Math 263

Exam I PRACTICE

February 5th 2015

The University of Arizona

Name: _____

Answers without adequate justification will not receive full credit, including multiple choice. Include units with your answer when appropriate, and box all answers unless an answer line is provided. By signing below I am agreeing to abide by the University of Arizona academic integrity policies and that all work done on this test is my own.

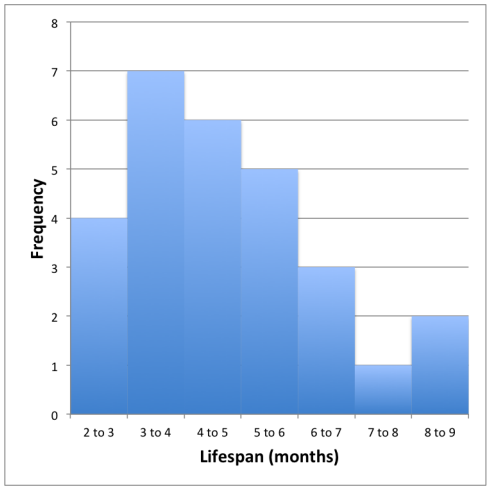
Signature: _____

Tips for Success:

- Look through the entire test before starting to prioritize questions.
- If you get stuck on a question, move on and come back to it later.
- Do a quick reality check after each question: does my answer make sense? Did I include units? Did I show all my work?
- Read over the entire test at the end to make sure you didn't miss anything.
- For each question: take a deep breath, think slowly and deliberately at first, then work quickly once you see what to do.

Special note for 263 exams: if you use a calculator to compute statistical quantities such as a 5 number summary, linear regression, normal proportion etc, please indicate which function you used i.e. 1-Var Stats, LinReg(ax+b), normalcdf().

1. The life span of 28 ant colonies is displayed in the histogram below:



   - How many ant colonies survived less than 3 months?
   - Approximately what was the median lifespan of the ant colonies?
   - Based on the histogram, which do you expect to be higher, the mean or the median?

   Suppose now that the ant colony life span is normally distributed with a mean of 5 months and a standard deviation of 1 month. (Is this reasonable given the empirical distribution?)

   - What proportion of ant colonies survive between 3 and 6 months? Find the z-scores, then draw a sketch on a standard normal before finding the value.
   - What is the lifespan of an ant colony at the 80th percentile?

2. (Mult. Choice) When using a histogram to display categorical values, you should make sure the categories are in alphabetical order.

   - True - Histograms are not useful if the categories are not in order.
   - True - Histograms can be used on any type of data.
   - False - You cannot use histograms to display categorical data.
   - False - The categories can be in any order when displaying categorical data.

3. For describing the distribution of a set of data, when is the five-number summary preferred over the mean and standard deviation?

   (a) When the distribution is reasonably symmetric.

   (b) When the distribution has little skewness and there are no outliers.

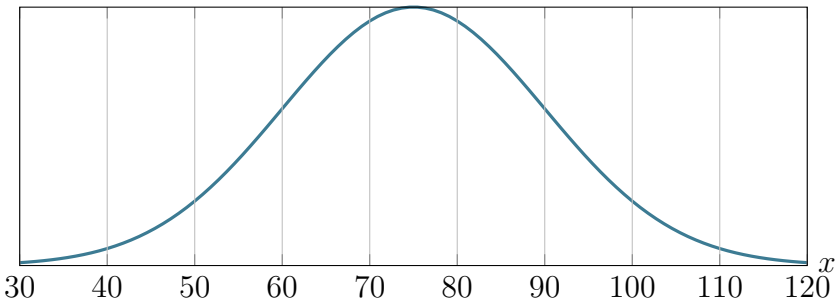
   (c) If the data are provided in increasing order of magnitude.

   (d) If the data exhibit skewness and there are strong outliers.

   (e) Never, because the mean and standard deviation are always more reliable.

4. Suppose the duration of a particular hand surgery is normally distributed with a mean of 75 minutes and a standard deviation of 15 minutes.

