Name: ______________________________

Practice Exam 4

Statistics 0800
FALL 2013
Dr. Nancy Pfenning

This is a closed book exam worth 150 points. You are allowed to bring a calculator, normal table, formula sheets, and a two-sided sheet of notes. If you want to receive partial credit for wrong answers, show your work.

1. (5 pts.) A 95% confidence interval for the mean weight of all airline passengers in 1995 was (154, 166). A 95% confidence interval for the mean weight of all airline passengers in 2002 was (172, 188). What does this suggest?

   (a) Mean weight has increased significantly since 1995.

   (b) Mean weight has increased since 1995, but the difference may not be statistically significant.

   (c) Smaller samples may have provided more convincing evidence of an increase.

2. (10 pts.) American adults were surveyed: if they could have only one child, would they have a preference for a particular sex, and if so, what would it be?

   (a) A 95% confidence interval for the population proportion of men with a preference who would want a boy is (.65, .75). We can conclude that of all men with a preference,

      i. a minority prefer boys
      ii. there is no statistical evidence of a minority or majority preferring boys
      iii. a majority prefer boys

   (b) A 95% confidence interval for the population proportion of women with a preference who would want a boy is (.42, .52). We can conclude that of all women with a preference,

      i. a minority prefer boys
      ii. there is no statistical evidence of a minority or majority preferring boys
      iii. a majority prefer boys

4. (20 pts.) For each of the following, circle the most appropriate test. Assume sample sizes are large.

   (a) Testing if bald men have a higher tendency for heart attacks than men who are not bald.
   (i) z test about a proportion (ii) z test about a mean (iii) two-sample z test
   (iv) chi-squared test

   (b) Testing if a majority of adult American Catholics are in favor of allowing women to be priests.
   (i) z test about a proportion (ii) z test about a mean (iii) two-sample z test
   (iv) chi-squared test

   (c) Testing if students' average number of hours slept in a night differs from 7 hours.
   (i) z test about a proportion (ii) z test about a mean (iii) two-sample z test
   (iv) chi-squared test

   (d) Looking at sample weights in 1995 and in 2003 to see if airline passengers overall weigh more now than they did in 1995.
   (i) z test about a proportion (ii) z test about a mean (iii) two-sample z test
   (iv) chi-squared test

5. (5 pts.) Suppose a chi-square test produced a P-value of .001. This indicates that (choose one):

   (a) the relationship between row and column variables is very strong.

   (b) we have very strong evidence of a relationship between row and column variables.

   (c) there is no statistical evidence of a relationship between the row and column variables.

6. (15 pts.) The proportion of all Americans who are left-handed is .1. Suppose I take repeated random samples of 16 Americans, recording each sample proportion of left-handers.

   (a) What should be the mean of all their sample proportions?____________

   (b) What should be the standard deviation of all their sample proportions?____________

   (c) The shape of the histogram for all of their sample proportions should be

      i. approximately normal because sample proportion is always normally distributed

      ii. approximately normal because the population is normally distributed

      iii. not necessarily normal because 16 is larger than 5

iv. not necessarily normal because we'd expect only 1 or 2 left-handers in each sample of 16.

7. (20 pts.) Values for all possible rolls of a die have mean 3.5, standard deviation 1.7. Consider the behavior of sample mean roll for repeated random samples of either 2 or 8 dice.

   (a) The mean of all sample means should be

      i. less for 8 dice than for 2 dice
      ii. approximately the same for 2 or 8 dice
      iii. more for 8 dice than for 2 dice

   (b) The standard deviation of all sample means should be

      i. less for 8 dice than for 2 dice
      ii. approximately the same for 2 or 8 dice
      iii. more for 8 dice than for 2 dice

   (c) The shape of the distribution of all sample proportions should be

      i. more normal for samples of 8 dice
      ii. the same regardless of sample size
      iii. less normal for samples of 8 dice

   (d) According to the Empirical Rule, almost all (99.7%) sample mean rolls of 8 dice should fall between ________________________and ________________________

8. (25 pts.) Suppose a survey found the mean weight of 100 airline passengers to be 180 lbs., with standard deviation 40 lbs.

   (a) Give a 95% confidence interval for the mean weight of all airline passengers.

   (b) To interpret your interval in (a), circle one of the following:

      i. 95% of sampled passengers' weights fall in this interval.
      ii. 95% of all passengers' weights fall in this interval.
      iii. The sample mean has a 95% probability of falling in this interval.
      iv. The population mean has a 95% probability of falling in this interval.
      v. We are 95% confident that sample mean falls in this interval.
      vi. We are 95% confident that population mean falls in this interval.

9. (5 pts.) Suppose the P-value for a two-sided hypothesis test is found to be .10. What would the P-value be if we changed to a one-sided test?____________

10. (10 pts.) Answer the following about role of sample size in tests of significance:

    (a) There is more of a risk of finding a statistically significant difference, even if the difference has no practical importance, when (i) the sample is small (ii) the sample is large (iii) the null hypothesis is false

    (b) There is more of a risk of failing to find a statistically significant difference, even when the actual difference exists and is important, when (i) the sample is small (ii) the sample is large (iii) the null hypothesis is true

11. (35 pts.) In a random sample of 662 American adults who expressed a preference for having a child of a particular gender, 385 preferred boys. We want to decide whether a majority of all American adults with a preference would rather have a boy:

    (a) The null hypothesis states that the population proportion preferring boys is (i) less than .50 (ii) equal to .50 (iii) greater than .50
    (iv) less than .58 (v) equal to .58 (vi) greater than .58

    (b) The alternative hypothesis states that the population proportion preferring boys is (i) less than .50 (ii) equal to .50 (iii) greater than .50
    (iv) less than .58 (v) equal to .58 (vi) greater than .58

    (c) Calculate the z statistic.

