Review of Part VI – Learning About the World

1. Crawling.

a) H_0: The mean age at which babies begin to crawl is the same whether the babies were born in January or July. $\left(\mu_{Jan} = \mu_{July} \text{ or } \mu_{Jan} - \mu_{July} = 0\right)$

H_A: There is a difference in the mean age at which babies begin to crawl, depending on whether the babies were born in January or July. $\left(\mu_{Jan} \neq \mu_{July} \text{ or } \mu_{Jan} - \mu_{July} \neq 0\right)$

Independent groups assumption: The groups of January and July babies are independent.
Randomization condition: Although not specifically stated, we will assume that the babies are representative of all babies.
10% condition: 32 and 21 are less than 10% of all babies.
Nearly Normal condition: We don't have the actual data, so we can't check the distribution of the sample. However, the samples are fairly large. The Central Limit Theorem allows us to proceed.

Since the conditions are satisfied, it is appropriate to model the sampling distribution of the difference in means with a Student's *t*-model, with 43.68 degrees of freedom (from the approximation formula).

We will perform a two-sample *t*-test. The sampling distribution model has mean 0, with standard error: $SE(\bar{y}_{Jan} - \bar{y}_{July}) = \sqrt{\dfrac{7.08^2}{32} + \dfrac{6.91^2}{21}} \approx 1.9596.$

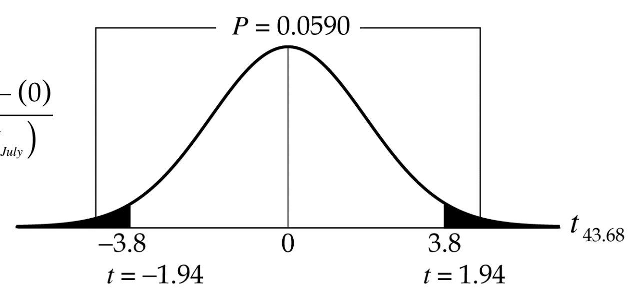
The observed difference between the mean ages is 29.84 – 33.64 = – 3.8 weeks.

Since the *P*-value = 0.0590 is fairly low, we reject the null hypothesis. There is some evidence that mean age at which babies crawl is different for January and June babies. June babies appear to crawl a bit earlier than July babies, on average. Since the evidence is not strong, we might want to do some more research into this claim.

$$t = \frac{\left(\bar{y}_{Jan} - \bar{y}_{July}\right) - (0)}{SE\left(\bar{y}_{Jan} - \bar{y}_{July}\right)}$$

$$t \approx \frac{-3.8 - 0}{1.9596}$$

$$t \approx -1.94$$



b) H_0: The mean age at which babies begin to crawl is the same whether the babies were born in April or October. $\left(\mu_{Apr} = \mu_{Oct} \text{ or } \mu_{Apr} - \mu_{Oct} = 0\right)$

H_A: There is a difference in the mean age at which babies begin to crawl, depending on whether the babies were born in April or October. $\left(\mu_{Apr} \neq \mu_{Oct} \text{ or } \mu_{Apr} - \mu_{Oct} \neq 0\right)$

The conditions (with minor variations) were checked in part a.

Since the conditions are satisfied, it is appropriate to model the sampling distribution of the difference in means with a Student's *t*-model, with 59.40 degrees of freedom (from the approximation formula).

We will perform a two-sample *t*-test. The sampling distribution model has mean 0, with standard error: $SE(\bar{y}_{Apr} - \bar{y}_{Oct}) = \sqrt{\frac{6.21^2}{26} + \frac{7.29^2}{44}} \approx 1.6404$.

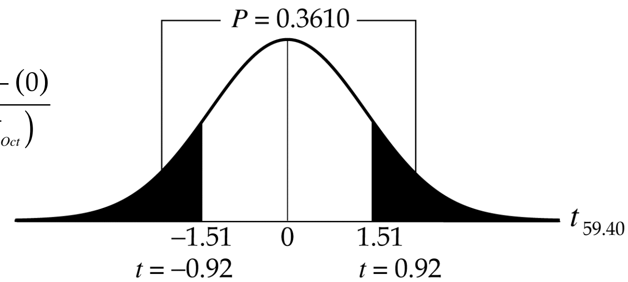
The observed difference between the mean ages is 31.84 – 33.35 = – 1.51 weeks.

Since the *P*-value = 0.3610 is high, we fail to reject the null hypothesis. There is no evidence that mean age at which babies crawl is different for April and October babies.

$$t = \frac{(\bar{y}_{Apr} - \bar{y}_{Oct}) - (0)}{SE(\bar{y}_{Apr} - \bar{y}_{Oct})}$$

$$t \approx \frac{-1.51 - 0}{1.6404}$$

$$t \approx -0.92$$



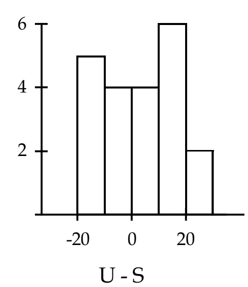
c) These results are not consistent with the researcher's claim. We have slight evidence in one test and no evidence in the other. The researcher will have to do better than this to convince us!

2. Mazes and smells.

H_0: The mean difference in maze times with and without the presence of a floral aroma is zero. $(\mu_d = 0)$

H_A: The mean difference in maze times with and without the presence of a floral aroma (unscented – scented) is greater than zero. $(\mu_d > 0)$

Paired data assumption: Each subject is paired with himself or herself.
Randomization condition: Subjects were randomized with respect to whether they did the scented trial first or second.
10% condition: We are testing the effects of the scent, not the subjects, so this condition doesn't need to be checked.
Nearly Normal condition: The histogram of differences between unscented and scented scores shows a distribution that could have come from a Normal population, and the sample size is fairly large.



Since the conditions are satisfied, the sampling distribution of the difference can be modeled with a Student's *t*-model with 21 – 1 = 20 degrees of freedom, $t_{20}\left(0, \frac{13.0087}{\sqrt{21}}\right)$.

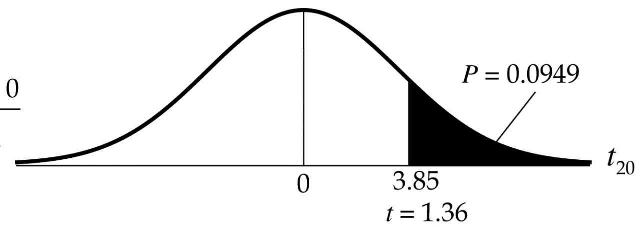
We will use a paired *t*-test, (unscented – scented) with $\bar{d} = 3.85238$ seconds.

Since the *P*-value = 0.0949 is fairly high, we fail to reject the null hypothesis. There is little evidence that the mean difference in time required to complete the maze is greater than zero. The floral scent didn't appear to cause lower times.

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}}$$

$$t = \frac{3.85238 - 0}{\frac{13.0087}{\sqrt{21}}}$$

$$t \approx 1.36$$



3. Women.

H_0 : The percentage of businesses in the area owned by women is 26%. ($p = 0.26$)
H_A : The percentage of businesses in the area owned by women is not 26%. ($p \neq 0.26$)

Random condition: This is a random sample of 410 businesses in the Denver area.
10% condition: The sample of 410 businesses is less than 10% of all businesses.
Success/Failure condition: $np = (410)(0.26) = 106.6$ and $nq = (410)(0.74) = 303.4$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.26$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.26)(0.74)}{410}} \approx 0.02166$.

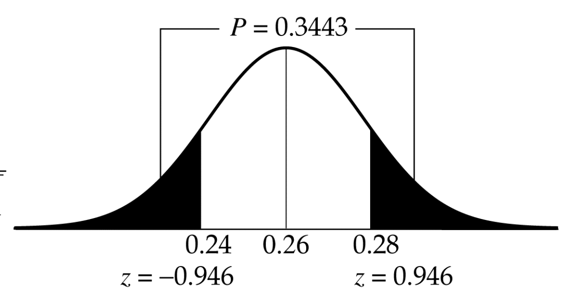
We can perform a one-proportion *z*-test. The observed proportion of businesses owned by women is $\hat{p} = \frac{115}{410} \approx 0.2805$.

