36-201 Spring 1999
Solutions to Homework 8

1. (a) (i) The margin of error for the proportion of females using Internet over a year ago is $\pm 2\ \sqrt{.5 \times .5/1106} = 0.030$. The margin of error for the proportion of females using Internet less than a year ago is $\pm 2\ \sqrt{.5 \times .5/879} = 0.034$, so
      * The percent of females among users who started using the Internet over a year ago is likely between 42 and 48% ($45 \pm 3\%$).
      * The percent of females among users who started using the Internet less than a year ago is likely between 49 and 55% ($52 \pm 3\%$).

   (ii) The percentages of females using the Internet more than one year ago and less than a year don't overlap. The lower bound of the interval for the proportion using Internet less than a year is greater than the upper bound of the interval for the proportion using Internet more than a year, so there was a statistical increase in the proportions.

   (b) (i) The listed sample sizes and margins of error are

sample size	2000	1993	1390	1016	1005	995	977	315
margin of error (%)	3	3	3	3.5	3.5	3.5	3.5	6

   The margin of error increases as the sample size decreases because the smaller the sample the lesser information that we have. This can also be seen from the formula of the margin of error $2\sqrt{.5 \times .5/\text{sample size}}$, which is decreasing in sample size.

   (ii) They take actions to compensate for biases in the sample by using of a random number generator to select phone numbers instead of using the phone book (eliminates the bias produced by being or not listed), designing a generator such that the sample covers the nation in approximately the same way telephone numbers do, designing the mechanism of selecting the youngest male over 18 years of age or the oldest female over 18 years of age (compensates for age biases), and weighting the sample so that it matches some population characteristics.

   (c) There are a lot of private questions or questions that we would not like to be released. For instance those related to the way the person spends his/her day, his/her political position, and personal questions as trustability of people, among others.

2. `Moore, 1.33 (p.41-42)`.

   (a) The value of the sample proportion who prefer balancing the budget is $\hat{p} = 702/1190 \times 100\% = 59\%$. The population parameter p represents the proportion of the adult population who prefer balancing the budget over cutting taxes.

   (b) We are 95% certain that the proportion of adults who prefer balancing the budget is between 55% and 63% ($59\% \pm 4\%$).

   (c) The margin of error of a 99% confidence interval would be greater than the margin of error of a 95% confidence interval.

   (d) Keeping the same confidence of 95%, we should increase the sample size in order to reduce the margin of error.

3. `Moore, 1.34 (p.42)`. Although we are 95% certain that the proportion favoring Bush is between 51% and 55%, there is still a 5% of the samples that lead us to intervals not containing the true population proportion. So we cannot be 100% certain.

`Moore, 1.38 (p.42).` We cannot make a confidence statement about the result because no sample size is given to compute the margin of error.

4. `Moore, 7.3 (p.413).` I got 20 heads out of 50 spins of a penny. Based on this result, the estimated probability of heads is $20/50 \times 100\% = 40\%$.

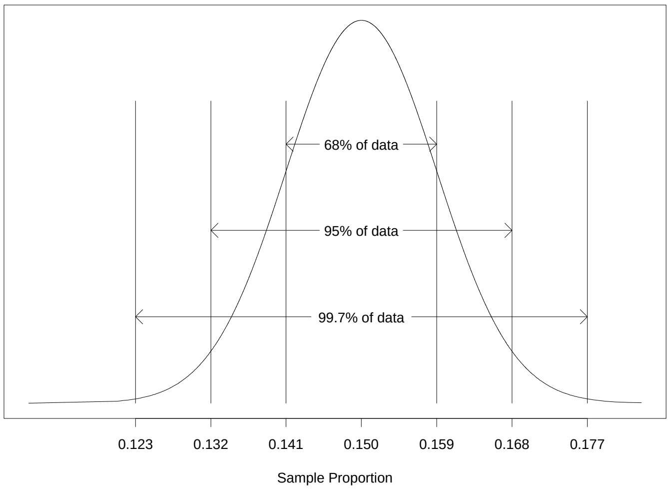
   `Moore, 7.6 (p.413).` That means that in the long run, we should expect to receive three of a kind in 1 out of 50 delivered hands (or 2% of the time).

5. `Moore, 7.9 (p.414).`

   (a) Probabilities must add to 1, so the probability that a woman says that her husband "Does less than his fair share" is 1 - 0.12 - 0.61 = 0.27.

   (b) Since the first two outcomes "Does more than his fair share" and "Does his fair share" are exclusive, the probability of "Does at least his fair share" is 0.12 + 0.61 = 0.73.

6. `Moore, 7.17 (p.416-417).`



   (a) The normal curve is symmetric around 0.15, so half of the area is to the left of 0.15 meaning that about half of the samples, or 50% of them, will have a proportion $\hat{p} \leq 0.15$.

   (b) Using the 68-95-99.7 rule, we know that approximately 68% of the area is between 0.141 and 0.159 (0.150 $\pm$ 0.009). So the probability that $\hat{p}$ takes a value between 0.141 and 0.159 is 0.68.

   (c) The probability that $\hat{p}$ **doesn't** lie between 0.141 and 0.159 is 1 - 0.68 = 0.32.