Practice Question Set #2 (Chapters 19 - 25)

Question	Reference	Notes	Page
1	19.6		525
2	19.12		525
3	19.14		5.25
4	19.30		527
5	20.2		549
6	20.4		549
7	20.10		550
8	20.12		550
9	21.12		580
10	21.14		580
11	22.10		604
12	22.14		605
13	22.16		605
14	22.26		606
15	22.28		606
16	23.14		646
17	23.30		648
18	24.10		681
19	24.18		682
20	24.42		687
21	25.14		708
22	25.30		712

1. (19.6)

More conclusions.

a) Not correct. This statement implies certainty. There is no level of confidence in the statement.

b) Not correct. We *know* that 56% of the spins in this experiment landed heads. There is no need to make an interval for the sample proportion.

c) Correct.

d) Correct.

e) Not correct. The interval should be about the proportion of heads, not the percentage of euros.

2. (19.12)

Contaminated chicken, second course.

a) **Independence Assumption**: It is very important that the researchers kept the chicken samples separated. Otherwise, *salmonella* could be transmitted from sample to sample.

Randomization Condition: It's not clear how the sample was chosen. We will assume the sample from 23 states is at least representative of all broiler chickens sold.

10% Condition: 525 is far less than 10% of all broiler chickens.

Success/Failure Condition: $n\hat{p} = 79$ and $n\hat{q} = 446$ are both greater than 10, so the sample is large enough.

b) Since the conditions are met, we can use a one-proportion z-interval to estimate the percentage of broiler chicken packages that are infected with *salmonella*.

$$\hat{p} \pm z^{*}\sqrt{\frac{\hat{p}\hat{q}}{n}} = (0.15) \pm 1.960\sqrt{\frac{(0.15)(0.85)}{525}} = (0.12, 0.18)$$

c) We are 95% confident that between 12% and 18% of all broiler chickens sold in U.S. food stores are infected with *salmonella*.

3. (19.14)

Canada's birthday 2009.

a) The margin of error is 1.96* √(0.21*0.79/1000) ≈ 0.025

b) We are 95% sure that the true proportion of Canadians who know Canada's age is within ±2.5% of 21%.

c) The margin of error would have to be smaller, if we wanted to be less certain of capturing the true proportion.

d) It is 1.645*√(0.21*0.79/1000) ≈ 0.021

e) Smaller samples are more variable, which translates to a larger margin of error.

4. (19.30)

Hiring.

a)

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

$$0.05 = 2.326\sqrt{\frac{(0.5)(0.5)}{n}}$$

$$n = \frac{(2.326)^2(0.5)(0.5)}{(0.05)^2}$$

$$n \approx 542 \text{ businesses}$$

In order to estimate the percentage of businesses planning to hire additional employees within the next 60 days to within 5% with 98% confidence, we would need a sample of at least 542 businesses. All decimals in the final answer must be rounded up, to the next integer.

b)

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

$$0.03 = 2.326\sqrt{\frac{(0.5)(0.5)}{n}}$$

$$n = \frac{(2.326)^2(0.5)(0.5)}{(0.03)^2}$$

$$n \approx 1503 \text{ businesses}$$

In order to estimate the percentage of businesses planning to hire additional employees within the next 60 days to within 3% with 98% confidence, we would need a sample of at least 1503 businesses. All decimals in the final answer must be rounded up, to the next integer.
(Alternatively, the margin of error is being decreased to 3/5 of its original size, so the sample size must increase by a factor of 25/9. $542(25/9) \approx 1506$ businesses. A bit off, because 542 was rounded, but close enough!

c)

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

$$0.01 = 2.326\sqrt{\frac{(0.5)(0.5)}{n}}$$

$$n = \frac{(2.326)^2(0.5)(0.5)}{(0.01)^2}$$

$$n \approx 13{,}526 \text{ businesses}$$

In order to estimate the percentage of businesses planning to hire additional employees within the next 60 days to within 1% with 98% confidence, we would need a sample of at least 13,526 businesses.
(Alternatively, the margin of error has been decreased to 1/5 of its original size, so a sample 25 times as large would be needed. 25(542) = 13,550. Close enough!
It would probably be very expensive and time consuming to sample that many businesses.

