AMS 102: PRACTICE FOR THE FINAL

SOLUTIONS

Chapter 21

21.7. $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 > \mu_2$.

We assume that we have two SRSs from normal populations. $SE \approx 0.078716$. The t statistic is $t = \dfrac{5.15 - 4.33}{SE} \approx 10.4$. Conservatively, we take 19 degrees of freedom. The result is significant at any reasonable significance level.

Thus we have strong evidence that the mean remating time is longer for large spermatophores.

21.14. Let p_1 = proportion of students with college graduate parents in 2004; p_2 = same proportion in 1978.

We will determine the 99% confidence interval for $p_1 - p_2$. The counts of successes and failures and the sample size are large enough to use large-sample procedures. The sample proportions are $\hat{p}_1 = \frac{1014}{2158} \approx 0.4699$ and $\hat{p}_2 = \frac{5617}{17554} \approx 0.3200$. The standard error for a confidence interval is $SE \approx 0.01131$ and the margin of error is $2.576SE \approx 0.02912$. This gives the interval 0.1208 to 0.1790.

21.15. Let p be the proportion of all students with college graduate parents in 2004.

We will determine the 99% confidence interval for p. The counts of successes and failures and the sample size are large enough to use large-sample procedures. In the previous problem we found $\hat{p} = 0.4699$. $SE = \sqrt{0.4699(1 - 0.4699)/2158} \approx 0.01074$. The 99% confidence interval is $\hat{p} \pm 2.576SE \approx 0.4699 \pm 0.02767 = 0.4422$ to 0.4976.

21.16. Let μ_1 = mean score for the students with at least one college graduate parent and μ_2 = mean score for other students.

$H_0 : \mu_1 = \mu_2$, $H_a : \mu_1 \neq \mu_2$.

We assume that our samples are SRS. Then $SE = \sqrt{\frac{28.6^2}{10.14} + \frac{21.3^2}{410}} \approx 1.4211$ and $t = \dfrac{\bar{x}_1 - \bar{x}_2}{SE} = 15.48$. This is significant for any reasonable α. Conclusion: reject H_0, that is at 99% confidence we can conclude that mean scores are different for both groups.

To find a 95% confidence interval for the difference, we take $\bar{x}_1 - \bar{x}_2 \pm t^* SE$, where $t^* = 1.984$ (100 degrees of freedom), and obtain the interval 19.18 to 24.82.

21.17. To find the 95% confidence interval for μ_1, $SE = 28.6/\sqrt{1014}$ and the interval is $\bar{x}_1 \pm t^* SE$ ($t^* = 1.962$ with 100 degrees of freedom), that is 315.24 to 318.76 points.

21.18. $H_0 : \mu_w = \mu_m$; $H_a : \mu_w \neq \mu_m$.

We assume that we have SRS and proceed to find SE. Since we are given standard errors (as opposed to deviations) for both samples, $SE = \sqrt{0.9^2 + 1.0^2} \approx 1.3454$. Then $t = \dfrac{305 - 308}{1.3454} \approx -2.23$. This is significant at $\alpha = 5\%$. We conclude that the mean mathematics score is different.

21.19. (a) Observational study.

(b) $H_0 : p_1 = p_2$ vs $H_a : p_1 < p_2$ (p_1 and p_2 are graduation rates for the VLBW male population and the control).

$\hat{p}_1 \approx 0.7397$, $\hat{p}_2 \approx 0.8283$. Pooled rate $\hat{p} = \dfrac{179 + 193}{242 + 243} \approx 0.7832$.

$SE \approx 0.03782$ and $z = \dfrac{0.7397 - 0.8283}{SE} \approx -2.34$. P-value$= P(z < -2.34) = 0.0096$. Significant for all reasonable α.

Conclusion: VLBW graduation rates are lower.

21.20. $H_0 : \mu_1 = \mu_2$ vs $H_a : \mu_1 < \mu_2$ (μ_1 and μ_2 are mean IQs for the VLBW male population and the control).

$SE \approx 2.02786$, $t = \dfrac{87.6 - 94.7}{SE} \approx -3.50$. We have more than 100 degrees of freedom; for any reasonable α ($\alpha \geq 0.01$) the result is significant.

Conclusion: VLBW men have lower IQs.

21.21. In both tests we will use the level of significance $\alpha = 0.05$.

For proportions using drugs we test $H_0 : p_1 = p_2$ vs $H_a : p_1 \neq p_2$.

$\hat{p}_1 \approx 0.2937$, $\hat{p}_2 \approx 0.4194$, $\hat{p} \approx 0.3560$. $SE \approx 0.06057$ and $z = (0.2937 - 0.4194)/SE \approx -2.08$. For the two-sided alternative, P-value$= 0.0376$. Significant.

For comparing IQs, we test $H_0 : \mu_1 = \mu_2$ vs $H_a : \mu_1 \neq \mu_2$.

