Dataset Choices: <http://archive.ics.uci.edu/ml/datasets.html?format=&task=cla&att=&area=&numAtt=&numIns=less100&type=&sort=nameUp&view=table>

Stat 3080 basketball dataset

Project Topic: [此处请插入耀眼的文本]

Group: Lingfeng Cao, Ni Zhuang, Shaoran Sun, Xiao Dong, Yilin Liu, Zhehao Li

**Project Abstract**

Abstract Structure:

Paragraph #1: Who are group members? Brief description of the data and the problem we want to address.

Paragraph #2: Why is it interesting to you? How is it relevant to our class?

Paragraph #3. Literature review: What have been said and done about the problem (data)?

Paragraph#4. What methods are you planning to use?

Sample paragraphs:

For our final project, we are interested in examining the statistical relationships/whatever related to the \_\_\_\_\_\_\_\_\_\_\_ (like NBA player stats, Pornhub user stats, certain behaviors, etc.), and we will be using the following datasets: \_\_\_\_\_\_\_\_\_.txt, \_\_\_\_\_\_\_\_\_.csv, and \_\_\_\_\_\_\_\_\_.avi.

Explain our motivation for selecting our project theme and datasets

After a brief examination of the attributes of the three data sets, we plan to use nonparametric methods to continue our study, because the normality assumption is not met and the sample size is not large enough (maybe).

Here are our proposed questions based on the data sets and our motivations and methodology for answering those questions:

1. One question that we can solve using the Materials from Chapter 0, 1, and 2

Explanation / Motivation:

Method / Statistical Tests to use:

2. One question that we can solve using the Materials from Chapter 3

Explanation / Motivation:

Method / Statistical Tests to use:

3. The first of the three question that we can solve using the Materials from Chapter 4, 5, or 8

Explanation / Motivation:

Method / Statistical Tests to use:

4. The second of the three question that we can solve using the Materials from Chapter 4, 5, or 8

Explanation / Motivation:

Method / Statistical Tests to use:

5. The third of the three question that we can solve using the Materials from Chapter 4, 5, or 8

Explanation / Motivation:

Method / Statistical Tests to use:

Tests we have learned:

**From Chapter 0,1,2:**

One sample methods:

Binomial Test

* Test if the mean is different from the median
* Tets if the observations in one treatment are different from each others

Two sample methods:

Two-sample permutation Test

* Test if the distribution from one treatment is different from that of the other treatment

Wilcoxon Rank-Sum Test

* Test if the distribution from one treatment is different from that of the other treatment
* Use Ranking methods

Mann-Whitney Test for Confidence Interval

* Two sample permutation test 🡪 get confidence interval from this test

Hodges-Lehmann Estimate

* Median of all the pairwise differences of the form Xi – Yj is called the Hodges- Lehmann Estimate of the delta
* To get confidence interval from this estimate

Scoring Systems

* Other ranking tests for two-sample permutation test

Sigel-Tukey and Ansari-Bradley Test

* Test for deviances
* However, assume the location parameters are the same

Test on Deviances

* RMD Test to test for deviances (Recommended over Siegel-Tukey or A-B Test)
* Location parameters are not the same

Kolmogorov- Smirnov Test

* This K-S test is the nonparametric analog to the two-sample t-test with unequal variances

From Chapter 3:

K-sample permutation test

* Test if for multiple treatments, they have equal distributions

Kruskal-Wallis Test

* Test if for multiple treatments, they have equal distributions
* Ranking based

Pairwise Test

* Determine where is the pairwise difference when rejecting the null of K-W
* Using two-sample permutation test or WRS

Bonferroni adjusted p-value

* Adjusted p-value for pairwise differences

Jonckheere-Terpstra Test

* Test for ordered multiple treatments alternatives

From Chapter 4,5,8:

Chapter 4: (Not really familiar with, you guys can go ahead filling this up)

Paired-comparison Permutation Test

Test for the Median of a Symmetric Population

Wilcoxon Signed-Rank Test

Randomized Complete Block Design (RCBD) Permutation Test

Friedman’s Test for RCBD

Cochran’s Q and Kendall’s W

Pages’ Test for Ordered Alternatives

Chapter 5:

Pearson’s Correlation

Permutation Test for ρ or β1

* Permutation test for testing correlation or β1in the regression line

Spearman Rank Correlation

* Standard Pearson’s correlation applied to the ranks of X and ranks of Y

Kendall’s τ

* Test if X and Y are concordant or discordant
  + Positive association
  + Negative association
  + No association

Contingency Tables (Chi-Square Test)

* Table methodology to test if X and Y are associated

Fisher’s Exact Test for a 2\*2 Contingency Table

* Specific situation for a 2\*2 Contingency Table
* Still test if X and Y are associated

Chapter 8:

???