5. (20.2)

More hypotheses.

a) H_0 : The proportion of high school graduates is 40%. ($p = 0.40$)
   H_A : The proportion of high school graduates is not 40%. ($p \neq 0.40$)

b) H_0 : The proportion of cars needing transmission repair is 20%. ($p = 0.20$)
   H_A : The proportion of cars needing transmission repair is less than 20%. ($p < 0.20$)

c) H_0 : The proportion of people who like the flavour is 60%. ($p = 0.60$)
   H_A : The proportion of people who like the flavour is greater than 60%. ($p > 0.60$)

6. (20.4)

Dice. Statement d is the correct interpretation of a *P*-value.

7. (20.10)

Got milk?

1) Null and alternative hypotheses should involve p, not $\hat{p}$.
2) The question asks if there is evidence that the 90% figure is *not accurate*, so a two-sided alternative hypothesis should be used. H_A should be $p \neq 0.90$.
3) One of the conditions checked appears to be $n > 10$, which is not a condition for hypothesis tests. The Success/Failure Condition checks $np = (750)(0.90) = 675 > 10$ and
   $nq = (750)(0.10) = 75 > 10$. Also, the 10% condition is not verified.
4) $SD(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.90)(0.10)}{750}} \approx 0.011$. The student used rounded values of $\hat{p}$ and $\hat{q}$.
5) Value of z is incorrect. The correct value is $z = \frac{0.876 - 0.90}{0.011} \approx -2.18$.
6) The *P*-value calculated is in the wrong direction. To test the given hypothesis, the lower-tail probability should have been calculated. The correct, two-tailed *P*-values is
   $P = 2P(z < -2.18) = 0.029$.
7) The *P*-value is misinterpreted. Since the *P*-value is so low, there is moderately strong evidence that the proportion of adults who drink milk is different than the claimed 90%. In fact, our sample suggests that the proportion may be lower. There is only a 2.9% chance of observing a $\hat{p}$ as far from 0.90 as this, simply from natural sampling variation.

8. (20.12)

Abnormalities.

a) H_0 : The percentage of children with genetic abnormalities is 5%.(p = 0.05)
H_A : The percentage of children with genetic abnormalities is greater than 5%. (p > 0.05)

b) **Independence Assumption:** There is no reason to think that one child having genetic abnormalities would affect the probability that other children have them.
Randomization Condition: This sample may not be random, but genetic abnormalities are plausibly independent. The sample is probably representative of all children, with regards to genetic abnomalities.
10% Condition: The sample of 384 children is less than 10% of all children.
Success/Failure Condition: np = (384)(0.05) = 19.2 and nq = (384)(0.95) = 364.8 are both greater than 10, so the sample is large enough.

c) The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p =$ 0.05 and

$$\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.05)(0.95)}{384}} \approx 0.0111.$$

We can perform a one-proportion z-test. The observed proportion of children with genetic abnormalities is $\hat{p} = \frac{46}{384} \approx 0.1198$.

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

$$z = \frac{0.1198 - 0.05}{\sqrt{\frac{(0.05)(0.95)}{384}}}$$

$$z \approx 6.28$$

The value of z is approximately 6.28, meaning that the observed proportion of children with genetic abnormalities is over 6 standard deviations above the hypothesized proportion. The P-value associated with this z score is 2×10^{-10}, essentially 0.

d) If 5% of children have genetic abnormalities, the chance of observing 46 children with genetic abnormalities in a random sample of 384 children is essentially 0.

e) With a p-value this low there is very strong evidence against the null hypothesis in favour of the alternative. This is very strong evidence that more than 5% of children have genetic abnormalities.

f) We don't know that environmental chemicals cause genetic abnormalities. We merely have evidence that suggests that a greater percentage of children are diagnosed with genetic abnormalities now, compared to the 1980s.

9. (21.12)

Superdads.

a) **Independence Assumption:** One man's response is not likely to have any effect on another man's response.
Randomization Condition: The men were contacted through a random telephone poll.
10% Condition: 712 men represent less than 10% of all men.
Success/Failure Condition: $n\hat{p} = (712)(0.22) = 157$ and $n\hat{q} = (712)(0.78) = 555$ 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 men who identify themselves as stay-at-home dads.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} = (0.22) \pm 1.96\sqrt{\frac{(0.22)(0.78)}{712}} = (19.0\%, 25.1\%)$$

We are 95% confident that between 19.0% and 25.1% of all men identify themselves as stay-at-home dads.

b) Since the confidence interval extends well below 25%, we can't be confident that over 25% of men are stay-at-home dads. The company should not buy the ads.

c) Since, 25% is within the confidence interval, Spike could claim that it's possible that the true proportion of stay-at-home dads is 25%, but we can never prove that the null hypothesis is true.