$SE \approx 1.7337$ and $t = (86.2 - 89.8)/SE \approx -2.08$. Significant for a large number of degrees of freedom.

21.24. Let p_1 be the proportion of walking flies responding to vibration, and p_2 be that proportion for resting flies.

$H_0 : p_1 = p_2$ versus $H_a : p_1 \neq p_2$.

One count is only 4; this makes the use of z procedures potentially risky. Nonetheless we proceed: $\hat{p}_1 = 54/64 \approx 0.8438$ and $\hat{p}_2 = 4/32 = 0.1250$, $\hat{p} = (54 + 4)/(64 + 32) \approx 0.6042$. $SE \approx 0.10588$, and $z = \dfrac{0.8438 - 0.1250}{SE} \approx 6.79$. The p-value is close to 0.

We conclude that we have strong evidence that walking and resting flies respond differently. Despite the low count of successes among resting flies, our conclusions are valid because the difference in proportions is very large.

21.31. A one-sample t for means (specifically, construct confidence interval for the population mean μ).

21.33. (a) We want to examine the mean of responses to the question, hence use the t-test for means.

(b) Use a matched-pairs test because we should keep each couple's responses together.

21.49. (a) We want to compare the proportions p_1 (microwaved crackers that show checking) and p_2 (control crackers that show checking). This can be done either by hypothesis testing or by constructing a confidence interval at, for instance, 95% for $p_1 - p_2$. We'll do the latter.

Since one of the counts is low (3) we will use the "plus four" method.

$\tilde{p}_1 = (3+1)/(65+2) \approx 0.0597$ and $\tilde{p}_2 = (57+1)/(65+2) \approx 0.8657$. The standard error $SE = \sqrt{\dfrac{\tilde{p}_1(1-\tilde{p}_1)}{67} + \dfrac{\tilde{p}_2(1-\tilde{p}_2)}{67}} \approx 0.05073$, so the 95% plus four confidence interval is $-0.8060 \pm 0.0994 = -0.9054$ to -0.7065.

We are 95% confident that on average microwaving reduces checking by 71% to 91%.

(b) We test $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 > \mu_2$.

We assume that the samples are SRSs from the two populations, and that the underlying distributions are close to normal. The standard error is $SE \approx \sqrt{\dfrac{33.62^2}{20} + \dfrac{21.62^2}{20}} \approx 9.0546$, and the test statistic is $t = \dfrac{x1 - x2}{SE} \approx 6.914$. This is significant at any reasonable confidence level.

(We could also compute the confidence interval: $x1 - x2 \pm t^* SE = 43.65$ to 81.55 psi. Here $t^* = 2.093$, with 19 degrees of freedom.)

We conclude that there is strong evidence that microwaved crackers withstand additional pressure.

21.51. Two of the counts are too small to perform a significance test safely (1 out of 11 blacks, 4 out of 31 whites).

Chapter 22

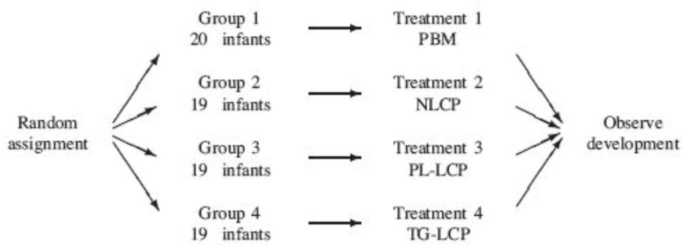
22.30. The two-way table is

	Yes	No
Phone	168	632
One-on-one	200	600
Anonymous	224	576

We test H_0: all proportions are equal, versus H_a: some proportions are different. To find the entries in the table, take (0.21)(800), (0.25)(800), and (0.28)(800). We find $X^2 = 10.619$ with 2 degrees of freedom. Hence

$p < 0.005$ and we have strong evidence that the contact method makes a difference in response.

22.31. (a) The diagram is shown below. To perform the randomization, label the infants 01 to 77, and choose pairs of random digits.



(b) The two-way table is

	PBM	NLCP	PC–LCP	TG–LCP	Total
Female	11	11	11	9	42
Male	9	9	8	10	35
Total	20	19	19	19	77

We find $X^2 = 0.568$, $df = 3$, and $P = 0.904$ and conclude that randomization worked.

22.44. (a) This is not an experiment; no treatment was assigned to the subjects.

(b) A high nonresponse rate might mean that our attempt to get a random sample was thwarted because of those who did not participate. However, this nonresponse rate is extraordinarily low.

(c) We will perform a chi-square test of the null hypothesis "there is no relationship between olive oil consumption and cancer."

All expected counts are much more than 5, so the chi-square test can be used.

With 4 degrees of freedom the chi-square statistic is $X^2 = 1.552$. If H_0 were true, the mean of X^2 would be 4. Since the value is lower than the mean, we should not reject H_0.

We conclude that high olive oil consumption is not more common among those with cancer. (In fact, computing conditional distribution of olive oil consumption confirms this.)