Since the *P*-value = 0.3443 is high, we fail to reject the null hypothesis. There is no evidence that the proportion of businesses in the Denver area owned by women is any different than the national figure of 26%.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

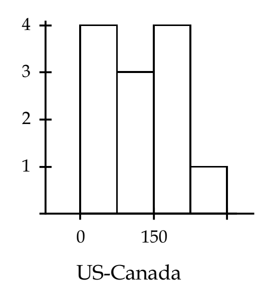
$$z = \frac{0.2805 - 0.26}{\sqrt{\frac{(0.26)(0.74)}{410}}}$$

$$z \approx 0.946$$



4. Drugs.

a) Paired data assumption: The data are paired by drug.
Randomization condition: These drugs DO NOT appear to be a random sample of all drugs. The names all begin with one of 5 letters!
10% condition: These drugs are less than 10% of all drugs.
Nearly Normal condition: The histogram of the differences in price is roughly unimodal and symmetric.

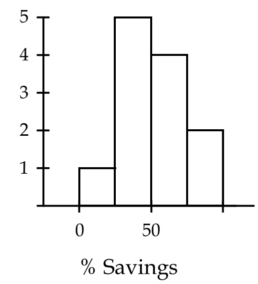


Since the conditions are satisfied, the sampling distribution of the difference can be modeled with a Student's *t*-model with 12 – 1 = 11 degrees of freedom. We will find a paired *t*-interval, with 95% confidence.

$$\bar{d} \pm t^*_{n-1}\left(\frac{s_d}{\sqrt{n}}\right) = 126 \pm t^*_{11}\left(\frac{76.2257}{\sqrt{12}}\right) \approx (77.57,\ 174.43)$$

We are 95% confident that the mean savings in drug cost when buying from the discount pharmacy is between about \$77.60 and \$174.40.

b) **Paired data assumption:** The data are paired by drug.
Randomization condition: These drugs DO NOT appear to be a random sample of all drugs. The names all begin with one of 5 letters!
10% condition: These drugs are less than 10% of all drugs.
Nearly Normal condition: The histogram of the percent savings is roughly unimodal and symmetric.



Since the conditions are satisfied, the sampling distribution of the difference can be modeled with a Student's *t*-model with 12 – 1 = 11 degrees of freedom. We will find a paired *t*-interval, with 95% confidence.

$$\bar{d} \pm t^*_{n-1}\left(\frac{s_d}{\sqrt{n}}\right) = 52.1667 \pm t^*_{11}\left(\frac{18.9681}{\sqrt{12}}\right) \approx (40.1\%,\ 64.2\%)$$

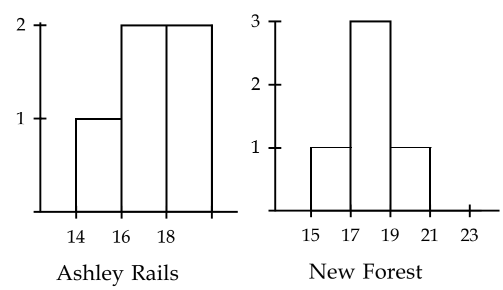
We are 95% confident the mean savings in drug cost when buying from the discount pharmacy is between 40.1% and 64.2%.

c) The analysis using the percents is more appropriate. It equalizes variability that is strictly due to the relative cost of the drugs. Some drugs are more expensive than others, and if just one of these drugs costs much less in Canada, it can pull the mean savings down.

d) These drugs are probably not a random sample of all drugs available. It is very unlikely that the 12 drugs chosen would have names beginning with one of a few letters. Also, the ad may have deliberately showcased those drugs that cost much less in Canada.

5. Pottery.

Independent groups assumption: The pottery samples are from two different sites.
Randomization condition: It is reasonable to think that the pottery samples are representative of all pottery at that site with respect to aluminum oxide content.
10% condition: The samples of 5 pieces are less than 10% of all pottery pieces.
Nearly Normal condition: The histograms of aluminum oxide content are roughly unimodal and symmetric.



Since the conditions are satisfied, it is appropriate to model the sampling distribution of the difference in means with a Student's *t*-model, with 7 degrees of freedom (from the approximation formula). We will construct a two-sample *t*-interval, with 95% confidence.

$$(\bar{y}_{AR} - \bar{y}_{NF}) \pm t^*_{df}\sqrt{\frac{s^2_{AR}}{n_{AR}} + \frac{s^2_{NF}}{n_{NF}}} = (17.32 - 18.18) \pm t^*_7\sqrt{\frac{1.65892^2}{5} + \frac{1.77539^2}{5}} \approx (-3.37,\ 1.65)$$

We are 95% confident that the difference in the mean percentage of aluminum oxide content of the pottery at the two sites is between –3.37% and 1.65%. Since 0 is in the interval, there is no evidence that the aluminum oxide content at the two sites is different. It would be reasonable for the archaeologists to think that the same ancient people inhabited the sites.

6. Streams.

Random condition: The researchers randomly selected 172 streams.
10% condition: 172 is less than 10% of all streams.
Success/Failure condition: $n\hat{p} = 69$ and $n\hat{q} = 103$ are both greater than 10, so the sample is large enough.

Since the conditions are met, we can use a one-proportion z-interval to estimate the percentage of Adirondack streams with a shale substrate.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} = \left(\frac{69}{172}\right) \pm 1.960 \sqrt{\frac{\left(\frac{69}{172}\right)\left(\frac{103}{172}\right)}{172}} = (32.8\%, 47.4\%)$$

We are 95% confident that between 32.8% and 47.4% of Adirondack streams have a shale substrate.

7. Gehrig.

H_0: The proportion of ALS patients who were athletes is the same as the proportion of patients with other disorders who were athletes. $\left(p_{ALS} = p_{Other} \text{ or } p_{ALS} - p_{Other} = 0\right)$

H_A: The proportion of ALS patients who were athletes is greater than the proportion of patients with other disorders who were athletes. $\left(p_{ALS} > p_{Other} \text{ or } p_{ALS} - p_{Other} > 0\right)$

Random condition: This is NOT a random sample. We must assume that these patients are representative of all patients with neurological disorders.
10% condition: 280 and 151 are both less than 10% of all patients with disorders.
Independent samples condition: The groups are independent.
Success/Failure condition: $n\hat{p}$(ALS) = (280)(0.38) = 106, $n\hat{q}$(ALS) = (280)(0.72) = 174, $n\hat{p}$(Other) = (151)(0.26) = 39, and $n\hat{q}$(Other) = (151)(0.74) = 112 are all greater than 10, so the samples are both large enough.

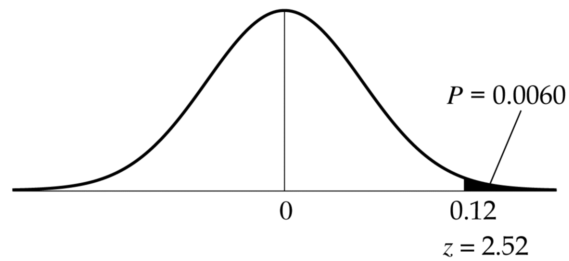
Since the conditions have been satisfied, we will model the sampling distribution of the difference in proportion with a Normal model with mean 0 and standard deviation estimated by

$$SE_{\text{pooled}}\left(\hat{p}_{ALS} - \hat{p}_{Other}\right) = \sqrt{\frac{\hat{p}_{\text{pooled}}\hat{q}_{\text{pooled}}}{n_{ALS}} + \frac{\hat{p}_{\text{pooled}}\hat{q}_{\text{pooled}}}{n_{Other}}} = \sqrt{\frac{(0.336)(0.664)}{280} + \frac{(0.336)(0.664)}{151}} \approx 0.0477.$$

The observed difference between the proportions is 0.38 – 0.26 = 0.12.

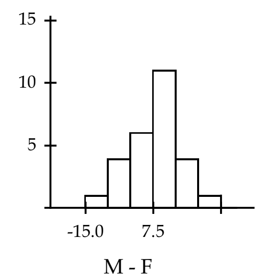
Since the P-value = 0.0060 is very low, we reject the null hypothesis. There is strong evidence that the proportion of ALS patients who are athletes is greater than the proportion of patients with other disorders who are athletes.

$$z = \frac{0.12 - 0}{0.0477}$$
$$z \approx 2.52$$



b) This was a retrospective observational study. In order to make the inference, we must assume that the patients studied are representative of all patients with neurological disorders.

