Math 180 Fall 2014 Midterm 2

*This take-home exam is due in class on Tuesday, December2. Each question is worth 50 points.*

Problem 1

Answer the questions in the material previously distributed containing the regression analysis in *Bazemore v. Friday*.

Problem 2

A drug store chain operating in the southeast was sued by the EEOC and a female pharmacist. They claimed that there was discrimination in the promotion of women from pharmacist to chief pharmacist. Here are the data for a certain cohort of employees showing the time in months from hire to promotion for men and women:

Men: 5, 7, 12, 14, 14, 14, 18, 21, 22, 23, 24, 25, 34, 34, 37, 47, 49, 64, 67, 69, 125, 192, 483

Women: 229, 453

One way to analyze data such as this is to rank order all the numbers (25 in all, including those of the two women) from lowest to highest and to compare the sums of the ranks of the women with those of the men. One can then apply the Wilcoxon rank sum test, for which one generally needs tables, but the problem is simple here because there are only two women.

1. What were the ranks of the women?
2. If the ranks were drawn at random, how many different pairs of ranks would have been possible for the women?
3. With this, what is the probability that the sum of the women’s ranks would be as high as actually seen?

(Adapted from Finkelstein and Levin. The case was *Capaci v. Katz & Besthoff, Inc.*)

Problem 3

The accompanying page from an issue of Evidence-Based Mental Health gives a brief description and commentary on an important study published in JAMA (Journal of the American Medical Association). Explain how the reported probabilities were calculated. If there is insufficient information in the summary, make any reasonable assumptions necessary to permit the computations but be absolutely explicit about what you are assuming. Compute the NNTs. Again, you may make any reasonable assumptions necessary to permit the calculations, but be absolutely explicit about what you are assuming. (Fluoxetine, also known by the trade names Prozac, Sarafem, Ladose, and Fontex, among others, is an antidepressant.)



Problem 4

The European roulette wheel has only 37 pockets, 18 red, 18 black, and one green 0. As everywhere, red and black each pay out 2 for 1. The European version of Atlantic City’s “surrender” is “*en prison*”. If you have bet on anything but 0 and the ball lands on zero, then your bet stays on the table through the next spin (*en prison*). If the ball lands on your number on the next spin, then your bet is returned to you (but nothing more). Able plays in Atlantic City with the surrender rule, Pierre plays in Paris with the *en prison* rule. If each bets $1 (or equivalent) on red 100 times, what is the probability that Pierre loses less than Able?

Problem 5

The distribution of eligible jurors in a certain district is 1/4 African-American, 1/3 Latino, and 5/12 Caucasian. By law, venires (the pools which must show up at the courthouse and from which the jurors are drawn) must be selected at random from the population of potentially eligible jurors. However, both defense and the prosecution may challenge potential jurors for cause (e.g., being a relative of the defendant or married to a police officer) and generally also get a limited number of “peremptory” challenges where no cause need be shown. In 20 successive criminal cases tried to 12- person juries this was the distribution of jurors.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group/Trial | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Caucasian | 7 | 4 | 7 | 5 | 9 | 5 | 6 | 3 | 7 | 6 | 6 | 7 | 8 | 5 | 7 | 4 | 7 | 6 | 9 | 5 |
| Latino | 2 | 4 | 2 | 4 | 3 | 4 | 1 | 5 | 4 | 3 | 5 | 2 | 4 | 3 | 2 | 4 | 4 | 5 | 2 | 4 |
| African-American | 3 | 4 | 3 | 3 | 0 | 3 | 5 | 4 | 1 | 3 | 1 | 3 | 0 | 4 | 3 | 4 | 1 | 1 | 1 | 3 |

In the next 20 trials this was the distribution.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group/Trial | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
| Caucasian | 5 | 4 | 6 | 5 | 7 | 5 | 4 | 3 | 5 | 6 | 4 | 7 | 6 | 5 | 5 | 4 | 5 | 6 | 7 | 5 |
| Latino | 4 | 4 | 3 | 4 | 5 | 4 | 3 | 5 | 4 | 3 | 5 | 2 | 4 | 3 | 3 | 4 | 4 | 5 | 3 | 4 |
| African-American | 3 | 4 | 3 | 3 | 0 | 3 | 5 | 4 | 3 | 3 | 3 | 3 | 2 | 4 | 4 | 4 | 3 | 1 | 2 | 3 |

After trial #20 a Caucasian convicted in trial #5 appealed. After trial #40 an African-American convicted in Trial #27 appealed. Setting aside any other possible grounds, what could you say on the basis of these figures alone? What would be the argument for the prosecution (which wants to preserve the convictions)? What would be the argument for the defense? Are there any ways of looking at these figures which would be better for one party than the other, for example, by consolidating some of the categories? Is there evidence for a change in the method of jury selection after Trial #20?

Problem 6

Here are some figures from the early part of the last century showing how often the teams in a particular English league scored a given umber of goals per match. (From Moroney, *Facts from Figures.)*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Number of goals | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Number of times teams scored that number of goals in a match | 95 | 158 | 108 | 63 | 40 | 9 | 5 | 2 |

If the numbers of goals followed a Poisson distribution, what would be the expected numbers be?

Use the chi-square distribution to see if the Poisson distribution is a good fit for this distribution of goals: How many degrees of freedom are there? What is the value of chi-square? What is the value of *p*? What factors might cause a deviation of the observed numbers of goals from a chi-square distribution? What are the mean and the variance of this distribution. (The mean and variance of a true Poisson distribution are identical.)