University of California, Santa Cruz
Department of Statistics
Baskin School of Engineering
Statistical Methods for the Biological,
Environmental, and Health Sciences
STAT 007

Answers to the Final

1. Hemoglobin levels in adult males are normally distributed with a mean 14.7 g/dL and a standard deviation 1.3 g/dL. The normal hemoglobin range for a men is 13.6 g/dL to 17.7 g/dL.

   (a) Find the probability that an adult male has hemoglobin levels in the normal range.

   Let x be the random variable that describes hemoglobin levels in adult males.

   $$\begin{aligned} P(13.6 < x < 17.7)\textbf{[1 pts]} &= P((13.6-14.7)/1.3 < z < (17.7-14.7)/1.3) \\ &= P(-0.8461 < z < 2.3076)\textbf{[2 pts]} \\ &= P(z < 2.3076) - P(z < -0.8461)\textbf{[1 pts]} \\ &= .9896 - 0.1977 = 0.7919\textbf{[1 pts]} \end{aligned}$$

   (b) If 9 adult males are randomly selected, find the probability that they have a mean hemoglobin level in the normal range.

   By the central limit theorem, $\bar{x}$ follows a normal distribution with mean 14.7 and standard deviation $\sigma/\sqrt{n} = 1.3/\sqrt{9}$ ***[3 pts]*** *.*

   $$\begin{aligned} P(13.6 < \bar{x} < 17.7) &= P((13.6-14.7)/(1.3/sqrt9) < z < (17.7-14.7)/(1.3/sqrt9)) \\ &= P(-2.5386 < z < 6.9230)\textbf{[1 pts]} \\ &= 0.9999 - 0.0001 = 0.9998.\textbf{[1 pts]} \end{aligned}$$

2. In a study of cell phone use and brain hemispheric dominance, an internet survey was emailed to 5000 subjects randomly selected from an online group whose focus is related to ears. 717 surveys were returned.

   (a) Construct and 90% confidence interval for the proportion of returned surveys.

   Note that $\hat{p} = 717/5000$ ***[1 pts]*** *and $z_{\alpha/2} = z_{0.1/2} = z_{0.05} = 1.645$* ***[1 pts]*** *.*
   Also $717/5000 - 1.645\sqrt{\frac{(717/5000)(1-717/5000)}{5000}} = 0.1352$ and
   $717/5000 + 1.645\sqrt{\frac{(717/5000)(1-717/5000)}{5000}} = 0.1515$. *Then, a 90% confidence interval for the proportion of returned surveys is $0.1352 < p < 0.1515$* ***[3 pts]***

(b) How many surveys must be emailed in order to be 90% confident that the sample proportion is in error by no more than 0.1 percentage points? Note that $E = 0.1\% = 0.001$**[1 pts]**, $z_{\alpha/2} = 1.645$**[1 pts]**, and $\hat{p} = 0.5$ because there is no information available regarding an estimate for p **[1 pts]** . So,

$$n = \frac{(z_{\alpha/2})^2 0.5(1-0.5)}{E^2} = \frac{1.645^2 0.25}{0.001^2} = 676506.2 \textbf{\textit{[1 pts]}}$$

So, the required sample size if 676507 ***[1 pts]*** .

3. In a study regarding body measurements, platelets counts (1000 cels / μL) were measured from 147 adult females. It it of interest to test the claim that the population of adult females has a mean platelet count less than 270. Use a 0.05 level of significance. From the study it was observed that the mean platelets counts were 240 and that the standard deviation was 130.

(a) Write the null and alternative hypothesis and compute the test statistic for this test of hypothesis.
$H_0 : \mu = 270$ *and* $H_1 : \mu < 270$ ***[2 pts]*** . *The test statistic is* $t^{stat} = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{240-270}{130/\sqrt{147}} = -2.7979$ ***[3 pts]*** .

(b) Make a decision to either reject or fail to reject the null hypothesis based on the critical value. Write your decision in simple non technical terms.
Because $t^{stat} = -2.7979$ *is smaller than the critical value* $-t_\alpha = -1.6553$ *(other accepted values are -1.660 and -1.653)* ***[1 pts]*** *from a Student t distribution with* $n - 1 = 146$ *degrees of freedom, we make the decision to reject the null hypothesis* ***[2 pts]*** . *In simple non technical terms we conclude that there is enough evidence to support the claim that the population of adult females has a mean platelet count less than 270* ***[2 pts]*** .

4. In a study involving 153 male students from UCSC, their body mass indexes and waist circumferences (cm) were recorded. From the study, it was observed that the sample linear correlation between the body mass index and waist circumference was 0.92.

(a) Use a 0.05 level of significance to test the claim that there is a linear correlation between the body mass index and waist circumference of male students at UCSC. Compute the test statistic and write the rejection region.

The test statistic is $t^{stat} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.92}{\sqrt{\frac{1-0.92^2}{153-2}}} = 28.8456$, ***[2 pts.]*** *the critical value from a Student t distribution with* $(n-2) = 151$ *degrees of freedom is* $t_{\alpha/2} = t_{0.05/2} = 1.972$ *(*1.984 *is also correct),****[1 pts.]*** *and the critical region is given by all the test statistics that are greater than* $t_{\alpha/2} = 1.972$ *or smaller than* $-t_{\alpha/2} = -1.972$ ***[2 pts.]***.

(b) The computed p-value for the above test is less than 0.0001. Make a decision to either reject or fail to reject the null hypothesis, write your decision in simple non technical terms, and mention the type of error you could be making in your decision.

Because the p-value is smaller than the level of significance 0.05, we reject the null hypothesis ***[2 pts.]*** *and conclude that there is enough evidence to conclude that there is a linear correlation between the body mass index and waist circumference of male students at UCSC* ***[2 pts.]****. The type of error that we could eb making is the type I error* ***[1 pts.]****.*

5. In a study involving 153 male students from UCSC, their body mass indexes and waist circumferences (cm) were recorded. From the study, it was observed that the sample mean of body mass index was 28.2, the sample standard deviation of the body mass index was 5.33, the sample mean of waist circumference was 100.1 cm, and the sample standard deviation of the waist circumference was 14.94 cm. Also, the sample linear correlation between the body mass index and waist circumference was 0.92.

   (a) Find the regression equation that has body mass index as the dependent variable and waist circumference as the independent variable.

   Note that $b_1 = r * s_y/s_x = 0.92 * 5.33/14.94 = 0.3282$ ***[1.5 pts.]*** *and* $b_0 = \bar{y} - b_1\bar{x} = 28.2 - 0.3282 * 100.1 = -4.6528$ ***[1.5 pts.]****. The regression equation is* $\hat{y} = -4.6528 + 0.3282x$ ***[2 pts.]****.*

   (b) Predict the value of the body mass index for a male student that has a waist circumference of 110 cm.

   we use the regression equation to predict the value of the body mass index . So, the predicted body mass index for a male student that has a waist circumference of 110 cm is $\hat{y} = -4.6528 + 0.3282 * 110 = 31.4492$. ***[5 pts.]***