Math 121 **Test 1 Solutions** **Spring 2006**

$n = 20$, $(52, 62.25, 68, 89.5, 94)$, $\overline{x} = 73$

1. (10 pts) $n = 20$, $(6, 7.75, 10, 10, 10)$, $\overline{x} = 8.8$.

   (a) A Type I error would be to believe that your quail-hunting buddy was not in your line of fire when in fact he was.

   (b) A Type II error would be to believe that your quail-hunting buddy was in your line of fire when in fact he was not.

   (c) A Type I error in this situation is more serious because it could lead to a serious injury or death.

2. (18 pts) $n = 20$, $(0, 8.5, 11, 13.5, 18)$, $\overline{x} = 10.4$. If you gave the wrong direction of extreme in part (a), I took that into account when grading the other parts. For example, if you thought the direction of extreme was to the right, then in part (b), you should get $\frac{10}{15}$ instead of $\frac{5}{15}$ because you went in the opposite direction.

   (a) The direction of extreme is to the left. That it the direction in which the alternative hypothesis becomes the more plausible explanation and the null hypothesis becomes less plausible.

   (b) To find α, count the number of points (X's) in the rejection region under the null hypothesis, which includes all values less than 4. The value of α is $\frac{5}{15}$.

   (c) To find β, count the number of points in the acceptance region under the alternative hypothesis. The value of β is $\frac{1}{15}$.

   (d) The p-value of 4 is the proportion of points at or to the left of 4 under the null hypothesis. That would be $\frac{10}{15}$.

   (e) This is similar to the previous part, but start at 1 and go left. The p-value is $\frac{1}{15}$.

3. (25 pts) $n = 20$, $(14, 18, 21, 24, 25)$, $\overline{x} = 20.9$.

   (a) The sample is the set of 1000 Palestinians, aged 15 and older, living in the West Bank, the Gaza Strip, and East Jerusalem who were interviewed.

   (b) The population should be as similar to the sample as possible. The best answer is that the population consists of all Palestinians aged 15 or over. The sample came only from the West Bank, the Gaza Strip, and East Jerusalem purely for convenience, not because they were interested only in those cities.

   (c) The sample size is 1000.

   (d) This was an observational study. The researchers did not manipulate the levels of any explanatory variables.

(e) The 51% figure is a statistic because it is calculated from a sample.

(f) The results are not statistically significant. The stated conclusion is that the percentages today are similar to what they were six years ago. That is in agreement with the null hypothesis. For the results to be significant, the researchers would have to reject the null hypothesis in favor of the alternative hypothesis.

(g) The variable is qualitative. The numbers 0 through 10 are used only as labels.

4. (14 pts) $n = 20$, $(6, 10, 11.5, 14, 14)$, $\overline{x} = 11.8$.

   (a) Set the seed to 73, and then use `randInt(1,1350)`. Press `ENTER` five times and get the numbers 576, 649, 610, 343, and 174.

   (b) To find k, divide 1350 by 50. The result is 27.

   (c) Set the seed to 48. Then use `randInt(1,27)` to get a random integer from 1 to 27. You should get 22. Then repeatedly add 27 until you have five numbers. The values are 22, 49, 76, 103, and 130.

5. (9 pts) $n = 20$, $(3, 5.75, 6, 7.5, 9)$, $\overline{x} = 6.2$.

   (a) It is response bias. From the wording of the questions, it is clear that there is a preferred answer. The subjects are more likely to give what they perceive to be the preferred answer.

   (b) I was looking for two types of bias *in addition* to response bias from part (a), but I didn't make that clear. The best answers are non-response bias, because of the questionnaires that were not returned, and selection bias, because they *sysematically* sampled from a small part of the population that was likely to have a substantially different opinion from the population at large.

6. (5 pts) $n = 20$, $(1, 1, 5, 5, 5)$, $\overline{x} = 3.65$. The researchers should double-blind the experiment. If the researchers divide the subjects into two groups and give one group the new drug and the other group a placebo, and if the researchers (the ones who make the observations) themselves do not know which subject is in which group, then there should be no bias in the observations.

7. (19 pts) $n = 20$, $(0, 6, 14, 17, 19)$, $\overline{x} = 11.45$.

   (a) The most important reason is that the data are not qualitative. The fact that they do not add up to 100 percent is not relevant. Qualitative data do not "add up" to 100 percent, either. Instead, we count the number of observations in each category and form percentages. That could be done in this example, but it would be far from the best way to go.

   (b) There are too few stems and too many leaves for each stem. We cannot discern the shape of the distribution.

(c) The stem-and-leaf plot with split stems:

0	2 3 4 4
0	6 6 7 8 8 8 8 9 9 9
1	0 0 1 1 3 3 3 3
1	5 7 8
2	1 3
2	
3	0 2

(d) The distribution is unimodal and skewed right. The skewness is not pronounced, so if you didn't mention it, I let it pass.

(e) Yes, the histogram also appears to be unimodal and skewed right.