8. Teen drinking.



Paired data assumption: The data are paired by country.
Randomization condition: We don't know the individuals in the different countries were chosen. We will assume that the individuals are representative of teens in each country.
10% condition: The teens sampled are less than 10% of all teens in the countries.
Normal population assumption: The histogram of differences is roughly unimodal and symmetric.

Since the conditions are satisfied, the sampling distribution of the difference can be modeled with a Student's t-model with 27 – 1 = 26 degrees of freedom. We will find a paired t-interval, with 95% confidence.

$$\bar{d} \pm t^*_{n-1}\left(\frac{s_d}{\sqrt{n}}\right) = 7.96296 \pm t^*_{26}\left(\frac{8.79459}{\sqrt{27}}\right) \approx (4.5\%, 11.4\%)$$

We are 95% confident that the mean percentage of 15-year-old boys in these countries who have been drunk at least twice is between 4.5% and 11.4% higher than the percentage of 15-year-old girls.

9. Babies.

H_0: The mean weight of newborns in the U.S. is 7.41 pounds, the same as the mean weight of Australian babies. $(\mu = 7.41)$

H_A: The mean weight of newborns in the U.S. is not the same as the mean weight of Australian babies. $(\mu \neq 7.41)$

Randomization condition: Assume that the babies at this Missouri hospital are representative of all U.S. newborns. (Given)
10% condition: 112 newborns are less than 10% of all newborns.
Nearly Normal condition: We don't have the actual data, so we cannot look at a graphical display, but since the sample is large, it is safe to proceed.

The babies in the sample had a mean weight of 7.68 pounds and a standard deviation in weight of 1.31 pounds. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean weight of U.S. newborns with a Student's *t* model, with $112 - 1 = 111$ degrees of freedom, $t_{111}\left(7.41, \frac{1.31}{\sqrt{112}}\right)$.

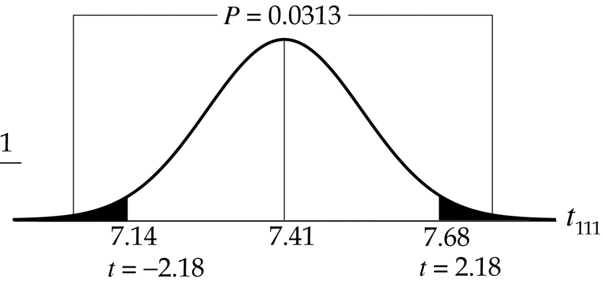
We will perform a one-sample *t*-test.

Since the *P*-value = 0.0313 is low, we reject the null hypothesis. If we believe that the babies at this Missouri hospital are representative of all U.S. babies, there is evidence to suggest that the mean weight of U.S. babies is different than the mean weight of Australian babies. U.S. babies appear to weigh more on average.

$$t = \frac{\bar{y} - \mu_0}{SE(\bar{y})}$$

$$t = \frac{7.68 - 7.41}{\frac{1.31}{\sqrt{112}}}$$

$$t \approx 2.18$$



10. Petitions.

a) $\frac{1772}{2000} = 0.886 = 88.6\%$ of the sample signatures were valid.

b) $\frac{250{,}000}{304{,}266} \approx 0.822 \approx 82.2\%$ of the petition signatures must be valid in order to have the initiative certified by the Elections Committee.

c) If the Elections Committee commits a Type I error, a petition would be certified when there are not enough valid signatures.

d) If the Elections Committee commits a Type II error, a valid petition is not certified.

e) H_0: The percentage of valid signatures is 82.2% ($p = 0.822$)
H_A: The percentage of valid signatures is greater than 82.2% ($p > 0.822$)

Random Condition: This is a simple random sample of 2000 signatures.
10% condition: The sample of 2000 businesses is less than 10% of all signatures.
Success/Failure condition: $np = (2000)(0.822) = 1644$ and $nq = (2000)(0.178) = 356$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.822$ and

$$\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.822)(0.178)}{2000}} \approx 0.00855.$$

We can perform a one-proportion *z*-test.

The observed proportion of valid signatures is $\hat{p} = \frac{1772}{2000} \approx 0.886$.

Since the *P*-value = 3.64×10^{-14} is low, we reject the null hypothesis. There is strong evidence that the percentage of valid signatures is greater than 82.2%. The petition should be certified.

$$z = \frac{\hat{p} - p_0}{\sigma(\hat{p})}$$

$$z \approx \frac{0.886 - 0.822}{0.008553}$$

$$z \approx 7.48$$

f) In order to increase the power of their test to detect valid petitions, the Elections Committee could sample more signatures.

11. Feeding fish.

a) If there is no difference in the average fish sizes, the chance of observing a difference this large, or larger, just by natural sampling variation is 0.1%.

b) There is evidence that largemouth bass that are fed a natural diet are larger. The researchers would advise people who raise largemouth bass to feed them a natural diet.

c) If the advice is incorrect, the researchers have committed a Type I error.

12. Risk.

These samples are independent, one sample of midsize cars and another of SUVs. The appropriate test is a two-sample t-test. There is no indication of random sampling. The back-to-back stemplot shows the distribution of the death rates for midsize cars is uniform, not unimodal and symmetric. The distribution of death rates for the SUVs has two outliers. With sample sizes this small, it is probably unwise to proceed with this inference.

Midsize		SUV
7	4	
4	5	5
4	6	02
6	7	6
8	8	
7	9	1
	10	9

8 | 2 = 82 deaths per million sales

13. Age.

a) **Independent groups assumption:** The group of patients with and without cardiac disease are not related in any way.
Randomization condition: Assume that these patients are representative of all people.
10% condition: 2397 patients without cardiac disease and 450 patients with cardiac disease are both less than 10% of all people.
Normal population assumption: We don't have the actual data, so we will assume that the population of ages of patients is Normal.

Since the conditions are satisfied, it is appropriate to model the sampling distribution of the difference in means with a Student's *t*-model, with 670 degrees of freedom (from the approximation formula). We will construct a two-sample *t*-interval, with 95% confidence.

$$(\bar{y}_{Card} - \bar{y}_{None}) \pm t^*_{df}\sqrt{\frac{s^2_{Card}}{n_{Card}} + \frac{s^2_{None}}{n_{None}}} = (74.0 - 69.8) \pm t^*_{670}\sqrt{\frac{7.9^2}{450} + \frac{8.7^2}{2397}} \approx (3.39, 5.01)$$

We are 95% confident that the mean age of patients with cardiac disease is between 3.39 and 5.01 years higher than the mean age of patients without cardiac disease.

b) Older patients are at greater risk for a variety of health problems. If an older patient does not survive a heart attack, the researchers will not know to what extent depression was involved, because there will be a variety of other possible variables influencing the death rate. Additionally, older patients may be more (or less) likely to be depressed than younger ones.

14. Smoking.

Randomization condition: Assume that these patients are representative of all people.
10% condition: 2397 patients without cardiac disease and 450 patients with cardiac disease are both less than 10% of all people.
Independent groups assumption: The group of patients with and without cardiac disease are not related in any way.
Success/Failure condition: $n\hat{p}$(cardiac) = (450)(0.32) = 144, $n\hat{q}$(cardiac) = (450)(0.68) = 306, $n\hat{p}$(none) = (2397)(0.237) = 568, and $n\hat{q}$(none) = (2397)(0.763) = 1829 are all greater than 10, so the samples are both large enough.

Since the conditions have been satisfied, we will find a two-proportion z-interval.

$$\left(\hat{p}_{Card} - \hat{p}_{None}\right) \pm z^* \sqrt{\frac{\hat{p}_{Card}\hat{q}_{Card}}{n_{Card}} + \frac{\hat{p}_{None}\hat{q}_{None}}{n_{None}}} = (0.32 - 0.237) \pm 1.960\sqrt{\frac{(0.32)(0.68)}{450} + \frac{(0.237)(0.763)}{2397}}$$
$$= (0.0367, 0.1293)$$

We are 95% confident that the proportion of smokers is between 3.67% and 12.93% higher for patients with cardiac disease than for patients without cardiac disease.

b) Since the confidence interval does not contain 0, there is evidence that cardiac patients have a higher rate of smokers than the patients without cardiac disease. The two groups are different.

c) Smoking could be a confounding variable. Smokers have a higher risk of other health problems that may be associated with their ability to survive a heart attack.

