Math 244 Exam #1

1) Reuters Health, March 28, 2011 (I USED 1-PROPZINT)

a. Find the 95% CI for the probability of type 2 diabetes in the experimental group, to the nearest tenth of a percent.—(0.07615, 0.10175) which in percent form is (7.6%, 10.2%)

b. Find the 95% CI for the probability of type 2 diabetes in the control group, to the nearest tenth of a percent.—(0.04932, 0.07068) which in percent form is (4.9%, 7.1%)

c. Draw both CI’s on a number line

![Number Line]

d. Yes, there is a statistically significantly increased risk of developing type II diabetes according to this data. The minimum measurable difference between the control group and the experimental group is approximately 0.5%. Because the two confidence intervals do not overlap each other we can be 90.25% confident that the results shown from the data show a significant increase in the risk of developing type 2 diabetes.

e. If two statistics have non—overlapping CI’s, then we are (sure/unsure) that one of the parameters being estimated is greater than the other. We can say, therefore,
that the two parameters estimated are statistically significantly different/not necessarily statistically significantly different. However, if two statistics have overlapping CI’s, then then [sic] we are sure/unsure that one of the parameters being estimated is greater than the other. We can say, therefore, that the two parameters estimated are statistically significantly different/not necessarily statistically significantly different.

2) Relative Risk (RR)
   a. Calculate p-hat for type 2 diabetes in experimental group, to the nearest percent.
      8.89% Answer: 9%
   b. Calculate p-hat for type 2 diabetes in control group to the nearest percent. 6%
      Answer: 6%
   c. Calculate the RR of type 2 diabetes while using Lipitor, using the formula above.
      9/6=1.5 Answer: 1.5
   d. Completed Statement: Individuals with cardiovascular disease who received a full course of Lipitor are 50% more likely to develop type 2 than those not using Lipitor
   e. Margin of Error
   f. Fill in blanks: Individuals with cardiovascular disease who received a full course of Lipitor are at least 17.92% more likely (and at most 86.37% more likely) to develop type 2 diabetes than those not using Lipitor.

3) Seasonally adjusted unemployment rate
   a. Give the 95% CI for the SAUR of Delaware in August 2014, to the nearest tenth of a percent. (0.05306, 0.06694) Answer based off sample: (5.3%, 6.7%)
b. (w) In March 2012, the SAUR in DE was 6.8%. Can we say that it has changed significantly from March 2012 to August 2014? Why or why not? **Answer:** No we cannot say that it has changed significantly from March 2012 to August 2014 because the margin of error for August 2014 is approximately 0.7%, which means that the two confidence intervals overlap each other. $6.8\% \pm 0.7\% = (6.1\%, 7.5\%)$ and $(5.3\%, 6.7\%)$ overlap each other. Thus if we have two areas that overlap we are unable to confirm that has significantly changed as the correct area could fit into either areas margin of error.

c. Draw both CI’s on the same number line

![Diagram showing confidence intervals]

d. (w) At what confidence would we have had to run our original CI’s to achieve 95% confidence for both? **Answer:** Because when using two 95% CI’s the confidence level we get is 90.25% (calculated from $95\%^2$) we see an equation: $95\%^2=90.25\%$ So to calculate what percentage of confidence we would need to run our CI’s to get 95% percent confidence we do this equation: $x\%^2=95\%$ and then solve for $x$. $x=\sqrt{.95}$ $x=.974679$. We then convert $x$ to percent by multiply it by 100 and get approximately 97.47%. Which, rounded to the nearest half of a
percent we get 97.5% as the confidence level we would need to use in our CI’s calculations in order to achieve a 95% confidence for both CI’s.

e. (w) What’s the overall confidence for all 50 results to the nearest percent?

Answer: We use a similar equation for this problem as we did in part (d), instead of x being the percent to the power of, we use x as the unknown answer of the overall confidence for 50 CI’s at 95% confidence. So the equation is 95%^50=x%. Entered into the calculator we use .95^50 and get approximately 0.07694 as an answer. We then multiply this by 100 to convert it into percent resulting in 7.694%. To the nearest percent the answer is 8% for the overall confidence of all 50 results.

f. What’s the chance that we’re wrong in at least one of them? To find this we use the 1 minus the Prob(all right) and come up with the Answer: 92.31%

4) Two studies, one where the MOE makes a claim statistically significantly believable, and another where it is not:

The following study by Rasmussen Reports is one that makes a claim where it is statistically significantly believable. Even with the 3% margin of error included with the study, we still are well above 50%. (80%±3%) creates the intervals of (77%, 83%) which does show the majority, as the study claims “Most Americans are still paying more” which is statistically significantly believable.
The following study by Rasmussen Reports is one that makes a claim where it is not statistically significantly believable. Because of the 3% margin of error we could fall beneath 50% (51%±3% which makes the intervals [48%, 54%]) and thus the claim that “just over half of Americans” is not statistically significantly believable.
5) Explain why we need two different kinds of CI (1—PropZIntervals, and TIntervals):

There are several reasons we need two different kinds of CI, one reason has to do with what parameter we are trying to estimate. With a 1-propZInterval we are trying to estimate a proportion and with a T Interval we are trying to estimate an average. Another reason we need two different kinds of CI has to do with what kind of data we have to use. When we have a sample size and the amount that tested positive (or negative for whatever we were sampling for) we use the 1-propzInterval because with
that data we can calculate a proportion which we then use 1-propZint to give us the interval for the proportion. The data used for constructing T Intervals is different; with T intervals we can either use data that we collected, in list form, or data used by having the average of the sample, the sample standard deviation, and also our sample size. This is the type of data used in measuring and calculating averages and so this data is what we use to construct T intervals for the average. Depending on which data we have will depend on what we are measuring and which CI constructor we will use, T Interval or 1-PropZ Interval. Another reason we need two different kinds of CI is concerned with how we get the data we are using. For a 1-PropZInt we get the data by simply taking a sample of our population, and determining how many are (a) rather than (b) for whatever being studied. For example, taking a sample from the population of people and determining how many people voted yes for a specific ballot. For T intervals we get data by taking a sample, and finding out how much of (a), for example, each single individual of your sample does or has. For example, taking a sample of 40 people and discovering how much each of those 40 people spent for a specific month. This data is then used to find the average and the standard deviation of the collected data and can be used to find the T Interval. Another reason we need two CI’s depends on what we want to say with our data. If we want to make a claim about the average of the population we would need to use the T interval, whereas if we wanted to make a claim about the proportion of the population we would need to use the 1-Propz Interval.