Final Project
Math 243 - Summer 2016

Instructions:

- Do not write answers on this page.
- You need to show some sort of work or justification for your answers, unless otherwise stated. Writing down the formula you are using is a good way to get partial credit.
- Solutions must be written neatly. Work out the answers on scratch paper before rewriting them for your submission. You will probably need more than a page (be liberal with space!), in which case the pages must be stapled together.
- You may work together, but each student must submit their own work.
- Your project must be turned in before the final exam on *Friday, August 12*.

Grading:

- This project is *optional*. If you choose not to do it, then your third exam will be weighted as 30% of your grade.
- If you submit the project it will account for 15% of your grade, and the third exam the remaining 15%.

Part 1. Data and Statistics

You will be taking data from the Bureau of Labor Statistics, which conducts its monthly Current Population Survey using $n = 60\,000$ households. To find their data, go to

http://data.bls.gov/cgi-bin/surveymost?bls

or from the www.bls.gov homepage, navigate via the "Data Tools" drop down to "Top Picks." To retrieve a particular data set, check the box (and as many others as you like) and click "Retrieve data" at the bottom of the page. On the next page you will have some display options. Notice that you can include annual averages.

1. Choose one of the first three options in the "Productivity" category. You should be shown data by quarter, from 2006 to Q1 2016. You will use this data for Part 2 below.
   (a) Write down the information in the data header: Series ID, sector, etc. You do not need to give me the data itself, but you will need to use it in what follows.
   (b) Describe, in complete sentences, what is being displayed in the table. Notice that a *percent change* is not treated like a proportion, so this is a quantitative variable.

2. Again from the list of data sets, select the "Unemployment Rate" data set from the "Employment" category. You should be shown data by month from 2006 to 2016. You will use this for Part 3 below.
   (a) Again, copy down the information in the header.
   (b) Describe, in complete sentences, what is being displayed in the table. Notice that this represents a *proportion* of the sample.

Part 2. **Analysis of Productivity**

Use the data from Problem 1 above.

3. Draw a histogram for all 41 quarters from 2006 to 2016. Use classes of width 2%, i.e. −2 to 0, 0 to 2, 2 to 4, etc. Clearly label the axes.

4. Compute $\bar{x}$, the mean quarterly change of the $n = 8$ quarters spanning the years 2014-2015. (You do not need to show work for this.)

5. Compute s, the standard deviation for 2014-2015. (You do not need to show work for this.)

6. We will construct a confidence interval for the average change, which we might want to think of as a predictor for the expected mean change μ in 2016.
- **(a)** Should you use a z-procedure or a t-procedure? Why?
- **(b)** Find the critical value for a 96% confidence level. (You do not need to show work for this.)
- **(c)** Compute the standard error.
- **(d)** Compute the 96% confidence interval.
- **(e)** In your own words, how might we interpret the meaning of this interval?

7. Test the claim "there is no difference in mean change for Quarter 2 (pop. 1) and mean change for Quarter 3 (pop. 2)":
- **(a)** State the null hypothesis.
- **(b)** State the alternative hypothesis.
- **(c)** Compute $\bar{x}_1$, the mean change for Quarter 2 for the $n_1 = 10$ years from 2006 to 2015. (You do not need to show work for this.)
- **(d)** Compute s_1, the standard deviation for Quarter 2 for the same years. (You do not need to show work for this.)
- **(e)** Compute $\bar{x}_2$, the mean change for Quarter 3 for the same years. (You do not need to show work for this.)
- **(f)** Compute s_2, the standard deviation for Quarter 3 for the same years. (You do not need to show work for this.)
- **(g)** Compute the test statistic for the difference in means $\bar{x}_1 - \bar{x}_2$.
- **(h)** Find the best possible range for the P-value, according to table C.
- **(i)** At the $\alpha = 0.05$ significance level, should you reject or accept the null hypothesis?
- **(j)** Explain your conclusion in terms of the original claim.

Part 3. **Analysis of Unemployment**

Use the data from Problem 2 above.

8. We will construct a confidence interval for the true unemployment rate p in June 2016.

(a) Should you use a z-procedure or a t-procedure? Why?

(b) Find the critical value for a 98.44% confidence level.

(c) Using the survey size $n = 60\,000$, compute a 98.44% confidence interval for the true mean unemployment in June 2016.

9. We will test the claim "the average monthly unemployment in the first half of 2016 is still greater than the mean monthly unemployment in 2007."

(a) Compute $\hat{p}_1$, the average monthly unemployment rate in the first six months of 2016. We will think of this as "the" proportion for the first half of 2016. It is actually a mean of several sample proportions, but forget this. (You do not need to show work for this.)

(b) Compute $\hat{p}_2$, the average monthly unemployment rate for 2007. We will think of this as "the" unemployment rate for 2007. (You do not need to show work for this.)

(c) State the null hypothesis.

(d) State the alternative hypothesis.

(e) Compute the standard error using the pooled proportion $\hat{p}$, using $n_1 = n_2 = 60\,000$, the sample size in the CPS.

(f) Compute the test statistic.

(g) Compute the P-value.

(h) At the $\alpha = 0.05$ significance level, is the result significant or not?

(i) Explain your conclusion in terms of the original claim.