

11.12 (a)  $\mu$  is the mean difference ( $\mu_A - \mu_B$ ) for the two varieties' yields in pounds per plant.

(b) We will test if variety A has a higher mean yield; that

$$H_0: \mu = 0$$

$$H_A: \mu > 0.$$

• We assume the 10 plots are a random sample of all available plots.

• We do not know if the yields are normally distributed, but 10 is a good sample size if the distribution is not too far from normal.

✓ The T1-B3 T-Test gives:  $t = 1.2953$ ,  $p\text{-value} = .1137$ .

Since the likelihood of our sample result ( $\bar{x} = 0.34$ ) is about 11.37% simply due to chance (assuming there is no difference) we do not have very significant evidence that there is a difference.

11.16 The hypotheses are  $H_0: \mu = 0$   
 $H_A: \mu > 0$ ,  $\mu$  is the mean change in score (after - before).

T1-B3 T-Test:  $t = 6.9$ ,  $p\text{-value} = 0.0000003459$ .

✓ This very low p-value indicates there is very significant evidence to support the neurobiological arguments.

### 11.22 (AN OFFICIAL WRITE-UP)

(a) From TABLE C, with 49 df  $\approx$  50,  $t^* = 2.403$

(b) We will REJECT  $H_0$  IF  $t$ , our TEST STATISTIC, IS GREATER THAN 2.403.

That is IF  $t = \frac{\bar{x} - 0}{s/\sqrt{n}} > 2.403$  WE REJECT, OR

$$\bar{x} > 2.403 \left( \frac{108}{\sqrt{50}} \right) = 36.702$$

(c) Power =  $P(\bar{x} > 36.702 \mid \mu = 100, \sigma = 108)$   
 $= \text{NORMALCDF}(36.702, 1e99, 100, 108/\sqrt{50}) = 0.99998$

This PROBABILITY (POWER) IS QUITE HIGH, SO A SAMPLE OF 50 (OR MAYBE EVEN FEWER) WOULD BE JUST FINE.

(d) A TYPE I ERROR MEANS WE WOULD BELIEVE THERE IS AN INCREASE IN SPENDING WHEN IN FACT THERE IS NOT; A TYPE II ERROR MEANS WE WOULD THINK THERE IS NO INCREASE IN SPENDING WHEN IN FACT THERE IS. TYPE I ERROR MAY LEAD TO A LOSS OF MONEY WHERE TYPE II ERROR MAY MEAN WE MISS AN OPPORTUNITY TO MAKE MORE MONEY... IN THIS CASE I THINK TYPE I ERROR IS MORE SERIOUS.

11.32 The 90% C.I. FOR THE MEAN DIFFERENCE (RIGHT - LEFT) IS:

$$(-21.17, -5.472) \text{ SECONDS.}$$

$$\bar{x}_{\text{RIGHT}} = 104.12 \text{ SECONDS, } \bar{x}_{\text{LEFT}} = 117.44 \text{ SECONDS}$$

$$\frac{\bar{x}_{\text{RIGHT}}}{\bar{x}_{\text{LEFT}}} = 0.8865 : \text{ WE MIGHT SAY THAT OVER TIME THE RIGHT-HAND THREAD IS ABOUT 12\% FASTER.}$$

### 11.36

(a) THE SCORES IN THIS SAMPLE APPEAR QUITE SKEWED TO THE RIGHT, AND YET THERE ARE NO APPARENT OUTLIERS. THE NORMAL PROBABILITY PLOT SHOWS NO MAJOR DEVIATIONS FROM NORMALITY, AND THERE IS GOOD REASON TO THINK THESE SCORES COME FROM A NORMAL DISTRIBUTION ANYWAY.

THE STANDARD DEVIATION OF THE SAMPLING DIST<sup>n</sup> OF  $\bar{x}$ , THE SAMPLE MEAN FROM  $n = 25$  IS APPROXIMATELY 0.2 POINTS, SO 25 IS A GOOD SAMPLE SIZE FOR ESTIMATING.

(b) SEE ABOVE

(c) WE CONSTRUCT A 95% C.I. FOR  $\mu$ , THE MEAN AP SCORE IN 2001.

