Stat 371-003
Supplementary problems to help you prepare for Midterm 2
Solutions

- **6.16 (pg 195)**

  (a) We first calculate the estimated standard error, $s/\sqrt{n} = 12.4/\sqrt{10} = 3.92$.

  We next find the 97.5 percentile of a t distribution with 9 degrees of freedom. From the table at the back of the book: 2.262.

  The 95% CI for the population mean is thus $13.0 \pm 2.262 \times 3.92 \approx 13.0 \pm 8.87 \approx (4.1, 21.9)$

  (b) These are the plausible values for the effect of exercise on blood HBE levels; i.e., the plausible values for the average change in blood HBE level following such a physical fitness program. To be more precise, this interval is calculated by a procedure that, in advance of gathering the data, would have a 95% chance of capturing the true effect of exercise on blood HBE levels.

- **6.20 (pg 196)**

  We first calculate the estimated standard error, $s/\sqrt{n} = 0.14/\sqrt{50} \approx 0.020$.

  We next find the 95th percentile of a t distribution with 49 degrees of freedom. From the table at the back of the book (looking at df=50): 1.676.

  The 90% confidence interval for the population mean is thus $1.20 \pm 1.676 \times 0.020 \approx 1.20 \pm 0.033 \approx (1.17, 1.23)$.

- **7.10 (pg 231)**

  We first calculate the estimated standard error, $\sqrt{(2.8)^2/489 + (2.9)^2/469} \approx 0.184$.

  We next find the 97.5 percentile of a t distribution with 950 degrees of freedom From the table at the back of the book (looking at df=1000): 1.962.

  The 95% confidence interval for the male-female difference in population averages is then $(45.8 - 40.6) \pm 1.962 \times 0.184 \approx 5.2 \pm 0.36 \approx (4.84, 5.56)$

- **7.14 (pg 232)**

  (a) We first calculate the estimated standard error of the difference between the sample averages,

  $$\sqrt{\frac{(10)^2}{10} + \frac{(8)^2}{10}} \approx 4.05$$

  We next find the 95th percentile of a t distribution with 17.2 degrees of freedom From the table at the back of the book (looking at df=17): 1.740.

  The 95% confidence interval for the male-female difference in population averages is then $(25 - 23) \pm 1.740 \times 4.05 \approx 2 \pm 7.0 \approx (-5.0, 9.0)$

(b) These are the plausible values for the effect of the antibiotic on prothrombin times; i.e, plausible values for the difference between the average prothombin times with and without the antibiotic treatment. To be more precise, this interval is calculated by a procedure that, in advance of gathering the data, would have a 95% chance of capturing the true effect of the antibiotic on prothombin times.

- **7.30 (pg 245)**

(a) We first calculate the estimated standard error of the difference between the sample averages

$$\sqrt{\frac{(2.87)^2}{60} + \frac{(3.52)^2}{50}} \approx 0.621$$

The t-statistic is $(78.42 - 80.44)/0.621 = -2.02/0.621 \approx -3.26$.

We compare this to the 97.5 percentile of a t distribution with 94.3 degrees of freedom. From the table at the back of the book (looking at df=100): 1.984.

Since $|-3.26| > 1.98$, we reject the null hypothesis.

The p-value is between 1% and 0.1%. (Using R, we find that is about 0.16%.)

(b) We conclude that males and females have different average tibia lengths.

(c) We males and females differ in their average tibia length, since the variability is large, we wouldn't be able to predict sex from tibia length very well.

(d) If the sample sizes were 6 and 5, the estimated SE for the difference between the sample averages would increase by a factor of $\sqrt{10}$, and so would be $0.621 \times \sqrt{10} \approx 1.96$.

The t-statistic would thus decrease by a factor of $\sqrt{10}$, and so would be $-3.26/\sqrt{10} \approx -1.03$.

We compare this to the 97.5 percentile of a t distribution with 7.8 df. From the table at the back of the book (looking at df=8): 2.306.

Since $|-1.03| < 2.306$, we would now *fail to reject* the null hypothesis, and would conclude that we have insufficient evidence to demonstrate that the sexes differ in tibia length.

The p-value is between 20% and 40%.

- **9.39 (pg 385)**

We first calculate the standard error of the sample average (of the differences): $1.20/sqrt15 = 0.310$.

The t-statistic is $-1/0.310 = -3.23$.

We compare this to the 97.5 percentile of a t distribution with 14 degrees of freedom. From the table at the back of the book: 2.145.

Since $|-3.23| > 2.145$, we reject the null hypothesis and conclude that catnip does have an effect on cats' "negative interactions".

The p-value is between 1% and 0.1%

- **10.10 (pg 401)**

  The observed and expected counts are as follows.

	Black	Brown	White	Total
Observed	40	59	42	141
Expected	35.25	70.50	35.25	141

  The χ^2 statistic is

$$X^2 = \frac{(40-35.25)^2}{35.25} + \frac{(59-70.50)^2}{70.50} + \frac{(42-35.25)^2}{35.25} \approx 3.81$$

  We compare this to the 95th percentile of a χ^2 distribution with 2 degrees of freedom. From Table 9: 5.99.

  Since $3.81 < 5.99$, we *fail to reject* the null hypothesis and conclude that the data conform reasonably well to the 1:2:1 ratio.

  The p-value is between 10% and 20%.

- **10.25 (pg 411)**

  The observed counts are as follows.

	Cold?		
No. relationships	No	Yes	Total
≤ 5	66	57	123
≥ 6	101	52	153
Total	167	109	276

  We get the expected count for the upper-left cell as $123 \times 167/276 \approx 74.4$. The others can be obtained by subtraction (e.g., $123 - 74.4 = 48.6$), and so we obtain:

	Cold?		
No. relationships	No	Yes	Total
≤ 5	74.4	48.6	123
≥ 6	92.6	60.4	153
Total	167	109	276

  The χ^2 statistic is:

$$X^2 = \frac{(66-74.4)^2}{74.4} + \frac{(57-48.6)^2}{48.6} + \frac{(101-92.6)^2}{92.6} + \frac{(52-60.4)^2}{60.4} \approx 4.36$$

  We compare this to the 95th percentile of a χ^2 distribution with 1 degree of freedom. From table 9: 3.84.

  Since $4.36 > 3.84$, we reject the null hypothesis and conclude that there is an association between the no. social relationships and the chance of getting a cold.

  The p-value is between 2% and 5%.