15. Computer use.

a) It is unlikely that an equal number of boys and girls were contacted strictly by chance. It is likely that this was a stratified random sample, stratified by gender.

b) **Randomization condition:** The teens were selected at random.
10% condition: 620 boys and 620 girls are both less than 10% of all teens.
Independent groups assumption: The groups of boys and girls are not paired or otherwise related in any way.
Success/Failure condition: $n\hat{p}$(boys) = (620)(0.77) = 477, $n\hat{q}$(boys) = (620)(0.23) = 143, $n\hat{p}$(girls) = (620)(0.65) = 403, and $n\hat{q}$(girls) = (620)(0.35) = 217 are all greater than 10, so the samples are both large enough.

Since the conditions have been satisfied, we will find a two-proportion z-interval.

$$\left(\hat{p}_B - \hat{p}_G\right) \pm z^* \sqrt{\frac{\hat{p}_B\hat{q}_B}{n_B} + \frac{\hat{p}_G\hat{q}_G}{n_G}} = (0.77 - 0.65) \pm 1.960\sqrt{\frac{(0.77)(0.33)}{620} + \frac{(0.65)(0.35)}{620}} = (0.070, 0.170)$$

We are 95% confident that the proportion of computer gamers is between 7.0% and 17.0% higher for boys than for girls.

c) Since the interval lies entirely above 0, there is evidence that a greater percentage of boys play computer games than girls.

16. Recruiting.

a) **Randomization condition:** Assume that these students are representative of all admitted graduate students.
10% condition: Two groups of 500 students are both less than 10% of all students.
Independent groups assumption: The groups are from two different years.
Success/Failure condition: $n\hat{p}$(before) = (500)(0.52) = 260, $n\hat{q}$(before) = (500)(0.48) = 240, $n\hat{p}$(after) = (500)(0.54) = 270, and $n\hat{q}$(after) = (500)(0.46) = 230 are all greater than 10, so the samples are both large enough.

Since the conditions have been satisfied, we will find a two-proportion *z*-interval.

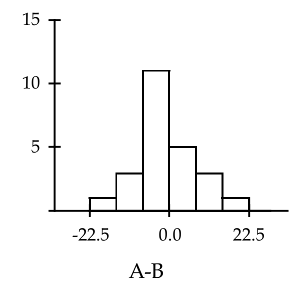
$$(\hat{p}_A - \hat{p}_B) \pm z^* \sqrt{\frac{\hat{p}_A\hat{q}_A}{n_A} + \frac{\hat{p}_B\hat{q}_B}{n_B}} = (0.54 - 0.52) \pm 1.960\sqrt{\frac{(0.54)(0.46)}{500} + \frac{(0.52)(0.48)}{500}} = (-0.042, 0.082)$$

We are 95% confident that the change in proportion of students who choose to enroll is between –4.2% and 8.2%.

b) Since 0 is contained in the interval, there is no evidence to suggest a change in the proportion of students who enroll. The program does not appear to be effective.

17. Hearing.

Paired data assumption: The data are paired by subject.
Randomization condition: The order of the tapes was randomized.
10% condition: We are testing the tapes, not the people, so this condition does not need to be checked.
Normal population assumption: The histogram of differences between List A and List B is roughly unimodal and symmetric.



Since the conditions are satisfied, the sampling distribution of the difference can be modeled with a Student's *t*-model with 24 – 1 = 23 degrees of freedom. We will find a paired *t*-interval, with 95% confidence.

$$\bar{d} \pm t^*_{n-1}\left(\frac{s_d}{\sqrt{n}}\right) = -0.\bar{3} \pm t^*_{23}\left(\frac{8.12225}{\sqrt{24}}\right) \approx (-3.76, 3.10)$$

We are 95% confident that the mean difference in the number of words a person might misunderstand using these two lists is between – 3.76 and 3.10 words. Since 0 is contained in the interval, there is no evidence to suggest that that the two lists are different for the purposes of the hearing test when there is background noise. It is reasonable to think that the two lists are still equivalent.

18. Cesareans.

H_0: The proportion of births involving cesarean deliveries is the same in Vermont and New Hampshire. $(p_{VT} = p_{NH} \text{ or } p_{VT} - p_{NH} = 0)$

H_A: The proportion of births involving cesarean deliveries is different in Vermont and New Hampshire. $(p_{VT} \neq p_{NH} \text{ or } p_{VT} - p_{NH} \neq 0)$

Random condition: Hospitals were randomly selected.
10% condition: 223 and 186 are both less than 10% of all births in these states.
Independent samples condition: Vermont and New Hampshire are different states!
Success/Failure condition: $n\hat{p}$(VT) = (223)(0.166) = 37, $n\hat{q}$(VT) = (223)(0.834) = 186, $n\hat{p}$(NH) = (186)(0.188) = 35, and $n\hat{q}$(NH) = (186)(0.812) = 151 are all greater than 10, so the samples are both large enough.

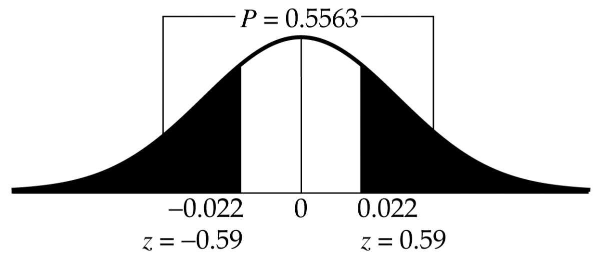
Since the conditions have been satisfied, we will model the sampling distribution of the difference in proportion with a Normal model with mean 0 and standard deviation estimated by

$$SE_{\text{pooled}}\left(\hat{p}_{VT}-\hat{p}_{NH}\right)=\sqrt{\frac{\hat{p}_{\text{pooled}}\hat{q}_{\text{pooled}}}{n_{VT}}+\frac{\hat{p}_{\text{pooled}}\hat{q}_{\text{pooled}}}{n_{NH}}}=\sqrt{\frac{(0.176)(0.824)}{223}+\frac{(0.176)(0.824)}{186}}\approx 0.03782.$$

The observed difference between the proportions is 0.166 – 0.188 = – 0.022.

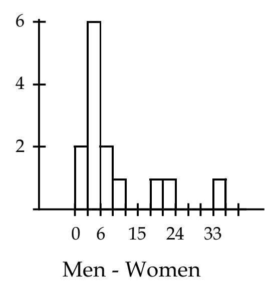
Since the *P*-value = 0.5563 is high, we fail to reject the null hypothesis. There is no evidence that the proportion of cesarean births in Vermont is different from the proportion of cesarean births in New Hampshire.

$$z=\frac{-0.22-0}{0.03782}$$
$$z\approx -0.59$$



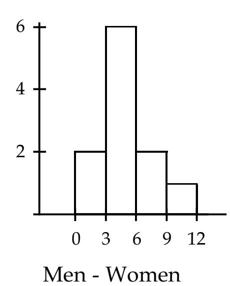
19. Newspapers.

a) An examination of a graphical display reveals Spain, Portugal, and Italy to be outliers. They are all Mediterranean countries, and all have a significantly higher percentage of men than women reading a newspaper daily.



b) H_0: The mean difference in the percentage of men and women who read a daily newspaper in these countries is zero. $(\mu_d = 0)$

H_A: The mean difference in the percentage of men and women who read a daily newspaper in these countries is greater than zero. $(\mu_d > 0)$



Paired data assumption: The data are paired by country.
Randomization condition: Samples in each country were random.
10% condition: The 1000 respondents in each country are less than 10% of the people in these countries.
Nearly Normal condition: With three outliers removed, the distribution of differences is roughly unimodal and symmetric.

Since the conditions are satisfied, the sampling distribution of the difference can be modeled with a Student's *t*-model with 11 – 1 = 10 degrees of freedom, $t_{10}\left(0,\frac{2.83668}{\sqrt{11}}\right)$.

