Math Stat, HW5, due April 7

Book problems:

Page 436–: 12, 13 (I think there's an $n-1$ missing in the book's answer to 13).

Practice problems:

1. Let $X_1, ..., X_n$ and $Y_1, ..., Y_m$ be independent samples from normal distributions with means μ_1 and μ_2, and the same variance σ^2.

(a) Suggest a symmetric confidence interval for a linear combination of the type $a\mu_1 + b\mu_2$ where a and b are known constants.

(b) Below are weights for Norwegian salmon and Canadian salmon. Norwegians claim that their salmon is at least twice as large, and Canadians claim this is not true. Use the result in **(a)** to investigate the claims. Use a 95% confidence level.

Norwegian: 40.2, 39.5, 35.8, 43.6
Canadian: 14.0, 20.1, 21.8

5. Consider two samples $X_1, ..., X_n$ and $Y_1, ..., Y_m$ where $n = 5$. If the individual sample variances are $s_X^2 = 9$ and $s_Y^2 = 4$, and the pooled sample variance is $s_p^2 = 6$, what is m?

6. The amount of carbon particles (μg/m^3) in the air was measured at various temperatures giving the following data:

$x : 8.3, 8.9, 9.6, 10.1, 10.3, 11.2$
$Y : 6.6, 7.3, 6.8, 7.9, 8.6, 9.6$

where x is degrees above 80 (for example, $x = 9.6$ means that the temperature was 89.6) Assume that the carbon amount Y has linear regression with respect to x. Find the estimated regression line and find 95% confidence intervals for the slope b and intercept a. To save you some tedious calculations, the sums you need are:

$$S_x = 58.4, \quad S_{xx} = 573.8, \quad S_Y = 46.8, \quad S_{xY} = 460.92$$

and

$$\sum_{k=1}^{n} (Y_k - \widehat{a} - \widehat{b} x_k)^2 = 1.15$$

Note that this is in the notation from class, not from the book.

7. A politician claims to have support from more than half of the voters. He bases this on an opinion poll where 523 out of a thousand voters gave him support. Does the poll support his claim?

8. The politician from Problem 7 above claims that his support has increased between two polls taken a year apart. The first poll gave him support from 410 out of 900. Do the two polls support his claim? Use a 95% confidence level.

9. In a Swedish opinion poll in the 1990s, a newspaper claimed that a majority of Swedes wanted to join the European Union. The claim was based on a opinion poll where 505 of 1,000 had expressed support for joining.

(a) Find a 95% confidence interval for the proportion p who wanted to join.

(b) Find a 50% confidence interval for p.

(c) In order to make the newspaper's claim, how low must the confidence level be?

Turn-in problems:

1. Let the random variable T have a t distribution with r degrees of freedom. Argue that T^2 has an $F_{m,n}$ distribution and find its two parameters m and n.

2(a) One hundred Americans are asked if they will travel abroad next summer and 20 answer yes. Find a 95% confidence interval for the true proportion of Americans who intend to travel abroad.

(b) If we want to take a new poll and have a margin of error which is about $\pm 3\%$ and a confidence level of 99%, how large should the sample be? Do it in two different ways.

3. In the 2004 election, George W Bush got 57% of the vote in Louisiana. In a 2005 opinion poll, 235 out of 413 expressed support for the president and in a 2006 poll, 312 out of 600 people expressed support for the president. Compute 95% confidence intervals to examine whether support in 2006 has decreased compared to **(a)** the election result **(b)** the support at the time of the 2005 poll.

4. In class we used the linear regression model $Y = a + bx + \epsilon$ for Edwin Hubble's data set on galaxies. As we know that the true regression line must go through the origin (the galaxy at distance 0 is the one we are in), it is better to use the model $Y = bx + \epsilon$ where $\epsilon \sim N(0, \sigma^2)$. As in class let

$$S_{xx} = \sum_{k=1}^{n} x_k^2$$

and

$$S_{xY} = \sum_{k=1}^{n} x_k Y_k$$

and remember that this notation is different from the book.

(a) Show that the MLE of b is

$$\widehat{b} = \frac{S_{xY}}{S_{xx}}$$

(b) Compute the estimated regression line $y = \widehat{b}x$ for Hubble's data.

5. In each of the following cases, state the relevant null hypothesis H_0 and alternative hypothesis H_A regarding an unknown probability p.

(a) A manufacturing process needs to be adjusted if more than 5% of the items are defective.

(b) A politician running for re-election decides to fire his campaign manager if his support drops below 40%.

(c) A computer program used to simulate coin flips must be adjusted if it does not give heads and tails in the correct proportions.