10. (21.14)

Spam.

a) Type II. The filter decided that the message was safe, when in fact it was spam.

b) Type I. The filter decided that the message was spam, when in fact it was not.

c) This is analogous to lowering alpha. It takes more evidence to classify a message as spam.

d) The risk of Type I error is decreased and the risk of Type II error has increased.

11. (22.10)

Prostate cancer.

a) This was an experiment. Men were randomly assigned to imposed treatments. They were assigned to either have prostate surgery or assigned to not have prostate surgery.

b) **Randomization Condition**: The men were randomly assigned to the two treatment groups.

Independent Samples Condition: The groups were assigned randomly, so the groups are not related.

Success/Failure Condition: $n\hat{p}$ (surg) = 16, $n\hat{q}$ (surg) = 331, $n\hat{p}$ (none) = 31, and $n\hat{q}$ (none) = 317 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.

c) $$(\hat{p}_{None} - \hat{p}_{Surg}) \pm z^* \sqrt{\frac{\hat{p}_{None}\hat{q}_{None}}{n_{None}} + \frac{\hat{p}_{Surg}\hat{q}_{Surg}}{n_{Surg}}}$$

$$= \left(\tfrac{31}{348} - \tfrac{16}{347}\right) \pm 1.960 \sqrt{\frac{\left(\frac{31}{348}\right)\left(\frac{317}{348}\right)}{348} + \frac{\left(\frac{16}{347}\right)\left(\frac{331}{347}\right)}{347}} = (0.006, 0.080)$$

We are 95% confident that the proportion of patients who die from prostate cancer after having no surgery is between 0.60% and 8.0% higher than the proportion of patients who die after having surgery.

d) Since 0 is not contained in the interval, there is evidence that surgery may be effective in preventing death from prostate cancer.

12. (22.14)

Anorexia.

a) **Randomization Condition:** The women were randomly assigned to the two treatment groups.

Independent Samples Condition: The groups were assigned randomly, so the groups are not related.

Success/Failure Condition: $n\hat{p}$ (Prozac) = 35, $n\hat{q}$ (Prozac) = 14, $n\hat{p}$ (placebo) = 32, and $n\hat{q}$ (placebo) = 12 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.

b)

$$(\hat{p}_{Proz} - \hat{p}_{Plac}) \pm z^* \sqrt{\frac{\hat{p}_{Proz}\hat{q}_{Proz}}{n_{Proz}} + \frac{\hat{p}_{Plac}\hat{q}_{Plac}}{n_{Plac}}}$$

$$= \left(\tfrac{35}{49} - \tfrac{32}{44}\right) \pm 1.960\sqrt{\frac{\left(\frac{35}{49}\right)\left(\frac{14}{49}\right)}{49} + \frac{\left(\frac{32}{44}\right)\left(\frac{12}{44}\right)}{44}} = (-0.20, 0.17)$$

c) We are 95% confident that the proportion of women taking Prozac deemed healthy is between 20% lower and 17% higher than the proportion of women taking a placebo. Prozac does not appear to be effective, since 0 is in the confidence interval. There is no evidence of a difference in the effectiveness of Prozac and a placebo.

13. (22.16)

Anorexia again.

a) H_0: The proportion of women taking Prozac who are deemed healthy is the same as the proportion of women taking the placebo who are deemed healthy.

$$(p_{Prozac} = p_{Placebo} \text{ or } p_{Prozac} - p_{Placebo} = 0)$$

H_A : The proportion of women taking Prozac who are deemed healthy is greater than the proportion of women taking the placebo who are deemed healthy. $(p_{Prozac} > p_{Placebo} \text{ or } p_{Prozac} - p_{Placebo} > 0)$

b) Since 0 is in the confidence interval, fail to reject the null hypothesis. There is no evidence that Prozac is an effective treatment for anorexia.

c) If we think that Prozac is not effective and it is, we have committed a Type II error.

d) We might overlook a potentially helpful treatment.

14. (22.26)

Gender gap.

a) **Randomization Condition**: The poll was probably random, although not specifically stated.
10% Condition: 473 and 522 are both less than 10% of all voters.
Success/Failure Condition: $n\hat{p}$ (men) = 246, $n\hat{q}$ (men) = 227, $n\hat{p}$ (women) = 235, and
$n\hat{q}$ (women) = 287 are all greater than 10, so both samples are large enough.