We will use a paired *t*-test, (Men – Women) with $\bar{d}=4.75455$.

Since the P-value = 0.0001 is very low, we reject the null hypothesis. There is strong evidence that the mean difference is greater than zero. The percentage of men in these countries who read the paper daily appears to be greater than the percentage of women who do so.

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{11}}}$$

$$t \approx \frac{4.75455 - 0}{\frac{2.83668}{\sqrt{11}}}$$

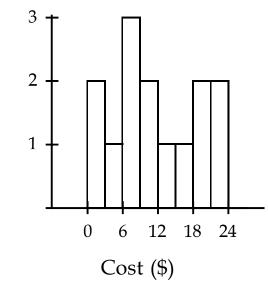
$$t \approx 5.56$$

20. Meals.

H_0: The college student's mean daily food expense is \$10. $(\mu = 10)$

H_A: The college student's mean daily food expense is greater than \$10. $(\mu > 10)$

Randomization condition: Assume that these days are representative of all days.
10% condition: 14 days are less than 10% of all days.
Nearly Normal condition: The histogram of daily expenses is fairly unimodal and symmetric. It is reasonable to think that this sample came from a Normal population.



The expenses in the sample had a mean of 11.4243 dollars and a standard deviation of 8.05794 dollars. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean daily expense with a Student's t model, with 14 – 1 = 13 degrees of freedom, $t_{13}\left(10, \frac{8.05794}{\sqrt{14}}\right)$.

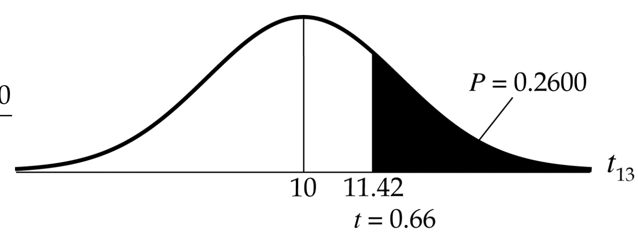
We will perform a one-sample t-test.

Since the P-value = 0.2600 is high, we fail to reject the null hypothesis. There is no evidence that the student's average spending is more than \$10 per day.

$$t = \frac{\bar{y} - \mu_0}{SE(\bar{y})}$$

$$t = \frac{11.4243 - 10}{\frac{8.05794}{\sqrt{14}}}$$

$$t \approx 0.66$$



21. Wall Street.

Random condition: 1002 American adults were randomly selected.
10% condition: 1002 is less than 10% of all American adults.
Success/Failure condition: $n\hat{p}$ = (1002)(0.60) = 601 and $n\hat{q}$ = (1002)(0.40) = 401 are both greater than 10, so the sample is large enough.

Since the conditions are met, we can use a one-proportion z-interval to estimate the percentage of American adults who agree with the statement.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} = (0.60) \pm 1.960\sqrt{\frac{(0.60)(0.40)}{1002}} \approx (57.0\%, 63.0\%)$$

We are 95% confident that between 57.0% and 63.0% of American adults would agree with the statement about Wall Street.

22. Teach for America.

H_0: The mean score of students with certified teachers is the same as the mean score of students with uncertified teachers. $(\mu_C = \mu_U \text{ or } \mu_C - \mu_U = 0)$

H_A: The mean score of students with certified teachers is greater than as the mean score of students with uncertified teachers. $(\mu_C > \mu_U \text{ or } \mu_C - \mu_U > 0)$

Independent groups assumption: The certified and uncertified teachers are independent groups.
Randomization condition: Assume the students studied were representative of all students.
10% condition: Two samples of size 44 are both less than 10% of the population.
Nearly Normal condition: We don't have the actual data, so we can't look at the graphical displays, but the sample sizes are large, so we can proceed.

Since the conditions are satisfied, it is appropriate to model the sampling distribution of the difference in means with a Student's *t*-model, with 86 degrees of freedom (from the approximation formula).

We will perform a two-sample *t*-test. The sampling distribution model has mean 0, with standard error: $SE(\bar{y}_C - \bar{y}_U) = \sqrt{\dfrac{9.31^2}{44} + \dfrac{9.43^2}{44}} \approx 1.9977$.

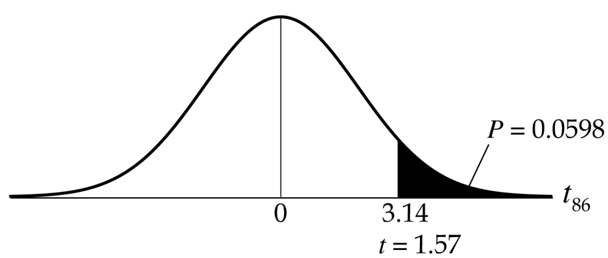
The observed difference between the mean scores is 35.62 – 32.48 = 3.14.

Since the *P*-value = 0.0598 is fairly high, we fail to reject the null hypothesis. There is little evidence that students with certified teachers had mean scores higher than students with uncertified teachers. However, since the *P*-value is not extremely high, further investigation is recommended.

$$t = \frac{(\bar{y}_C - \bar{y}_U) - (0)}{SE(\bar{y}_C - \bar{y}_U)}$$

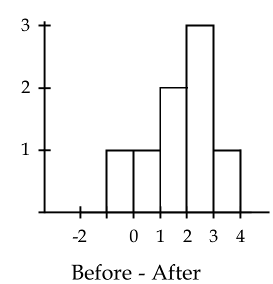
$$t \approx \frac{3.14}{1.9977}$$

$$t \approx 1.57$$



23. Legionnaires' disease.

Paired data assumption: The data are paired by room.
Randomization condition: We will assume that these rooms are representative of all rooms at the hotel.
10% condition: Assume that 8 rooms are less than 10% of the rooms. The hotel must have more than 80 rooms.
Nearly Normal condition: The histogram of differences between before and after measurements is roughly unimodal and symmetric.



Since the conditions are satisfied, the sampling distribution of the difference can be modeled with a Student's *t*-model with 8 – 1 = 7 degrees of freedom. We will find a paired *t*-interval, with 95% confidence.

$$\bar{d} \pm t_{n-1}^{*}\left(\frac{s_d}{\sqrt{n}}\right) = 1.6125 \pm t_7^{*}\left(\frac{1.23801}{\sqrt{8}}\right) \approx (0.58,\ 2.65)$$

We are 95% confident that the mean difference in the bacteria counts is between 0.58 and 2.65 colonies per cubic foot of air. Since the entire interval is above 0, there is evidence that the new air-conditioning system was effective in reducing average bacteria counts.

24. Teach for America, Part II.

H_0: The mean score of students with certified teachers is the same as the mean score of students with uncertified teachers. $(\mu_C = \mu_U \text{ or } \mu_C - \mu_U = 0)$

H_A: The mean score of students with certified teachers is different than the mean score of students with uncertified teachers. $(\mu_C \neq \mu_U \text{ or } \mu_C - \mu_U \neq 0)$

Mathematics: Since the *P*-value = 0.0002 is low, we reject the null hypothesis. There is strong evidence that students with certified teachers have different mean math scores than students with uncertified teachers. Students with certified teachers do better.

Language: Since the *P*-value = 0.045 is fairly low, we reject the null hypothesis. There is evidence that students with certified teachers have different mean language scores than students with uncertified teachers. Students with certified teachers do better. However, since the *P*-value is not extremely low, further investigation is recommended.

25. Bipolar kids.

a) **Random condition:** Assume that the 89 children are representative of all children with bipolar disorder.
10% condition: 89 is less than 10% of all children with bipolar disorder.
Success/Failure condition: $n\hat{p}$ = 26 and $n\hat{q}$ = 63 are both greater than 10, so the sample is large enough.