(0.571, 1.508). WE ARE 95% CONFIDENT THE MEAN SCORE FOR 2001 IS IN THIS INTERVAL

### 11.40

(a) If the loggers were aware that the effects of logging were being assessed, they may have changed their behavior in ways that would change how logging affects the rainforests. By “blinding” the loggers we are controlling the effect on damage caused by the variation in whether they know or don’t know they’re being assessed.

(b) We will test whether  $\mu_1$ , the mean number of tree species in areas logged 8 years ago is less than  $\mu_2$ , the mean number of tree species in unlogged areas:  $H_0: \mu_1 = \mu_2$  vs.  $H_A: \mu_1 < \mu_2$ .

The study explained that the plots can be considered a random sample of plots. The number of species on logged plots is notably skewed to the left, but has no outliers. The number of species on unlogged plots is more symmetric with no outliers. Since  $9 + 12 = 21$  is sort of small, considering the shape of the logged plots distribution, we proceed with caution. We will also note that the groups of plots are independent, and that 21 is surely less than all possible plots of that size in the forest.

The TI-84 2-SampTTest:  $t = -2.114$ ,  $df = 14.79$ ,  $p\text{-value} = 0.0259$ . A p-value this low gives reasonable evidence against the null hypothesis. That is, these results give significant evidence that the mean number of tree species in logged areas is less than that in unlogged areas.

(c) The 90% confidence interval for the difference in mean number of species ( $\mu_1 - \mu_2$ ) is  $(-7.015, -0.6517)$ . We are 90% confident that the true mean difference is in this interval.

### 11.42

(a) Since we have such a large combined sample size, we can safely use  $t$ -procedures.

(b) We will construct a 90% C.I. for  $\mu_1 - \mu_2$ , the difference between the mean summer earnings of male and female students, where  $\mu_1$  is the mean earnings for males, and  $\mu_2$  is the mean earnings for females. We assume the sample is random, and that they are independent. Also it is reasonable that  $1296 < 0.1$  (all possible undergraduates).

The 90% C.I. from the TI-84 2-SampTInt is  $(413.62, 634.64)$ . We are 90% confident that the true difference in mean earnings (males – females) is in this interval.

(c) This systematic sample is not officially an SRS, but it is random.

**(d) We would need to know if the list is a list of ALL undergraduates, or just from a particular college or colleges.**

**11.50**

- (a) There are two samples of size 50 each. This is a two-sample test.
- (b) The tires are paired together on the same car. We will measure the difference in the wear on the tires from each car. This is a matched pairs, single sample test.
- (c) Since each person is subjected to both treatments, each person acts as his or her own control in this matched pairs design. This is a matched pairs t test.
- (d) The two different samples mean that their responses are independent. This is a two-sample test.
- (e) Again, each women's weight difference is recorded. While this is not a good design because of a lack of a control group, this still is a matched pairs design.

**11.51**

Group	Treat.	$n$	$\bar{x}$	$s$
1	IDX	10	116	17.7087
2	Untreated	10	88.5	6.0083

We would use 9 df in our two-sample t-test.

**11.52**

We will test whether  $\mu_1$ , the mean life-span after infection with Scrapie in IDX treated hamsters, is greater than  $\mu_2$ , the mean life-span in untreated hamsters. That is, we test the hypotheses

$$H_0: \mu_1 = \mu_2$$

$$H_A: \mu_1 > \mu_2$$

Since we are not sure if our sample of 20 hamsters is an SRS of the population, we may not be able to generalize our results to all hamsters. The randomization of our experiment, however tries to control for this.

Assume that the life-spans for treated and untreated hamsters is approximately normal, and since  $n_1$  and  $n_2$  are both greater than 5, we can proceed with our two-sample test.

TI-83 2-Sample T-Test:  $t = 4.6503$  with 11.04 df,  $p$ -value = 0.000348

A  $p$ -value this low provides strong evidence against our null hypothesis. We therefore have strong evidence that the IDX treated hamsters live longer.

The 95% confidence interval for the mean increase in life-span provided by IDX treatment is (14.491, 40.509) days. We are 95% confident that the true mean increase in life-span is in this interval.

**11.55**

(a) We examine the differenced data,  $X = (\text{Reference Drug Absorption} - \text{Generic Drug Absorption})$ .

