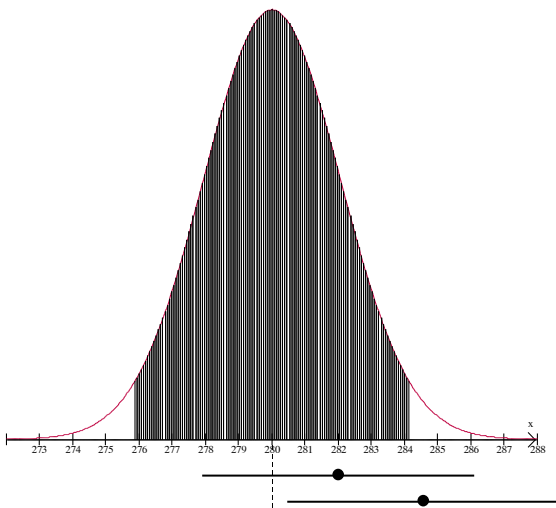
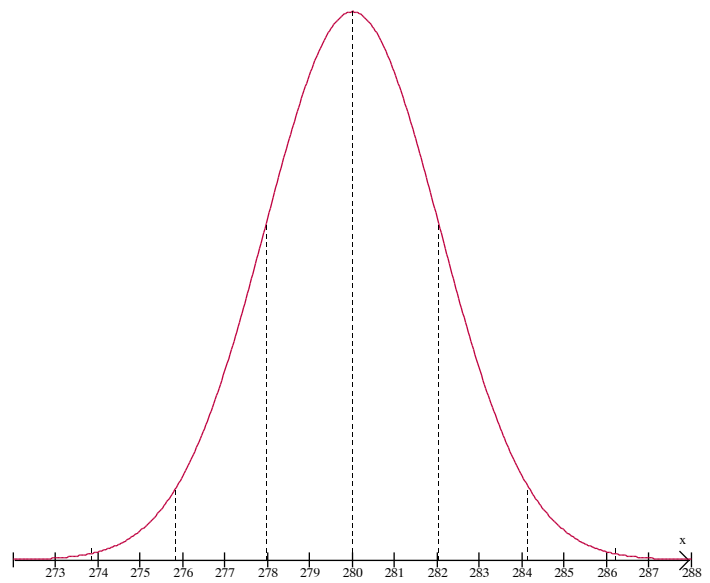


10.1

- (a) The 95% confidence interval is $0.47 \pm 0.03 = (0.44, 0.50)$.
- (b) We know that every time we go out and randomly sample from the population, the sample proportion of women who feel they don't get enough time will vary from sample to sample just due to chance variation. We can't make the claim that the actual proportion is 47%, but we can say it is probably something close to 47%. Our confidence interval helps us get an idea of "how close" it might be.
- (c) 95% confidence means that if we were to sample many, many times, and construct many, many confidence intervals, about 95% of the confidence intervals would contain the actual population proportion.

10.2

- (a) Since 840 is a "large" sample, we know that the sampling distribution of \bar{x} is approximately normal, with mean, $\mu_{\bar{x}} = \mu_x = 280$, and standard deviation $\sigma_{\bar{x}} = \frac{60}{\sqrt{840}}$. (Note that we can use this standard deviation because we were told the population is large.)
- (c) The missing number is $2\sigma_{\bar{x}} = 4.1404$ points



- (e) In 95% of all samples the true mean will be covered by the confidence interval we construct.

10.3

The student is NOT right. The truth is that about 95% of all young men have scores within 2 standard deviations of the true mean, whatever that mean is. We should say that "it is likely that the true mean is score of young men is somewhere between 267.8 and 276.2."

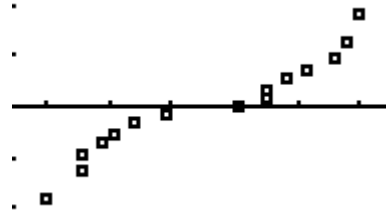
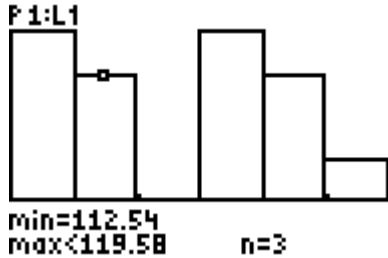
10.8

(a) We will construct a 90% confidence interval for the mean, μ , yield for this variety of corn.

We assume that the sample of plots is a random sample.

It is reasonable to assume that $n < 0.1$ (All plots of this variety of corn).

The histogram and normal probability plot show that we have no reason to doubt that the yield is normally distributed.



So we proceed

The 90% confidence interval is (119.55, 128.05) bushels per acre. We are 90% confident that the mean yield for this variety of corn is in this interval.

(b) The 95% confidence interval is (118.74, 128.86) bushels per acre.

(c) The 99% confidence interval is (117.15, 130.45) bushels per acre.

(f) As the confidence level increases, so does the margin of error.

