X}_m) + Var(\bar{Y}_n) = \frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}.$$

As these sample sizes are $m = 40$ and $n = 50$, and are therefore considered large, it acceptable to use the sample standard deviations in place of the population standard deviations in the formula, so this gives

$$Var(U) = \frac{(20)^2}{40} + \frac{(25)^2}{50} = \frac{(20)(20)}{2(20)} + \frac{(25)(25)}{2(25)} = (20)/2 + (25)/2 = 10 + 12.5 = 22.5.$$

Since the standard error is the square root of the variance of U, it follows that the standard error is simply $\sqrt{22.5}$.

FINAL ANSWER: $\sqrt{22.5}$.