    (d) Find the approximate p-value.

    (e) Can we conclude that a majority of adult Americans with a preference would rather have a boy? (i) yes, definitely (ii) no, not at all (iii) results are borderline

Conditions for Proportions

1. (a) A population has a fixed proportion falling into the category of interest; **or** (b) A repeatable situation has a long-run proportion of occurrences of a certain outcome; **and**

2. (a) A random sample of a given size is taken from the population; **or** (b) The situation is repeated independently a given number of times; **and**

3. (a) Sample size **or** (b) number of repetitions is large enough that there are at least 5 occurring both in and out of the category of interest.

Rules for Sample Proportions

If numerous samples or repetitions of the same size are taken,

1. (center) The mean of all sample proportions will be true proportion for the population.

2. (spread) Standard deviation of proportions = standard error =

$$\sqrt{\frac{\text{population proportion}(1 - \text{population proportion})}{\text{sample size}}}$$

3. (shape) The frequency curve made from proportions from the various samples will be approximately bell-shaped/normal.

Probability Intervals based on Empirical Rule

1. Probability is 68% that sample proportion falls within 1 standard error of population proportion.

2. Probability is 95% that sample proportion falls within 2 standard errors of population proportion.

3. Probability is 99.7% that sample proportion falls within 3 standard errors of population proportion.

Confidence Interval for Proportion

A 95% confidence interval for unknown population proportion is sample proportion plus or minus 2 standard errors, which is approximately

$$\text{sample proportion} \pm 2\sqrt{\frac{\text{sample proportion}(1 - \text{sample proportion})}{\text{sample size}}}$$

Conditions for Means

1. Samples must be random; **and**
2. Either (a) the population of measurements is bell-shaped; then any sample size is OK; **or** (b) if the population is not bell-shaped, then the sample size must be large enough

Rules for Sample Means

If numerous samples or repetitions of the same size are taken from a population of values for a measurement variable, and sample means are found,

1. (center) The mean of all sample means will equal population mean
2. (spread) Standard deviation of sample means = standard error =

$$\frac{\text{population standard deviation}}{\sqrt{\text{sample size}}}$$

3. (shape) The frequency curve made from means from the various samples will be approximately bell-shaped/normal.

Probability Intervals based on Empirical Rule

1. Probability is 68% that sample mean falls within 1 standard error of population mean.
2. Probability is 95% that sample mean falls within 2 standard errors of population mean.
3. Probability is 99.7% that sample mean falls within 3 standard errors of population mean.

Confidence Interval for Mean

A 95% confidence interval for unknown population mean is sample mean plus or minus 2 standard errors, which is approximately

$$\text{sample mean} \pm 2\frac{\text{sample standard deviation}}{\sqrt{\text{sample size}}}$$

Confidence Interval for Difference in Means

1. Collect large samples of observations independently for two groups. Compute sample mean and standard deviation for each.
2. Compute each standard error of the mean $SEM = \frac{\text{sample standard deviation}}{\sqrt{\text{sample size}}}$
3. Compute SE of the difference = $\sqrt{(SEM_1)^2 + (SEM_2)^2}$

4. A 95% confidence interval for difference in two population means is

$$\text{difference in sample means} \pm 2\sqrt{(SEM_1)^2 + (SEM_2)^2}$$

Steps for Testing Hypotheses about Proportions

1. Determine null and alternative hypotheses:

   null hypothesis: population proportion = ____________

   alternative hypothesis: population proportion $\left\{ \begin{array}{c} > \\ < \\ \neq \end{array} \right\}$ ____________

2. Collect and summarize data, including sample size, sample proportion,

$$\text{standard error} = \sqrt{\frac{\text{population proportion}(1 - \text{population proportion})}{\text{sample size}}}$$

$$\text{test statistic} \quad z = \frac{\text{sample proportion} - \text{population proportion}}{\text{standard error}}$$

   [Where population proportion is as stated in the null hypothesis.]

3. Determine p-value = unlikelihood of the test statistic, assuming the null hypothesis is true

4. Make a decision: If the p-value is "small", reject the null hypothesis and conclude the alternative is true. (Results are **statistically significant**.) If the p-value is not so small, we cannot reject the null hypothesis.

Steps for Testing Hypotheses about Means

1. Determine null and alternative hypotheses:
   null hypothesis: population mean = ____________

   alternative hypothesis: population mean $\left\{ \begin{array}{c} > \\ < \\ \neq \end{array} \right\}$ ____________

2. Collect and summarize data, including the test statistic

$$z = \frac{\text{sample mean} - \text{population mean}}{\text{standard error}} = \frac{\text{sample mean} - \text{population mean}}{\frac{\text{standard deviation}}{\sqrt{\text{sample size}}}}$$

   [Where population mean is as stated in the null hypothesis.]

3. Determine p-value = unlikelihood of the test statistic, assuming the null hypothesis is true

4. Make a decision: If the p-value is "small", reject the null hypothesis and conclude the alternative is true. (Results are **statistically significant**.) If the p-value is not so small, we cannot reject the null hypothesis.