Since the conditions are met, we can use a one-proportion *z*-interval to estimate the percentage of children with bipolar disorder who might be helped by this treatment.

$$\hat{p} \pm z^{*}\sqrt{\frac{\hat{p}\hat{q}}{n}} = \left(\frac{26}{89}\right) \pm 1.960\sqrt{\frac{\left(\frac{26}{89}\right)\left(\frac{63}{89}\right)}{89}} \approx (19.77\%, 38.66\%)$$

We are 95% confident that between 19.77% and 38.66% of children with bipolar disorder will be helped with medication and psychotherapy.

b)

$$ME = z^{*}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.06 = 1.960\sqrt{\frac{\left(\frac{26}{89}\right)\left(\frac{63}{89}\right)}{n}}$$

$$n = \frac{(1.960)^2\left(\frac{26}{89}\right)\left(\frac{63}{89}\right)}{(0.06)^2}$$

$$n \approx 221 \text{ children}$$

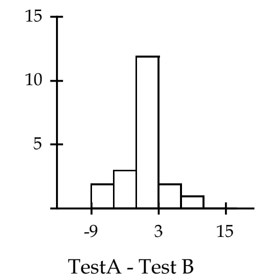
In order to estimate the proportion of children helped with medication and psychotherapy to within 6% with 95% confidence, we would need a sample of at least 221 children. All decimals in the final answer must be rounded up, to the next person.
(For a more cautious answer, let $\hat{p} = \hat{q} = 0.5$. This method results in a required sample of 267 children.)

26. Online testing.

a) H_0: The mean difference in the scores between Test A and Test B is zero. $(\mu_d = 0)$

H_A: The mean difference in the scores between Test A and Test B is not zero. $(\mu_d \neq 0)$

Paired data assumption: The data are paired by student.
Randomization condition: The volunteers were randomized with respect to the order in which they took the tests and which form they took in each environment.
10% condition: We are testing the difficulty of the tests, not the people, so we don't need to check this condition.
Nearly Normal condition: The distribution of differences is roughly unimodal and symmetric.



Since the conditions are satisfied, the sampling distribution of the difference can be modeled with a Student's *t*-model with 20 – 1 = 19 degrees of freedom, $t_{19}\left(0, \frac{3.52584}{\sqrt{20}}\right)$.

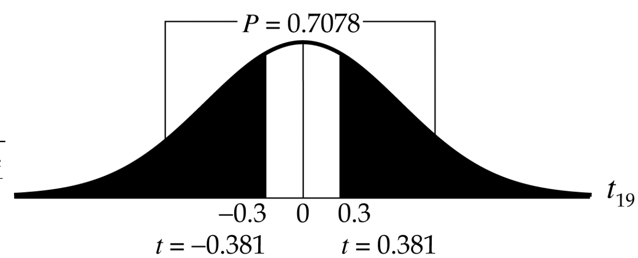
We will use a paired *t*-test, (Test A – Test B) with $\bar{d} = 0.3$.

Since the *P*-value = 0.7078 is very high, we fail to reject the null hypothesis. There is no evidence of that the mean difference in score is different from zero. It is reasonable to think that Test A and Test B are equivalent in terms of difficulty.

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}}$$

$$t \approx \frac{0.3 - 0}{\frac{3.52584}{\sqrt{20}}}$$

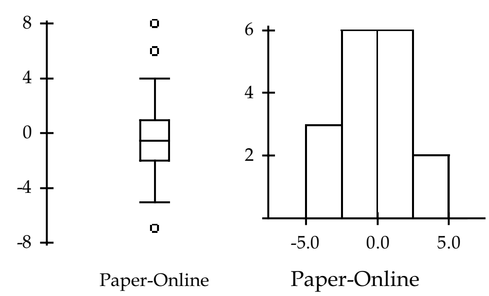
$$t \approx 0.381$$



b) H_0: The mean difference between paper scores and online scores is zero. $(\mu_d = 0)$

H_A: The mean difference between paper scores and online scores is not zero. $(\mu_d \neq 0)$

Paired data assumption: The data are paired by student.
Randomization condition: The volunteers were randomized with respect to the order in which they took the tests and which form they took in each environment.
10% condition: We are testing the formats of the test (paper/online), not the people, so we don't need to check this condition.



Nearly Normal condition: The boxplot of the distribution of differences is shows three outliers: students 3, 10, and 17. With these outliers removed, the histogram is roughly unimodal and symmetric.

Since the conditions are satisfied, the sampling distribution of the difference can be modeled with a Student's *t*-model with 17 – 1 = 16 degrees of freedom, $t_{16}\left(0, \frac{2.26222}{\sqrt{17}}\right)$.

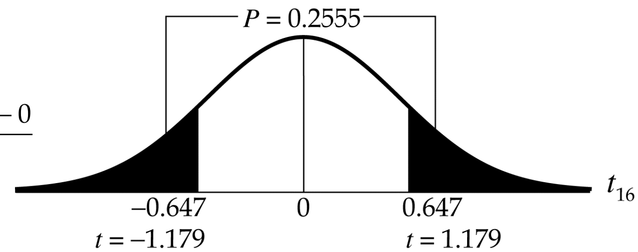
We will use a paired *t*-test, (Paper – Online) with $\bar{d} = -0.647059$.

Since the P-value = 0.2555 is high, we fail to reject the null hypothesis. There is no evidence that the mean difference in score is different from zero. It is reasonable to think that paper and online tests are equivalent in terms of difficulty.

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}}$$

$$t \approx \frac{-0.647059 - 0}{\frac{2.26222}{\sqrt{17}}}$$

$$t \approx -1.179$$



27. Bread.

a) Since the histogram shows that the distribution of the number of loaves sold per day is skewed strongly to the right, we can't use the Normal model to estimate the number of loaves sold on the busiest 10% of days.

b) **Randomization condition:** Assume that these days are representative of all days.
10% condition: 100 days are less than 10% of all days.
Nearly Normal condition: The histogram is skewed strongly to the right. However, since the sample size is large, the Central Limit Theorem guarantees that the distribution of averages will be approximately Normal.

The days in the sample had a mean of 103 loaves sold and a standard deviation of 9 loaves sold. Since the conditions are satisfied, the sampling distribution of the mean can be modeled by a Student's t- model, with 103 – 1 = 103 degrees of freedom. We will use a one-sample t-interval with 95% confidence for the mean number of loaves sold.
(By hand, use $t^*_{50} \approx 2.403$ from the table.)

c) $\bar{y} \pm t^*_{n-1}\left(\frac{s}{\sqrt{n}}\right) = 103 \pm t^*_{102}\left(\frac{9}{\sqrt{103}}\right) \approx (101.2, 104.8)$

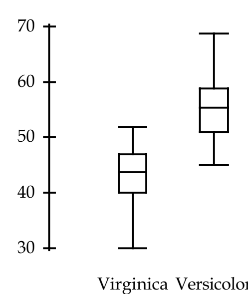
We are 95% confident that the mean number of loaves sold per day at the Clarksburg Bakery is between 101.2 and 104.8.

d) We know that in order to cut the margin of error in half, we need to a sample four times as large. If we allow a margin of error that is twice as wide, that would require a sample only one-fourth the size. In this case, our original sample is 100 loaves; so 25 loaves would be a sufficient number to estimate the mean with a margin of error twice as wide.

e) Since the interval is completely above 100 loaves, there is strong evidence that the estimate was incorrect. The evidence suggests that the mean number of loaves sold per day is greater than 100. This difference is statistically significant, but may not be practically significant. It seems like the owners made a pretty good estimate!

28. Irises.

a) Parallel boxplots of the distributions of petal lengths for the two species of flower are at the right. No units are specified, but millimeters seems like a reasonable guess.



b) The petals of *versicolor* are generally longer than the petals of *virginica.* Both distributions have about the same range, and both distributions are fairly symmetric.

c) **Independent groups assumption:** The two species of flowers are independent.
Randomization condition: It is reasonable to assume that these flowers are representative of their species.
10% condition: Two samples of 50 flowers each are less than 10% of all flowers.
Nearly Normal condition: The boxplots show distributions of petal lengths that are reasonably symmetric with no outliers. Additionally, the samples are large.

