Math 121 Test #1 Solutions Fall 2006

Five-number summary: $(52, 67, 74, 85, 91)$, $\overline{x} = 74$.

1. (15 pts) $(2, 8, 11, 12, 15)$, $\overline{x} = 10.0$.

   (a) Because the values in the first picture become less likely as you move away from the center in either direction, the direction of extreme is *two-sided.*

   (b) The values at least as extreme as 2 are 2 and 1 on the left and 6 and 7 on the right. This is the rejection region.

   (c) $\alpha = \frac{6}{20}$ because 6 of the 20 vouchers in Box 1 are in the rejection region.

   (d) The acceptance region consists of the values 3, 4, and 5. Therefore, $\beta = \frac{4}{20}$ because 4 of the 20 vouchers in Box 2 are in the acceptance region.

   (e) The p-value is computed in pretty much the same way that α was. The values that are at least as extreme as 3 are 3, 2, 1 and 5, 6, 7. In Box 1, that accounts for 14 out of 20 vouchers, so p-value $= \frac{14}{20}$.

2. (12 pts) $(2, 10, 12, 12, 12)$, $\overline{x} = 10.4$.

   (a) The null hypothesis should be H_0: The new type of airbag is not better than the old type. The alternative hypothesis should be H_1: The new type of airbag is better than the old type.

   (b) In general, a Type I error is to reject H_0 when it is true. In this situation, that means to conclude that the new type of airbag is better than the old type when in fact it is not better.

   (c) The phrase "statistically significant" always means that the null hypothesis was rejected, so the alternative hypothesis was supported.

3. (10 pts) $(5, 5, 5, 10, 10)$, $\overline{x} = 6.9$.

   (a) They are using a *cluster sample.* Only certain classes were selected and everyone in those classes was in the sample.

   (b) The three numbers turn out to be 432, 232, and 304. I did not count off if you did not use a seed of 87, provided your numbers appeared to be random in the range 1 to 500.

4. (18 pts) $(4, 10.75, 12, 15, 17)$, $\overline{x} = 12.2$.

   (a) This was an observational study. Surely the researchers would not make people obese on purpose or give them diabetes on purpose just for the sake of the experiment.

   (b) One explanatory variable was the person's obesity, as measured by the body-mass index (BMI). The other explanatory variable was whether the person had diabetes.

(c) The response variable was whether the person suffered "acute organ failure or death from organ failure" during the specified time period.

(d) The best answer is that there was no control group. We talked about control groups only in the context of experiments. However, there were several groups that were being compared: obese people with diabetes, obese people without diabetes, normal-weight people with diabetes, and normal-weight people without diabetes. One might think of the normal-weight people without diabetes as like a control group, so I counted that as correct.

(e) The article states (2nd paragraph) that "obese people who did not have diabetes had the same risk of death or organ failure as normal-weight people without the disease." Thus, the null hypothesis was supported.

(f) This is a *confounding* variable. Confounding variables are explanatory variables, so I gave partial credit for that answer, but they are explanatory variables that were not taken into account in the study.

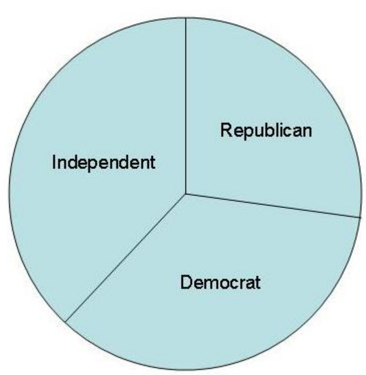
5. (10 pts) $(3, 7, 7, 10, 10)$, $\overline{x} = 7.6$.
It was sufficient either to name the types or to describe them. One type of bias is *selection* bias. They surveyed only people who belonged to their organization, so their responses would be pretty predictable. A second type of bias is *response* bias. The wording of the question is clearly leading the subject to the desired answer. The third type of bias is *nonresponse* bias. Because the survey was mailed out, those with neutral opinions are much less likely to return the survey.

The design of the survey did not include experimenter bias. The responses were simply "Agree," "Disagree," or "Undecided." No room was left for the experimenters to interpret the responses.

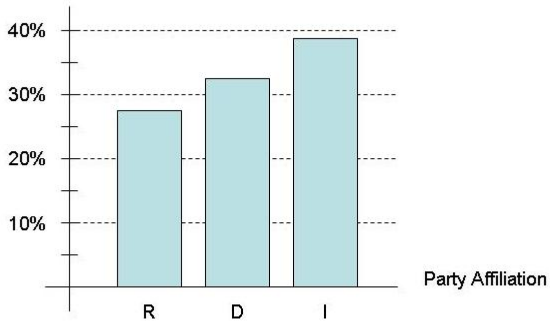
6. (13 pts) $(2, 9, 9, 13, 13)$, $\overline{x} = 9.8$.

(a) For the groups to be represented properly in the sample, the sample should be 33% Democrats, 28% Republicans, and 39% Independents. Since there are 500 people in the sample, that would be $500 \times 0.33 = 165$ Democrats, $500 \times 0.28 = 140$ Republicans, and $500 \times 0.39 = 195$ Independents.

(b) Pie chart:



Bar graph:



(c) If you drew a pie chart, then it was designed to facilitate the comparison of each category to the whole. I accepted answers like "to show the percentage of each group," but that is not really a very good answer. The numbers themselves already show that more clearly than the pie chart does.

If you drew a bar graph, then it was designed to facilitate comparisons between groups.

Neither graph is designed to show the distribution.

7. (12 pts) $(3, 8, 11, 12, 12)$, $\overline{x} = 10.0$.

   (a) The control group gives the researchers something to compare the treatment group to. If there is a difference, then the difference is likely attributable to the new drug.

   (b) "Double-blind" means that neither the subjects nor those making the observations know which subjects are in the treatment group and which are in the control group.

   (c) Double-blindedness is meant to eliminate response bias and experimenter bias.

8. (10 pts) $(0, 5, 7.5, 10, 10)$, $\overline{x} = 7.0$.

   (a) Quantitative discrete. The values are numerical, but the only possible values are whole numbers.

   (b) Qualitative. This is not numerical.

   (c) Quantitative continuous. It is numerical, but there are no values that can be ruled out beforehand, within a reasonable interval. All values are possible.

   (d) Quantitative discrete. We are counting people, so only whole numbers are possible.

   (e) The values would be "yea," "nay," "abstain," or "not present," (or something similar). These are not numbers, so the data are qualitative.