Since the conditions are met, we can use a one-proportion z-interval to estimate the percentage of men who may vote for the candidate.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} = (0.52) \pm 1.960\sqrt{\frac{(0.52)(0.48)}{473}} = (47.5\%, 56.5\%)$$

We are 95% confident that between 47.5% and 56.5% of men may vote for the candidate.

b) Since the conditions were met in part a), we can use a one-proportion z-interval to estimate the percentage of women who may vote for the candidate.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} = (0.45) \pm 1.960\sqrt{\frac{(0.45)(0.55)}{522}} = (40.7\%, 49.3\%)$$

We are 95% confident that between 40.7% and 49.3% of women may vote for the candidate.

c) The 95% confidence intervals overlap, which might make you think that there is no evidence of a difference in the proportions of men and women who may vote for the candidate. However, if you think that, don't delay! Move on to part…

d) Most of the conditions were checked in part a). We only have one more condition to check:
Independent Samples Condition: There is no reason to believe that the samples of men and women influence each other in any way.

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

$$(\hat{p}_M - \hat{p}_W) \pm z^* \sqrt{\frac{\hat{p}_M \hat{q}_M}{n_M} + \frac{\hat{p}_W \hat{q}_W}{n_W}}$$

$$= (0.52 - 0.45) \pm 1.960 \sqrt{\frac{(0.52)(0.48)}{473} + \frac{(0.45)(0.55)}{522}} = (0.008, 0.132)$$

We are 95% confident that the proportion of men who may vote for the candidate is between 0.8% and 13.2% higher than the proportion of women who may vote for the candidate.

e) The interval does not contain zero. There is evidence that the proportion of men may vote for the candidate is greater than the proportion of women who may vote for the candidate.

f) The two-proportion method is the proper method. By attempting to use two, separate, confidence intervals, you are adding standard deviations when looking for a difference in proportions. We know from our previous studies that *variances* add when finding the standard deviation of a difference. The two-proportion method does this.

15. (22.28)

Intentional walk.

H_0: The proportion of innings in which the Giants score is the same whether or not Barry Bonds is walked. $(p_{Walk} = p_{Not} \text{ or } p_{Walk} - p_{Not} = 0)$

H_A: The proportion of innings in which the Giants score is different when Barry Bonds is walked than when he isn't walked. $(p_{Walk} > p_{Not} \text{ or } p_{Walk} - p_{Not} > 0)$

Randomization Condition: We must assume that these innings are typical of results we anticipate in future innings.

10% Condition: 79 and 298 are both less than 10% of all innings.

Independent Samples Condition: Assume that the samples are independent of each other.

Success/Failure Condition: $n\hat{p}$ (walk) = 37, $n\hat{q}$ (walk) = 42, $n\hat{p}$(not walk) = 107, and $n\hat{q}$ (not walk) = 191 are all greater than 10, so both samples are 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:

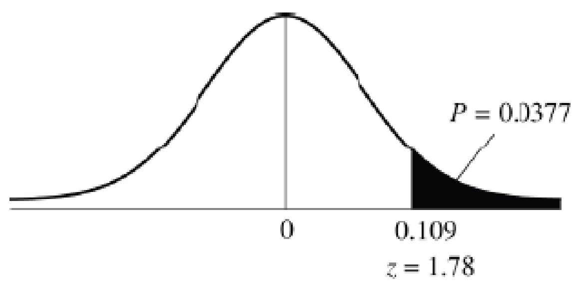
$$SE_{\text{pooled}}(\hat{p}_{Walk} - \hat{p}_{Not}) = \sqrt{\frac{\hat{p}_{\text{pooled}}\hat{q}_{\text{pooled}}}{n_{Walk}} + \frac{\hat{p}_{\text{pooled}}\hat{q}_{\text{pooled}}}{n_{Not}}} = \sqrt{\frac{\left(\frac{144}{377}\right)\left(\frac{233}{377}\right)}{79} + \frac{\left(\frac{144}{377}\right)\left(\frac{233}{377}\right)}{298}} \approx 0.06148.$$

The observed difference between the proportions is:
0.468 – 0.359 = 0.109.