The histogram looks fairly symmetric, but we note a large difference in the mean ( $\bar{x} = -37$ ) and the median,  $-215.5$ , suggesting some problems with the data. The boxplot shows that subject 15 is an outlier with a difference of 2353 absorption score. While the normal probability plot is roughly linear, the presence of the outlier and the small sample size make any conclusions from the  $t$  procedures unreliable.

(b) Nonetheless, we construct a 90% confidence interval for  $\mu = \mu_{\text{reference}} - \mu_{\text{generic}}$ , the mean

difference in absorption score for the reference drug and the generic drug. The 90% C.I. is (−451, 376.95). We are 90% confident that the true mean difference is in this interval.

### 11.56

(a) It would be appropriate in examining the coached students only to do a matched pairs test using the differences of the individual scores, or the “Gain”. This is because the scores for each student on their second try in the coached group are dependent on their scores from their first try. In a two-sample test, we require independence of the variables.

(b) We will test whether  $\mu$ , the mean gain in SAT score for coached students is greater than 0. That is we test  $H_0: \mu = 0$  vs.  $H_A: \mu > 0$ .

We are told the sample of coached students is random, and the sample size is 427, which is quite large enough to ensure the sampling distribution is approximately normal. Also it is reasonable that  $427 < 0.1(\text{Population of coached students.})$

The TI83 t-test gives:  $t = 10.157$  with  $p\text{-value} = 3.77 \times 10^{-22}$ . This  $p\text{-value}$  indicates we have very significant evidence that there is a gain in the mean score of coached students on their second try with the SAT.

(c) The 99% confidence interval for the mean gain is (21.612, 36.388). We are 99% confident that the actual mean gain that coach students attain is in this interval.

### 11.57

(a) We now test if the mean gain for coached students,  $\mu_{\text{coached}}$ , is higher than the mean gain for uncoached, students,  $\mu_{\text{uncoached}}$ . That is we test  $H_0: \mu_{\text{coached}} - \mu_{\text{uncoached}} = 0$  vs.  $H_A: \mu_{\text{coached}} - \mu_{\text{uncoached}} > 0$ .

Note that we have a large sample of uncoached students, so the sampling distribution should be close to normal.

The TI83 2-sample T-test gives:  $t = 2.646$ ,  $df = 534.45$ ,  $p\text{-value} = 0.00419$

Since we would expect a difference in sample gains this large or larger only 0.42% of the time assuming there is no difference in gains, we have significant evidence to support the claim that the mean gain for coached students is higher than that of uncoached students.

(b) The 99% confidence interval for the difference in mean gain (coached – uncoached) is (0.184, 15.816). We are 99% confident that the true difference in mean gain is in this interval.

(c) I do NOT think, from this study, that the coaching courses are worth paying for. While there is evidence of a statistically significant difference in mean gain for the coached students, it is not practically significant, because on average it is only a difference of at most 16 points.

### 11.58

**Simply put we cannot say that coaching caused the coached students' higher gain because this was an observational study, not a designed experiment. There were no steps taken to control for lurking or confounding variables that could have affected the response.**

## 11.62

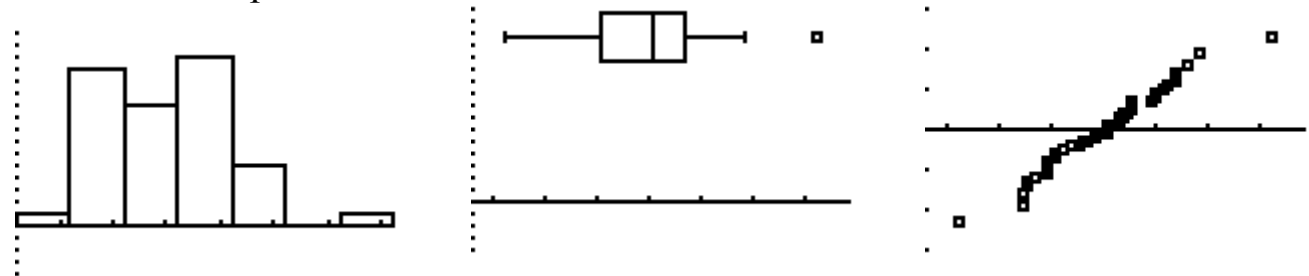
(a) This is clearly a two-sample  $t$  test setting. We are working with two different populations, those who experience the “undermining” information and those who don't.