Since the conditions are satisfied, it is appropriate to model the sampling distribution of the difference in means with a Student's *t*-model, with 97.92 degrees of freedom (from the approximation formula). We will construct a two-sample *t*-interval, with 95% confidence.

$$(\bar{y}_{Ver} - \bar{y}_{Vir}) \pm t^*_{df}\sqrt{\frac{s^2_{Ver}}{n_{Ver}} + \frac{s^2_{Vir}}{n_{Vir}}} = (55.52 - 43.22) \pm t^*_{97.92}\sqrt{\frac{5.51895^2}{50} + \frac{5.36158^2}{50}} \approx (10.14,\ 14.46)$$

d) We are 95% confident the mean petal length of *versicolor* irises is between 10.14 and 14.46 millimeters longer than the mean petal length of *virginica* irises.

e) Since the interval is completely above 0, there is strong evidence that the mean petal length of *versicolor* irises is greater than the mean petal length of *virginica* irises.

29. Insulin and diet.

a) H_0: People with high dairy consumption have IRS at the same rate as those with low dairy consumption. $\left(p_{High} = p_{Low} \text{ or } p_{High} - p_{Low} = 0\right)$

H_A: People with high dairy consumption have IRS at a different rate than those with low dairy consumption. $\left(p_{High} \neq p_{Low} \text{ or } p_{High} - p_{Low} \neq 0\right)$

Random condition: Assume that the people studied are representative of all people.
10% condition: 102 and 190 are both less than 10% of all people.
Independent samples condition: The two groups are not related.
Success/Failure condition: $n\hat{p}$(high) = 24, $n\hat{q}$(high) = 78, $n\hat{p}$(low) = 85, and $n\hat{q}$(low) = 105 are all greater than 10, so the samples are both large enough.

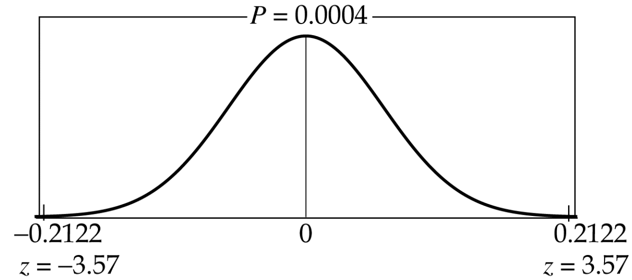
Since the conditions have been satisfied, we will model the sampling distribution of the difference in proportion with a Normal model with mean 0 and standard deviation estimated by

$$SE_{\text{pooled}}\left(\hat{p}_{High} - \hat{p}_{Low}\right) = \sqrt{\frac{\hat{p}_{\text{pooled}}\hat{q}_{\text{pooled}}}{n_{High}} + \frac{\hat{p}_{\text{pooled}}\hat{q}_{\text{pooled}}}{n_{Low}}} = \sqrt{\frac{(0.373)(0.627)}{102} + \frac{(0.373)(0.627)}{190}} \approx 0.05936.$$

The observed difference between the proportions is 0.2352 – 0.4474 = – 0.2122.

Since the P-value = 0.0004 is very low, we reject the null hypothesis. There is strong evidence that the proportion of people with IRS is different for those who with high dairy consumption compared to those with low dairy consumption. People who consume dairy products more than 35 times per week appear less likely to have IRS than those who consume dairy products fewer than 10 times per week.

$$z = \frac{-0.2122 - 0}{0.05936}$$
$$z \approx -3.57$$



b) There is evidence of an association between the low consumption of dairy products and IRS, but that does not prove that dairy consumption influences the development of IRS. This is an observational study, and a controlled experiment is required to prove cause and effect.

30. Speeding.

a) H_0 : The percentage of speeding tickets issued to black drivers is 16%, the same as the percentage of registered drivers who are black. (p = 0.16)

H_A : The percentage of speeding tickets issued to black drivers is greater than 16%, the percentage of registered drivers who are black. (p > 0.16)

Random condition: Assume that this month is representative of all months with respect to the percentage of tickets issued to black drivers.
10% condition: 324 speeding tickets are less than 10% of all tickets.
Success/Failure condition: $np = (324)(0.16) = 52$ and $nq = (324)(0.84) = 272$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.16$ and
$\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.16)(0.84)}{324}} \approx 0.02037$.

We can perform a one-proportion z-test.
The observed proportion of tickets issued to black drivers is $\hat{p} = 0.25$.

Since the P-value = 4.96×10^{-6} is very low, we reject the null hypothesis. There is strong evidence that the percentage of speeding tickets issued to black drivers is greater than 16%.

$$z = \frac{\hat{p} - p_0}{\sigma(\hat{p})}$$
$$z \approx \frac{0.25 - 0.16}{0.02037}$$
$$z \approx 4.42$$

b) There is strong evidence of an association between the receipt of a speeding ticket and race. Black drivers appear to be issued tickets at a higher rate than expected. However, this does not prove that racial profiling exists. There may be other factors present.

c) Answers may vary. The primary statistic of interest is the percentage of black motorists on this section of the New Jersey Turnpike. For example, if 80% of drivers on this section are black, then 25% of the speeding tickets being issued to black motorists is not an usually high percentage. In fact, it is probably unusually low. On the other hand, if only 3% of the motorists on this section of the turnpike are black, then there is even more evidence that racial profiling may be occurring.

31. Rainmakers?

Independent groups assumption: The two groups of clouds are independent.
Randomization condition: Researchers randomly assigned clouds to be seeded with silver iodide or not seeded.
10% condition: We are testing cloud seeding, not clouds themselves, so this condition doesn't need to be checked.
Nearly Normal condition: We don't have the actual data, so we can't look at the distributions, but the means of group are significantly higher than the medians. This is an indication that the distributions are skewed to the right, with possible outliers. The samples sizes of 26 each are fairly large, so it should be safe to proceed, but we should be careful making conclusions, since there may be outliers.

Since the conditions are satisfied, it is appropriate to model the sampling distribution of the difference in means with a Student's *t*-model, with 33.86 degrees of freedom (from the approximation formula). We will construct a two-sample *t*-interval, with 95% confidence.

$$(\bar{y}_S - \bar{y}_U) \pm t^*_{df}\sqrt{\frac{s_S^2}{n_S} + \frac{s_U^2}{n_U}} = (441.985 - 164.588) \pm t^*_{33.86}\sqrt{\frac{650.787^2}{26} + \frac{278.426^2}{26}} \approx (-4.76, 559.56)$$

We are 95% confident the mean amount of rainfall produced by seeded clouds is between 4.76 acre-feet less than and 559.56 acre-feet more than the mean amount of rainfall produced by unseeded clouds.

Since the interval contains 0, there is little evidence that the mean rainfall produced by seeded clouds is any different from the mean rainfall produced by unseeded clouds. However, we shouldn't place too much faith in this conclusion. It is based on a procedure that is sensitive to outliers, and there may have been outliers present.

32. Fritos.

a) H_0: The mean weight of bags of Fritos is 35.4 grams. $(\mu = 35.4)$

H_A: The mean weight of bags of Fritos is less than 35.4 grams. $(\mu < 35.4)$

b) **Randomization condition:** It is reasonable to think that the 6 bags are representative of all bags of Fritos.
10% condition: 6 bags are less than 10% of all bags.
Nearly Normal condition: The histogram of bags weights shows one unusually heavy bag. Although not technically an outlier, it probably should be excluded for the purposes of the test. (We will leave it in for the preliminary test, then remove it and test again.)

c) The bags in the sample had a mean weight of 35.5333 grams and a standard deviation in weight of 0.450185 grams. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean weight of bags of Fritos with a Student's t model, with 6 – 1 = 5 degrees of freedom, $t_5\left(35.4, \frac{0.450185}{\sqrt{6}}\right)$.

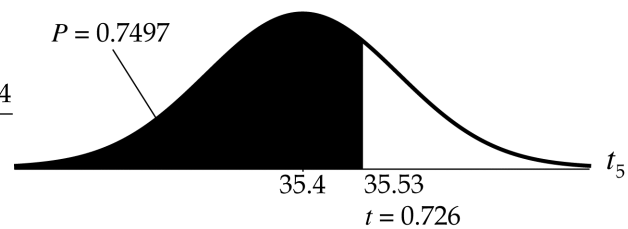
We will perform a one-sample t-test.