Since the P-value = 0.0377 is low, we reject the null hypothesis. There is evidence that the proportion of innings in which the Giants score is higher when Bonds is walked than when he isn't walked. The opposing teams should stop walking Barry Bonds.

$$z = \frac{0.109 - 0}{0.06148}$$

$$z \approx 1.78$$



16. (23.14)

Parking.

a) **Randomization Condition**: The weekdays were not randomly selected. We will assume that the weekdays in our sample are representative of all weekdays.
10% Condition: 44 weekdays are less than 10% of all weekdays.
Nearly Normal Condition: We don't have the actual data, but since the sample of 44 weekdays is fairly large it is okay to proceed.
The weekdays in the sample had a mean revenue of \$126 and a standard deviation in revenue of \$15. The sampling distribution of the mean can be modeled by a Student's t model, with 44 – 1 = 43 degrees of freedom. We will use a one-sample t -interval with 90% confidence for the mean daily income of the parking garage. (By hand, use $t_{40}^{*} \approx 1.684$)

b) $\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 126 \pm t_{43}^{*}\left(\frac{15}{\sqrt{44}}\right) \approx (122.2, 129.8)$

c) We are 90% confident that the interval \$122.20 to \$129.80 contains the true mean daily income of the parking garage. (If you calculated the interval by hand, using $t_{40}^{*} \approx 1.684$ from the table, your interval will be (122.19, 129.81), ever so slightly wider from the interval calculated using technology. This is not a big deal.)

d) 90% of all random samples of size 44 will produce intervals that contain the true mean daily income of the parking garage.

e) Since the interval is completely below the \$130 predicted by the consultant, there is evidence that the average daily parking revenue is lower than \$130.

17. (23.30)

Fuel economy

a) H_0: The mean mileage of the cars in the fleet is 9 L/100km ($\mu = 9$)
H_a: The mean mileage of the cars in the fleet is greater than 9 L/100km ($\mu > 9$)

b) **Randomization Condition**: The 50 trips were selected randomly.
10% Condition: 50 trips are less than 10% of all trips.
Nearly Normal Condition: We don't have the actual data, so we cannot look at the distribution of the data, but the sample is large, so we can proceed.

c)
Since the conditions for inference are satisfied, we can model the sampling distribution of the mean mileage of cars with $N(9, \sigma/\sqrt{n})$. Since we do not know σ, we will estimate the SE of the sample mean with $SE = s/\sqrt{n}$, and use a Student's t model, with 50-1 = 49 degrees of freedom.

$$t = \frac{\bar{y} - \mu_0}{SE(\bar{y})} = (9.4 - 9)/(1.81/\sqrt{50}) = 1.563$$

d) $p-value = P(t_{49} > 1.563) \in (0.05, 0.10)$ using the t-table

e, f) With a p-value between 0.05 and 0.10 there is moderate to weak evidence against the null hypothesis in favour of the alternative.

18. (24.10)

CPMP and word problems.

H_0: The mean score of CPMP students is the same as the mean score of traditional students. $\left(\mu_C = \mu_T \text{ or } \mu_C - \mu_T = 0\right)$

H_A: The mean score of CPMP students is different from the mean score of traditional students. $\left(\mu_C \neq \mu_T \text{ or } \mu_C - \mu_T \neq 0\right)$

Independent Groups Assumption: Scores of students from different classes should be independent.

Randomization Condition: Although not specifically stated, classes in this experiment were probably randomly assigned to either CPMP or traditional curricula.

10% Condition: 320 and 273 are less than 10% of all students.

Nearly Normal Condition: We don't have the actual data, so we can't check the distribution of the sample. However, the samples are 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 590.05 degrees of freedom (from the approximation formula).

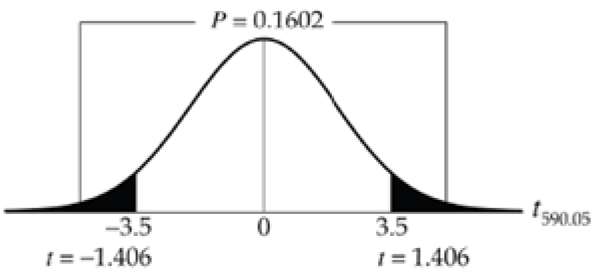
We will perform a two-sample *t*-test. The sampling distribution model has mean 0, with standard error: $SE(\bar{y}_C - \bar{y}_T) = \sqrt{\frac{32.1^2}{320} + \frac{28.5^2}{273}} \approx 2.489.$

The observed difference between the mean scores is 57.4 – 53.9 = 3.5.

$$t = \frac{(\bar{y}_c - \bar{y}_T) - (0)}{SE(\bar{y}_c - \bar{y}_T)}$$

$$t = \frac{3.5}{2.489}$$

$$t \approx 1.406$$



By hand approximate $df = \min(n_1 - 1, n_2 - 1) = 272.$

$p - value = 2P(t_{272} > 1.406) \in (0.10, 0.20)$ using the t-table

With a p-value between 0.10 and 0.20 we have weak evidence against the null hypothesis.