(b) I would use 44 df in a conservative application of the test.

(c) We are considering the sampling distribution of sample means (with  $n_1+n_2=90$ ) from these populations. The sampling distribution with such a large sample size will be, by the CLT, approximately normal.

## 11.63

Here are some plots...



(a) We notice that the plots indicate moderate symmetry despite an outlier. The npp isn't too far from being linear. According to the instructions of the question, we proceed.

(b) We construct a 95% confidence interval for  $\mu$ , the mean length of great white sharks. We consider the sample an SRS from the all great whites, and that  $44 < 1/10$  (All Great White Sharks).

The 95% C.I. is (14.811, 16.362) feet. We are 95% confident that the mean length of great white sharks is in this interval.

We will test the hypotheses:  $H_0: \mu=20$ , vs  $H_A: \mu \neq 20$ . Based on our interval above, noting that 20 feet is not in the interval, we have evidence at the 5% significance level to reject  $H_0$ . That is we have evidence to support the claim that the mean length of great white sharks is not 20 feet.

(c) We need to know whether the sample is a random sample of all great white sharks.

## 11.64

(a) Use a matched pairs test. Measure the difference in weight gain for the littermates.

(b) This is a two sample setting. Measure the two mean “effects” from each of the 100 plot groups.

(c) This is a two sample setting. Measure the mean salaries from the two groups (Male and Female).

(d) This is a matched pairs test. Measure the difference of “effect” (A – B) from each of the 100 plots.

(e) This is a matched pairs test setting. Measure the difference in calculation time (Calculator A – Calculator B) for each of the 100 participants.

### 11.68

(a) We will test whether  $\mu_1$ , the mean increase (Post - Pre) in test score for those receiving the positive subliminal message is greater than  $\mu_2$ , the mean increase for those receiving the neutral message. That is, we test the hypotheses,

$$H_0: \mu_1 = \mu_2, \text{ vs } H_A: \mu_1 > \mu_2$$

The plots of the differences for each of the groups look fairly symmetric with no major deviations from normality. We assume reasonably that  $18 < 0.1$  (All students who failed the Assessment test and are remediated). The participants were randomly assigned to treatment groups, but are not an SRS. Application to the whole population may not be warranted.

The TI-84 2-SampTTest gives:  $t = 1.9135$ , 13.9 df,  $p\text{-value} = 0.038$ . So we have significant evidence at both the 10% and the 5% level against the null. That is there is strong statistical evidence supporting that the mean increase in test score (Post – Pre) is higher for those receiving positive subliminal messages than for those who don't.

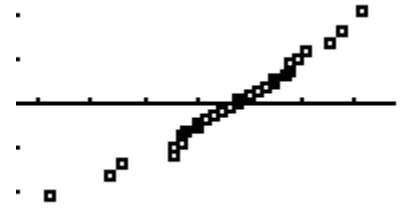
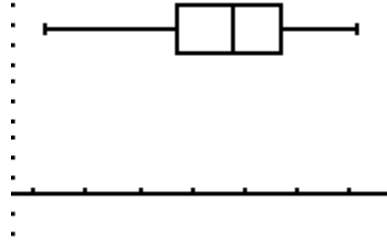
(b) The 90% C.I. for  $\mu_1 - \mu_2$  is (0.23943, 6.0506). We are 90% confident that the true difference in mean gain (tx – ctrl) is in this interval. So we expect on average for the mean increase in test score for those who receive a subliminal message to be between 0.2 and 6.0 points higher than those who don't.

### 11.70

It would certainly not be appropriate to apply our inference procedures to census data. If we had good census data, we would KNOW the mean population of an Indiana county, and there would be no need to estimate it with a confidence interval.

## 11.71

Here are plots of the measurements...



The histogram is very symmetric, there are no outliers, and the npp is quite linear. It seems clear that these measurements quite closely follow a normal distribution. Consider the measurements a random sample from all possible measurements, and that the population of all measurements is infinitely large.

The 95% confidence interval for  $\mu$ , the mean density of the earth is (5.3639, 5.532) times the density of water. We are 95% confident that the mean density of the earth is in this interval.