   (a) What proportion of surgeries last longer than 90 minutes? Sketch the region on the normal distribution below corresponding to the proportion of surgeries lasting longer than 90 minutes. Compare the answer you get using the 68-95-99.7 rule and the z-table.



   (b) Sketch the region corresponding to surgeries lasting between 50 and 60 minutes.

   (c) Find the z-scores for surgeries of 50 and 60 minutes

   (d) What percentage of surgeries last between 50 and 60 minutes?

   (e) Suppose you have a surgery that lasts 100 minutes. What percentage of people have a *longer* surgery than you?

   (f) Suppose that the doctor tells you 'that was easy, 99% of these surgeries last longer than yours did'. How long was yours?

5. True or false:

   (a) When analyzing correlation, the explanatory variable and response variable can be interchanged and the correlation coefficient will remain the same.

   (b) When analyzing regression, the explanatory variable and response variable can be interchanged and the least-squares regression line will remain the same.

6. The government wants to investigate average weekly household spending on tobacco products, T, and average household spending on alcohol, A for a group of domestic offenders. The data is below:

Offender ID	Alcohol ($)	Tobacco ($)
1	0	30
2	15	25
3	30	45
4	40	85
5	85	75
6	95	70
7	120	70
8	130	65
9	150	20

   (a) Use average household spending on alcohol as the explanatory variable and find the equation of the least-squares regression line.

   (b) What does your regression line predict the average weekly spending on tobacco products will be for someone who spends an average of $100 per week on alcohol?

   (c) Is it reasonable to conclude from this data that there is a causal relationship between the data?

   (d) Comment on any / all of the following. Outliers, influential data, strength of correlation, etc etc etc.

7. The thorax lengths of a population of fruit flies follow a normal distribution with mean $\mu = 0.800$mm and a standard deviation of $\sigma = 0.078$ mm.

   (a) What percent of flies have thorax lengths smaller than 1mm?

   (b) What percent of flies have thorax length greater than the mean?

   (c) What thorax lengths make up the bottom 15% of all thorax lengths?

   (d) Suppose we collect 600 fruit flies. About how many of these do we expect to have thorax lengths greater than 1mm?

8. The following 20 test scores were recorded on an exam:

   21, 29, 48, 55, 67, 70, 71, 75, 79, 81, 81, 86, 89, 90, 90, 92, 97, 98, 99, 100

   (a) Find the mean and the 5-number summary for the data and draw a box-and-whiskers plot, including the mean

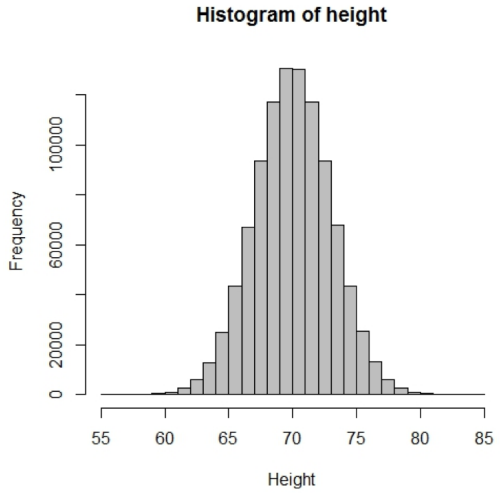
   (b) Based on your box-and-whiskers plot, do you think the data is skewed? Why?

   (c) Use the 1.5IQR rule to identify possible outliers.

   (d) Find the standard deviation for the data.

   (e) Compute *and interpret* the z-score for the following exam scores: 21, 55, 79, 89, 99

9. The following histogram shows the distribution of heights among US men:



(a) Describe the form of the distribution in height:

- One peak and is approximately symmetric.
- Two peaks and is skewed to the right
- Two peaks and is skewed to the left.
- One peak and is skewed to the right.
- One peak and is skewed to the left.

(b) Where is the mean location in relationship to the median?