

### UTAustinX: UT.7.20x Foundations of Data Analysis - Part 2



▶ Important Pre-Course Survey

- Contact Us
- ▶ How To Navigate the Course
- Discussion Board
- Office Hours
- Week 0: Introduction to Data (Optional Review)
- ▶ Week 1: Sampling
- Week 2: **Hypothesis** Testing (One **Group Means**)

#### Readings

Reading Check due May 03, 2016 at 17:00 UTC

#### **Lecture Videos**

Comprehension Check due May 03, 2016 at 17:00 UTC

#### **R Tutorial Videos**

#### Pre-Lab

Pre-Lab due May 03, 2016 at 17:00 UTC

Lab

Week 2: Hypothesis Testing (One Group Means) > Problem Set > Question 1

■ Bookmark

# Question 1

## How much money do professional bull riders earn by participating in an event?

- 1. Create a new variable that equals the "average earnings per event" in the 2012 season for each bull rider in the dataset. Call this new variable "earnings per"
- 2. Make a histogram of your "earnings per event" variable.
- 3. Use this data to answer the following questions.

Use the "BullRiders.csv" dataset to answer the following questions. Instructions for installing "BullRiders.csv" can be found under the **Examine the Data** unit in this week's **Pre-Lab** section.

(1/1 point)

1a. Have we met the assumptions for being able to calculate a 95% confidence interval to estimate the true mean earnings-per-event for a professional bull rider (using t)? Use the histogram to help answer this question.

- Yes, the distribution of "earnings\_per" looks realively normal
- No, the distribution of "earnings\_per" is positively skewed, with an outlier
- No, the distribution of "earnings\_per" is negatively skewed, with an outlier
- No, the distribution of "earnings\_per" looks realively normal, but contains outliers

Lab due May 03, 2016 at 17:00 UTC

#### **Problem Set**

Problem Set due May 03, 2016 at 17:00 UT You have used 1 of 1 submissions

When a variable is highly skewed, we can transform the data into a shape that allows us to conduct our analysis.

- 1. Create a new variable that is the log of your "earnings\_per" variable.
- 2. Here is the code to make a log transformation of a variable:

bull\$newvariable <- log(bull\$originalvariable)</pre>

3. Now use this new variable to answer the following questions.

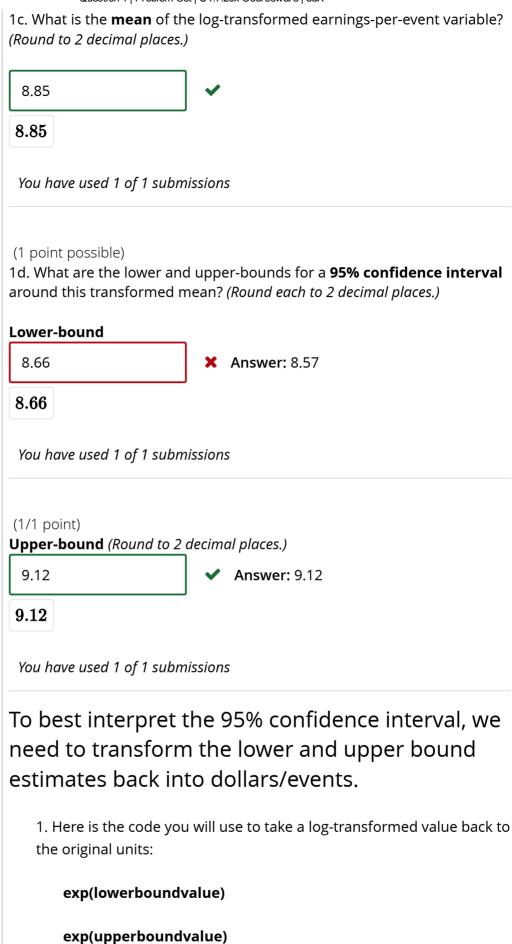
(1/1 point)

1b. Make a **histogram** of this log-transformed variable. Notice how the distribution shape has changed. Can we reliably calculate a 95% confidence interval for the mean of this transformed variable?

- Yes, the distributuon of the log-transformed variable looks relatively normal (some slight positive skew).
- No, the distributuon of the log-transformed variable still shows severe positive skew.
- No, the distributuon of the log-transformed variable is negatively skewed.
- No, the distributuon of the log-transformed variable has many outliers.

You have used 1 of 1 submissions

(1/1 point)



2. Run this code on the **unrounded original values**. Then answer the

questions that follow.

© All Rights Reserved



© edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.



