Since the P-value = 0.7497 is high, we fail to reject the null hypothesis. There is no evidence to suggest that the mean weight of bags of Fritos is less than 35.4 grams.

$$t = \frac{\bar{y} - \mu_0}{SE(\bar{y})}$$

$$t = \frac{35.5333 - 35.4}{\frac{0.450185}{\sqrt{6}}}$$

$$t = 0.726$$



d) With the one unusually high value removed, the mean weight of the 5 remaining bags is 35.36 grams, with a standard deviation in weight of 0.167332 grams. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean weight of bags of Fritos with a Student's t model, with 5 – 1 = 4 degrees of freedom, $t_4\left(35.4, \frac{0.167332}{\sqrt{5}}\right)$.

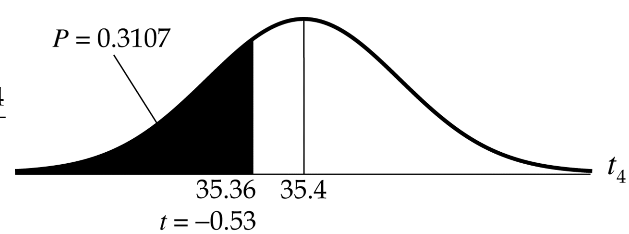
We will perform a one-sample t-test.

Since the P-value = 0.3107 is high, we fail to reject the null hypothesis. There is no evidence to suggest that the mean weight of bags of Fritos is less than 35.4 grams.

$$t = \frac{\bar{y} - \mu_0}{SE(\bar{y})}$$

$$t = \frac{35.36 - 35.4}{\frac{0.167332}{\sqrt{5}}}$$

$$t = -0.53$$



e) Neither test provides evidence that the mean weight of bags of Fritos is less than 35.4 grams. It is reasonable to believe that the mean weight of the bags is the same as the stated weight. However, the sample sizes are very small, and the tests have very little power to detect lower mean weights. It would be a good idea to weigh more bags.

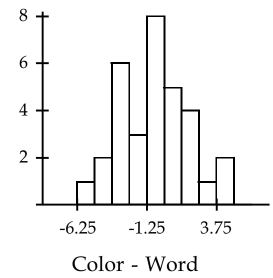
33. Color or text?

a) By randomizing the order of the cards (shuffling), the researchers are attempting to avoid bias that may result from volunteers remembering the order of the cards from the first task. Although mentioned, hopefully the researchers are randomizing the order in which the tasks are performed. For example, if each volunteer performs the color task first, they may all do better on the text task, simply because they have had some practice in memorizing cards. By randomizing the order, that bias is controlled.

b) H_0: The mean difference between color and word scores is zero. $(\mu_d = 0)$

H_A: The mean difference between color and word scores is not zero. $(\mu_d \neq 0)$

Paired data assumption: The data are paired by volunteer.
Randomization condition: Hopefully, the volunteers were randomized with respect to the order in which they performed the tasks, and the cards were shuffled between tasks.
10% condition: We are testing the format of the task (color/word), not the people, so we don't need to check this condition.
Nearly Normal condition: The histogram of the differences is roughly unimodal and symmetric.



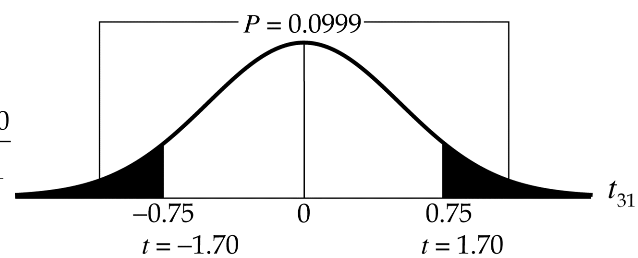
Since the conditions are satisfied, the sampling distribution of the difference can be modeled with a Student's t-model with 32 – 1 = 31 degrees of freedom, $t_{31}\left(0, \dfrac{2.50161}{\sqrt{32}}\right)$.

We will use a paired t-test, (color – word) with $\bar{d} = -0.75$.

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}}$$

$$t \approx \frac{-0.75 - 0}{\frac{2.50161}{\sqrt{32}}}$$

$$t \approx -1.70$$



Since the P-value = 0.0999 is high, we fail to reject the null hypothesis. There is no evidence that the mean difference in score is different from zero. It is reasonable to think that neither color nor word dominates perception. However, if the order in which volunteers took the test was not randomized, making conclusions from this test may be risky.

34. And it means?

a) The margin of error is $\dfrac{(2391 - 1644)}{2} = \$373.50$.

b) The insurance agent is 95% confident that the mean loss claimed by clients after home burglaries is between \$1644 and \$2391.

c) 95% of all random samples of this size will produce intervals that contain the true mean loss claimed.

35. Batteries.

a) Different samples have different means. Since this is a fairly small sample, the difference may be due to natural sampling variation. Also, we have no idea how to quantify "a lot less" with out considering the variation as measured by the standard deviation.

b) H_0: The mean life of a battery is 100 hours. $(\mu = 100)$

H_A: The mean life of a battery is less than 100 hours. $(\mu < 100)$

c) **Randomization condition:** It is reasonable to think that these 16 batteries are representative of all batteries of this type
10% condition: 16 batteries are less than 10% of all batteries.
Normal population assumption: Since we don't have the actual data, we can't check a graphical display, and the sample is not large. Assume that the population of battery lifetimes is Normal.

d) The batteries in the sample had a mean life of 97 hours and a standard deviation of 12 hours. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean battery life with a Student's *t* model, with 16 – 1 = 15 degrees of freedom, $t_{15}\left(100, \frac{12}{\sqrt{16}}\right)$, or $t_{15}(100, 3)$

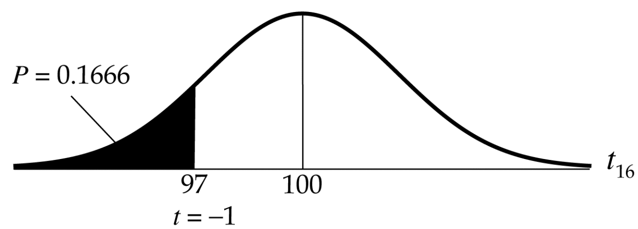
We will perform a one-sample *t*-test.

Since the *P*-value = 0.1666 is greater than $\alpha = 0.05$, we fail to reject the null hypothesis. There is no evidence to suggest that the mean battery life is less than 100 hours.

$$t = \frac{\bar{y} - \mu_0}{SE(\bar{y})}$$

$$t = \frac{97 - 100}{\frac{12}{\sqrt{16}}}$$

$$t = -1$$



e) If the mean life of the company's batteries is only 98 hours, then the mean life is less than 100, and the null hypothesis is false. We failed to reject a false null hypothesis, making a Type II error.

36. Hamsters.

a) **Randomization condition:** Assume that these litters are representative of all litters.
10% condition: 47 litters are less than 10% of all litters.
Nearly Normal condition: We don't have the actual data, so we can't look at a graphical display. However, since the sample size is large, the Central Limit Theorem guarantees that the distribution of averages will be approximately Normal, as long as there are no outliers.

The litters in the sample had a mean size of 7.72 baby hamsters and a standard deviation of 2.5 baby hamsters. Since the conditions are satisfied, the sampling distribution of the mean can be modeled by a Student's *t*- model, with 47 – 1 = 46 degrees of freedom. We will use a one-sample *t*-interval with 90% confidence for the mean number of baby hamsters per litter.

$$\bar{y} \pm t^*_{n-1}\left(\frac{s}{\sqrt{n}}\right) = 7.72 \pm t^*_{46}\left(\frac{2.5}{\sqrt{47}}\right) \approx (7.11, 8.33)$$

We are 90% confident that the mean number of baby hamsters per litter is between 7.11 and 8.33.

b) A 98% confidence interval would have a larger margin of error. Higher levels of confidence come at the price of less precision in the estimate.

c) A quick estimate using *z* gives us a sample size of about 25 litters. Using this estimate, $t^*_{24} = 2.064$ at 95% confidence. We need a sample of about 27 litters in order to estimate the number of baby hamsters per litter to within 1 baby hamster.

$$ME = t^*_{24}\left(\frac{s}{\sqrt{n}}\right)$$

$$1 = 2.064\left(\frac{2.5}{\sqrt{n}}\right)$$

$$n = \frac{(2.064)^2(2.5)^2}{(1)^2}$$

$$n \approx 27$$