

### 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**)
- ▶ Week 3: Hypothesis Testing (Two Group Means)
- Week 4: **Hypothesis Testing** (Categorical Data)

Readings

Week 4: Hypothesis Testing (Categorical Data) > Lab > Analyze the Data

Analyze the Data

■ Bookmark

# **Primary Research Questions**

- 1. Are each of the four musical genres equally represented on Austin City Limits?
- 2. Are some genres more likely to draw a large (100K+) Twitter following than others?

## **Analysis**

Let's break this question down into the different descriptive statistics that you will need to construct your answer. Be sure that your R output includes all of the following components.

#### **Goodness of Fit Test:**

- 1. Create a table to show the counts of each genre.
- 2. Create a vector of expected proportions.
- 3. Check the expected counts assumption.
- 4. Run a chi square test.
- 5. Interpret the chi square statistic and p-value.

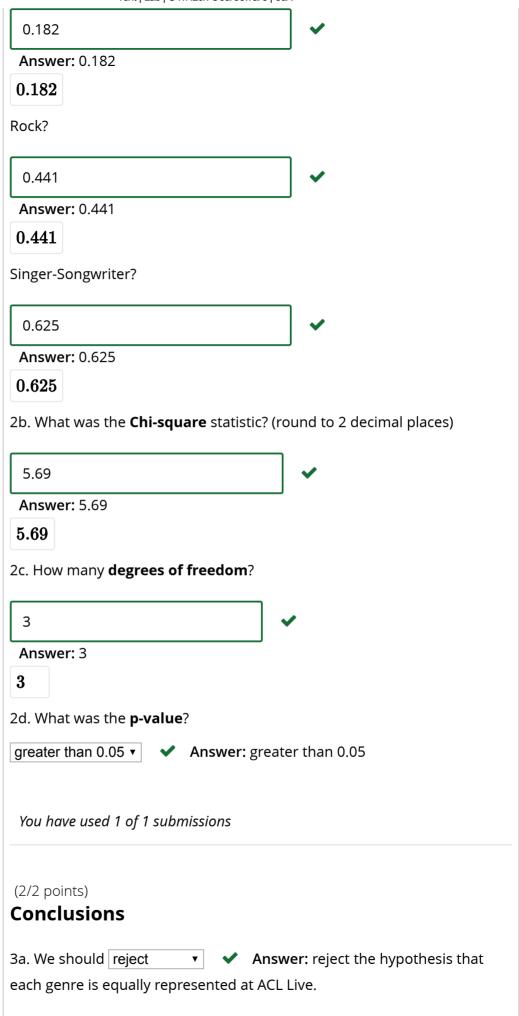
### **Test of Independence:**

- 1. Create a two-way table for genre and Twitter following.
- 2. Check the expected counts assumption.
- 3. Run a chi square test.
- 4. Interpret the chi square statistic and p-value.

(4/4 points)

## **Goodness of Fit Test**

Reading Check due 1a. What was the **expected count** of artists for each genre? May 03, 2016 at 17:00 UTC 29 **Lecture Videos** Comprehension Check Answer: 29 due May 03, 2016 at 17:00 UTC 29 **R Tutorial Videos** 1b. What was the **Chi-square** statistic? (round to 2 decimal places) Pre-Lab Pre-Lab due May 03, 2016 at 17:00 UTC 70.41 Lab **Answer:** 70.41 Lab due May 03, 2016 at 17:00 UTC 70.41 **Problem Set** 1c. How many **degrees of freedom**? Problem Set due May 03, 2016 at 17:00 UT @ 3 Answer: 3 3 1d. What was the **p-value**? less than 0.05 Answer: less than 0.05 You have used 1 of 1 submissions (7/7 points) **Test of Independence** 2a. Using the data from your two-way table, compute the proportion of artists in each genre with 100K+ Twitter followers. (Round to 3 decimal places). Country? 0.353 **Answer:** 0.353 0.353Jazz?



that genre is independent of Twitter followers. You have used 1 of 1 submissions

© 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.

