10.12

(a) We will construct a 98% confidence interval for μ , the mean of the repeated measurements of the 10g weight. We are given that the distribution of all measurements is normally distributed, we consider the repeated weighings an SRS of all weighings, and it reasonable to assume that there will be at least 50 measurements over the lifetime of the scale. The TI-84 98% confidence interval is (10.002, 10.003). We are 98% confident that the true mean of repeated measurements is in this interval.

(b) We require that the margin of error,

$$2.3263 \frac{0.0002}{\sqrt{n}} \leq .0001$$

$$2.3263 \frac{0.0002}{.0001} \leq \sqrt{n} \quad \text{So } n = 22, \text{ the required sample size.}$$

$$n \geq 21.647$$

10.20

(a) The population of interest is all hotel managers of hotels sized 200 to 500 rooms. One shortcoming is that the sample only included hotel managers from Detroit and Chicago, Midwestern metropolitan areas. The sample is not representative of the population. Another major problem is the non-response bias caused by the 76% non-response rate.

(b) We will be measuring the sample mean score of all of the responses. The sample means do not take on integer values. The sample means can take on many, many values between 1 and 7, so \bar{X} is very nearly continuously distributed. In addition, because we have a pretty large sample size, the sampling distribution of \bar{X} is expected to be normal.

(c) We will construct a 95% confidence interval for μ , the mean satisfaction score for hotel managers and their software.

We don't have an SRS, so we cited our concern above. Since $n = 135$ is large, and since it is reasonable to assume that $135 < 0.1(\text{Population of hotel managers})$, we will proceed with the cautions mentioned above.

The 95% confidence interval is $\bar{x} \pm z * \frac{\sigma}{\sqrt{n}} = 5.396 \pm 1.9599 \frac{1.75}{\sqrt{135}} = (5.100, 5.691)$. We are 95% confident that the mean satisfaction score with the software is in this interval.

(d) The 99% confidence interval for the mean satisfaction score with the training is

$\bar{x} \pm z * \frac{\sigma}{\sqrt{n}} = 4.398 \pm 2.576 \frac{1.75}{\sqrt{135}} = (4.010, 4.786)$ We are 99% confident that the mean satisfaction score with the training is in this interval.

10.22

(a) The stemplot is shown here. The normal probability plot is shown as well.

Stem	Leaf
1	4 2 1
1	8
2	2 3 2 3
2	9 7 8 6
3	4 0 3
3	5 5
4	0

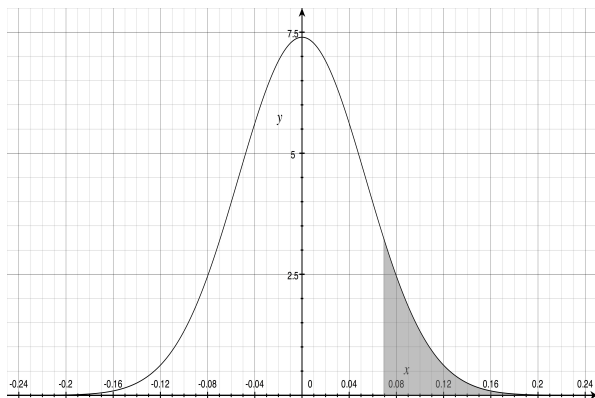


Since the stemplot does not show skewness, or obvious outliers, and the normal probability plot is fairly linear, we have no reason to doubt that the measurements come from a normal distribution.

(b) The 90% confidence interval is $(22.565, 28.7680) \mu/\text{hr}$. We are 90% confident that the true mean closing rate for the wounds is in this interval.

(c) The friend's interval is wider. In order to be more confident of "capturing" the true mean in our interval, we would need a wider interval.

10.36



(a) the curve is sketched below.

(b) $p\text{-value} = 0.1004$

(c) The result is NOT significant at the 5% level since the p-value is greater than 5%. That means we do not have significant evidence that the mean increase is greater than 0%.

10.40

(a) The test statistic is $\frac{0.4365 - 0.5}{\frac{0.2887}{10}} = -2.1995$

(b) The critical values corresponding to $\alpha = 0.05$ is $z^* = \pm \text{invnorm}(0.025) = \pm 1.9599$. Since the test statistic is less than the critical value, then the result is significant at the 5% level.

(c) The critical values corresponding to $\alpha = 0.01$ is $z^* = \pm \text{invnorm}(0.005) = \pm 2.5758$. Since the test statistic is not less than the critical value, then the result is not significant at the 1% level.

(d) The test statistic, z lies between -2.054 and -2.326 in Table C. The corresponding p-values (for the two sided test) that the p-value of our result lies between are 0.02 and 0.04.

10.42

(a) Referring to Table C in the back, the critical value for the one-sided test at the 0.05 level is 1.645. Clearly the test statistic is greater, so the result is significant at the 5% level.

(b) Referring again to Table C for the critical value at the $\alpha = 0.01$ level, we have $z^* = 2.326$. Thus the result is also significant at the 1% level.