19. (24.18)

Handy.

a) Males: $\bar{y}_M \pm t^*_{n-1}\left(\frac{s}{\sqrt{n}}\right) = 19.39 \pm t^*_{49}\left(\frac{2.52}{\sqrt{50}}\right) \approx (18.67,\ 20.11)$

We are 95% confident that males can place between 18.67 and 20.11 pegs on average.

Females: $\bar{y}_F \pm t^*_{n-1}\left(\frac{s}{\sqrt{n}}\right) = 17.91 \pm t^*_{49}\left(\frac{3.39}{\sqrt{50}}\right) \approx (16.95, 18.87)$

We are 95% confident that females can place between 16.95 and 18.87 pegs on average.

b) It may appear to suggest that there is no difference in the mean number of pegs placed by males and females, but a two-sample *t*-interval should be constructed to assess the difference in mean number of pegs placed.

c) $(\bar{y}_M - \bar{y}_F) \pm t^*_{df}\sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}} = (19.39 - 17.91) \pm t^*_{90.49}\sqrt{\frac{2.52^2}{50} + \frac{3.39^2}{50}} \approx (0.29, 2.67)$

d) We are 95% confident that the mean number of pegs placed by males is between 0.29 and 2.67 pegs higher than the mean number of pegs placed by females.

e) The two-sample *t*-interval is the correct procedure.

f) If you attempt to use two confidence intervals to assess a difference in means, you are actually adding standard deviations. But it's the variances that add, not the standard deviations. The two-sample difference of means procedure takes this into account.

Note: In part (c), you can use the approximate df = 49 if calculating by hand.

20. (23.42)

Harder water.

The margin of error of the CI is given by:

$$ME = t^*_{n-1}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

Using $z^* = 1.96$ for t^*_{n-1} and setting each standard deviation to 150 and using equal sample sizes,

$$ME = 1.96\sqrt{\frac{150^2}{n} + \frac{150^2}{n}} = 1.96\sqrt{\frac{45000}{n}} = 50$$

and $\sqrt{n} = \frac{1.96 \times \sqrt{45000}}{50} = 8.32$. $n = 69$. This is large enough.

21. (25.14)

Windy, part III.

H_0: The mean difference between wind speeds at the two sites is zero. $(\mu_{2-4} = 0)$

H_A: The mean difference between wind speeds at the two sites is different than zero. $(\mu_{2-4} \neq 0)$

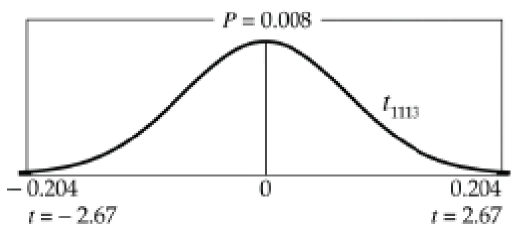
Since the conditions are satisfied (in a previous exercise), the sampling distribution of the difference can be modelled with a Student's *t*-model with 1114 – 1 = 1113 degrees of freedom, $t_{1113}\left(0, \frac{2.551}{\sqrt{1114}}\right)$.

We will use a paired *t-Test*, with $\bar{d} = 0.204$.

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

$$t = \frac{0.204 - 0}{\frac{2.551}{\sqrt{1114}}}$$

$$t \approx 2.67$$



$p - value = 2P(t_{1113} > 2.67) \in (0.005, 0.01)$ using the t-table

With a p-value between 0.005 and 0.01 there is strong evidence against the null hypothesis in favour of the alternative.

22. (25.30)

Boston startup years 2009.

a) Even if the individual times show a trend of improving speed over time, the differences may well be independent of each other. They are subject to random year-to-year fluctuations, and we may believe that these data are representative of similar races. After removing the three initial years, the remaining part of the histogram is a bit skewed and possibly bimodal, but we can probably use paired-*t* methods with caution.

b) $\bar{d} \pm t^*_{n-1}\left(\frac{s_d}{\sqrt{n}}\right) = -17.6583 \pm t^*_{29}\left(\frac{18.8731}{\sqrt{30}}\right) \approx (-24.71, -10.61)$

We are 95% confident that female wheelchair marathoners average between 10.61 and 24.71 minutes faster than male runners in races such as this.

c) Since the interval does not contain zero, we would reject the null hypothesis of no difference at a significance level of 0.05. There is strong evidence that female wheelchair marathoners finish faster than male runners, on average.