36-201 Spring 1999
Solutions to Homework 6

1. Moore, 5.18 (p.300-301).

   (a) The total number of Alaska Airlines delayed fights is 501 while the their total number of flights is 3775. So 501/3775 $\times$ 100 % = **13.3%** of Alaska Airlines flights were delayed.

   The total number of America West delayed fights is 787 while their total number of flights is 7225. So 787/7225 $\times$ 100 % = **10.9%** of America West flights were delayed.

   (b) For each airline, we find those percentages as the rate of delayed flights in each airport over the total number of flights of the airline in the same airport. For instance the percentage of delayed flights of Alaska Airlines in Los Angeles is 62/559 $\times$100 % = 11.1%, and similarly for the remaining airports and airline.

Airport	Alaska Airlines	America West
Los Angeles	11.1 %	14.4%
Phoenix	5.2%	7.9%
San Diego	8.6%	14.5%
San Francisco	16.9%	28.7%
Seattle	14.2%	23.3%

   (c) As the book says, America West does worse at every one of the five airports, yet it does better overall. The next table shows the percentages of flights in each airport, by airline (column percentages):

Airport	Alaska Airlines	America West
Los Angeles	14.8 %	11.2%
Phoenix	6.2%	72.8%
San Diego	6.1%	6.2%
San Francisco	16.0%	6.2%
Seattle	56.9%	3.6%
Total	100%	100%

   From the table in (b) we see that both airlines do best at Phoenix, and from the table above we see that America West has actually the majority of its flights there (while Alaska Airlines has only 6.2%). On the other hand, both airlines have a high percentage of delayed flights in Seattle, however America West has only a 3.6% of its flights in Seattle (while Alaska Airlines has 56.9%).

   So, the large percentage of flights that America West flies out of Phoenix, where being on-time is easy, together with the large percentage of flights that Alaska Airlines flies out of Seattle, where being on time is hard (due to weather?) makes America West look better overall.

2. Moore, 1.1 (p.9).

   The population is all employed adult women. The sampling frame is the 520 women in the list of members of the local business and professional women's club. The actual sample corresponds to the 68 women who returned their questionnaires.

Moore, 1.8 (p.11).

The Hite's sampling method is biased because of voluntary response; those women who are more angry towards men are more likely to respond and send the questionnaire back than those having more favorable opinions. The 91% reported as the percentage of women who initiated their divorce is likely to be overestimating the percentage in the population since the sample was biased towards women with bad marriage experiences.

Moore, 1.11 (p.17).

We begin by assigning two digit labels from 01 to 30 to the students in the class, in alphabetical order:

Label	Lastname	Label	Lastname	Label	Lastname
01	Anderson	11	Gutierrez	21	Patnaik
02	Aspin	12	Green	22	Pirelli
03	Bennett	13	Harter	23	Rao
04	Bock	14	Henderson	24	Rider
05	Breiman	15	Hughes	25	Robertson
06	Castillo	16	Johnson	26	Rodriguez
07	Dixon	17	Kempthorne	27	Siegel
08	Edwards	18	Liang	28	Tompkins
09	Fernandez	19	Laskowsky	29	Vandegraff
10	Gupta	20	Olds	30	Wang

The first two digit numbers at line 145 are 19 68 71 26 33 57 85 79 58 06 09 93. We discard those numbers greater than 30 and choose the underlined labels, corresponding to students Laskowsky, Rodriguez, Castillo and Fernandez.

If we choose to label the students by row (instead of by column) then the selected students are Dixon, Laskowsky, Pirelli and Rao.

3. Moore, 2.1 (p.101).

No, this is not an experiment but an observational study since there are no treatments applied. The explanatory variable is Gender and the response variable is Political Party.

Moore, 2.3 (p.101).

This not an experiment because no treatment is assigned to the persons in the sample. It is an observational study because we only observe preferences between the new an the competing brand muffins.

Moore, 2.8 (p.103).

(a) This is not an experiment but an observational study because there is no imposed treatment.

(b) The explanatory variable is whether the family lives in public housing or not. The response variable(s) are the variables measuring family stability (the book does not list them).

(c) Yes, there may be confounding in this study because the acceptance/rejection criteria for giving public housing to families may have included an evaluation of family stability. So we may find differences in stability between these two groups just because groups were created in that way, and not due to the fact that they are or are not living in public housing.

4. Moore, 2.18 (p.111).

   (a) It is likely that students who are more comfortable with the subject choose to take the self-paced instruction so their scores are likely to be higher even if the teaching methods are equally effective.

   (b) Assign labels from 01 to 30 to each student. From Table A in the Appendix, starting at line 108 we get the first 15 two digit numbers between 01 to 30 (discard repeats):

   60 94 07 20 24 17 86 82 49 43 61 79 09 06 56 87 96 41 88 83 36 00
   91 93 65 15 41 23 96 38 85 45 34 68 16 83 48 54 19 79 38 44 84 87
   89 18 33 82 46 97 39 36 44 20 06 76 68 80 87 08 81 48 66 94 87 60
   51 30 92 97 00 41 27 12

   The students having the underlined labels are assigned to the self-paced instruction course and the remaining 15 students assigned to the classroom instruction course.

5. Moore, 2.33 (p.124).

   (a) A control group is necessary in this experiment since people may have decreased their anxiety just because they took a class (not necessarily a meditation class).

   (b) The experimenter may be expecting to find a positive effect on anxiety in the group of people who took meditation class. The knowledge of who did meditate and who did not may bias his diagnosis.

   (c) We may build a control group by choosing a group of people with similar characteristics to the experimental group and making them take other kinds of classes. We should diagnose them before and after taking the classes, as we do with the experimental group. To blind the diagnosis we should have each diagnosis be made by someone who doesn't know which class was taken by each evaluated person.

Moore, 2.35 (p.125).

Although the groups are randomized at the beginning of the experiment, making the two groups comparable, dropouts and non-adherence may not occur at random, making the final groups not comparable. If, for instance, one treatment has undesirable side effects it may have higher dropout rate than the other, making it harder to compare the effects of the treatments themselves on the target disease.