(c) The p-value of our result is between 0.005 and 0.01 (from Table C).

10.51

Saying that a result is significant at the 1% level means that the probability of the result occurring purely due to chance is no more than 1%. So if a result is significant at the 1% level, then we know for sure that the probability that the result occurs purely by chance will necessarily be lower than 5%. Therefore it is significant at the 5% level already.

If a result is significant at the 5% level it may or may not be significant at the 1% level.

10.52

Note that we never calculate the probability of the null hypothesis being true.. It is either true or it is not true. There is not some “proportion of times that it is true”. So saying that we are calculating a probability that the null is true is sort of silly.

10.64

(a) The probability that any of the specific tests shows a difference significant at the 5% level is of course, 5%.

(Remember that if you interpreted this question as asking, “What is the probability that exactly one of the 77 tests shows significance?”, you must do a binomial probability calculation.)

(b) Let X represent the number of tests in the list of 77 that show a difference significant at the 5% level. Then X is binomially distributed with $n = 77$, and $p = 0.05$.

So the likelihood of obtaining 2 or more “successes” (tests showing significant results) of the 77 tests is $P(X \geq 2) = 1 - \text{binomcdf}(77, 0.05, 1) = 0.9027$. Since this is a very likely occurrence, we should not be surprised that it actually happened in this situation.

10.66

(a) The two hypotheses can be written

H_0 : The patient is healthy.

H_A : The patient is sick.

What is really true. (Only God knows)

		H_0 is True	H_A is True
<u>What the machine actually decides</u>	Reject H_0	Type I Error	
	Do Not Reject H_0		Type II Error

Type I error means that the machine will refer the patient when there is nothing wrong. This is called a “false positive”. Type II error means that the machine will not refer the patient when something is wrong. This is called a “false negative.”

(b) I think it is better to reduce the probability of Type II error. It is better to have a doctor find nothing wrong, as in the case of a machine false positive, than to let an illness go undetected, as in the case of a false negative.

10.68

(a) Since the rejection criteria is reject if $\bar{x} > 0$, then assuming $\mu = 0$, the probability of rejection is 50%.

(b) $P(\text{Type II error}) = P(\text{Not rejecting when } \mu = 0.3) = P(x \leq 0 | \mu = 0.3) = \text{normcdf}(-1E99, 0, 0.3, 1/3) = 0.184$

(c) $P(x \leq 0 | \mu = 1) = \text{normcdf}(-1E99, 0, 1, 1/3) = 0.00135$

10.72

(a) Power = $1 - P(\text{Type II Error}) = 1 - P(\text{Do not Reject } H_0 | \mu = 299) = 1 - P(\bar{x} > 297.985) = 0.2037$

(b) Power = $1 - P(\text{Type II Error}) = 1 - P(\text{Do not Reject } H_0 | \mu = 295) = 1 - P(\bar{x} > 297.985) = 0.9926$

(c) Since the alternative $\mu = 290$ is even farther from the null $\mu_0 = 300$ than $\mu = 295$, we expect the power against 290 to be even higher than the power against 295.

10.84

I think this student is essentially correct. When we conduct a test of significance, we are assuming that our null hypothesis is true in our probability calculations. We calculate a p-value, which is the probability that our sample result will occur due to chance variation under the assumption that the null is true. If this p-value is low, then we have a result that is unlikely due to chance variation alone. There must be some other explanation for our sample result.

10.86

(a) We will test whether μ , the mean SSHA score for students over 30 is higher than the mean score for U.S. College students ($=115$). That is, we test the hypotheses:

$$H_0: \mu = 115$$

$$H_A: \mu > 115$$

The TI-83 z-test: $z = 3.0511$, and the p -value = 0.0013. Since the test result will only occur about 0.13% of the time due to chance variation, assuming the mean is really 115, we have evidence to support the claim that the mean score for adults 30 years old and older is higher than 115.

(b) The conditions required by this test are that the sample is an SRS and that the distribution of the SSHA scores is approximately normal. It is more important that the sample be an SRS. If the population distribution is not approximately normal, the Central Limit Theorem says that with a sample size of 20 the sampling distribution will be close to normal. But if we do not have an SRS, the laws of probability on which we base the calculation of our p -value do not apply.

10.88

(a) This study is an observational study (not an experiment) because there are no treatments applied to the subjects. That is they did not require some to use the cell phone and some not too. They, the researchers, only observed existing cell phone use in the subjects.

(b) No statistically significant association means that the probability of an association occurring in this test was not so low that it seemed unlikely simply do to chance.

(c) Note that 5% of 20 is 1. So if we run 20 separate tests of significance, it is not surprising that 5% of them would turn out to show significance. This is what we would expect due to chance.

It is interesting to note that if you increase the sample size (number of observations) of that one particular type of cancer, the sampling distribution becomes less variable, and we would find the association between the occurrence of the cancer and the cell phone use less and less likely. So it is also not surprising that the study makes the (incorrect) statement that risk appears to decrease with greater cell phone use.