## 36-201 Spring 1999 Solutions to Homework 7

- (a) (i) The study "New Marrow May Help With SCIDS" is an observational study since the doctors merely observed and reported the results in a group of transplanted patients. There was no control or treatment comparison group. The study "Study Questions Growth Hormones" may have been an experiment since the report states results as "increasing the target heights", meaning that there are previous studies determining expected sizes. However, that is not clearly stated.
  - (ii) If the first study were to be done as an experiment, we would need to randomly select children to have transplants and to prevent other children from having transplants, which is probably unethical.
  - (b) (i) In both cases, we are not given the results of a similar group with no intervention. In particular, the first study does not compare its results to any other treatment, so we don't know how "good" are the results they are reporting. The second study states a comparison with respect to expected heights, however we are not told about the source of the expected heights.

More generally, not having a control group means that we can't distinguish between improvements that would have happened anyway and improvements due to the treatments.

- (ii) In both cases, having a control group means leaving a group of children with no treatment. In the first case, it is clear that it might have been unethical since the life of the child is at risk. In the second case, it is not that clear since most of the cases were cosmetical cases.
- (iii) In both studies, there was a strong expectation about what would happen if the treatment were withheld ("SCIDS once considered almost always fatal" and "average of about 2 inches taller than would have been expected otherwise"). Especially in the bone marrow study, there is no expectation of spontaneous recovery so the improvement in life expectancy could be attributable to the transplant. In the hormone study, the improvement was small so it is believable that the hormone (or anything else that happened to the children in the study period) was simply ineffective.
- 2. <u>Moore</u>, 1.17 (p.31). The reported **6.5%** is a statistic since it represents the percentage of unemployed in the 60,000 household sample, not in the total population.

<u>Moore</u>, 1.18 (p.31). The number **2.503** represents the value of the parameter since it is the average diameter of all balls in the carload. The number **2.515** is a statistic since it is the average diameter of a sample taken from the balls in the carload.

<u>Moore</u>, 1.19 (p.31). 43 (out of 100) is a statistic since is the number of unlisted telephones in the sample. The reported 52% is a parameter since that is the actual percentage of unlisted phone numbers, out of the total population of phones in Los Angeles.

<u>Moore</u>, 1.20 (p.31). **68%** is a parameter since it is the actual percentage of voters who are registered Republicans. On the other hand, **73%** is a statistic since that is the percentage obtained in a sample of voters.

- 3. Moore, 1.22 (p.31-32).
  - (a) High bias, low precision.
  - (b) Low bias, high precision.

- (c) Low bias, low precision.
- (d) High bias, high precision.

<u>Moore</u>, 1.23 (p.32). If both samples are taken under the same circumstances, increasing the size will not reduce the bias. However, it will improve the precision of the results since there will be less uncertainty.

- 4. Moore, 1.44 (p.52).
  - (a) FBI reported less rapes than the National Crime Victimization Survey because their source is the list of *reported* rapes, which is clearly a subset of the total number or rapes.
  - (b) There is no margin of error associated to the FBI data because they are reporting the exhaustive list of reported rapes, not a probability-based or random sample.
  - (c) The FBI data is based on rapes reported by local agencies, who may under-report crimes, as discussed in part (a). On the other hand, there may be bias in the National Crime Victimization Survey because people don't understand what exactly constitutes a crime or may remember dates inaccurately.
- 5. <u>Moore</u>, <u>1.52</u> (<u>p.55</u>). The two questions produced different responses because they are worded differently. For example, the first question is confusing since the response "*impossible*" implies a double negative ("*impossible that the Holocaust never happened*") which produces confusion among respondents. This is not true in the second question, since neither answer implies a double negative.

Moore, 1.53 (p.56-57). The main advantage of using Closed questions is that they are easier to process, while Open questions can get vague or hard to categorize responses.

On the other hand, the advantage of using Open questions is that they can uncover aspects that we don't think about when making the questionnaire. Closed questions can totally miss an important point that is in people's mind.