3.8. The number 3.9% was calculated from the sample of 60,000 households. It is a statistic.

3.13. I used R to take simple random samples. I first assigned the following numerical codes to the club members:

Alonso	1	Darwin	6	Hernstein	11	Myrdal	16	**Vogt**	21
Binet	2	Epstein	7	**Jimenez**	12	**Perez**	17	Went	22
Blumenbach	3	Ferri	8	Luo	13	**Spencer**	18	Wilson	23
Chase	4	**Gonzales**	9	**Moll**	14	Thomson	19	Yerkes	24
Chen	5	Gupta	10	**Morales**	15	Toulmin	20	Zimmer	25

Female club members are shown **bolded**. In R, I entered the following commands:

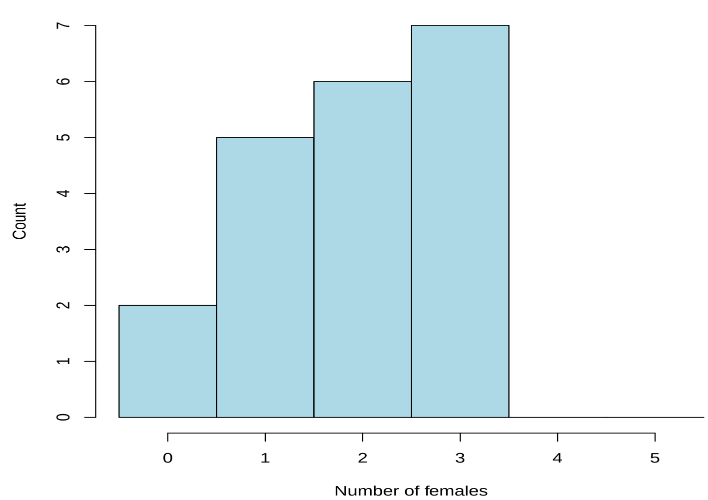
```
> members = seq(1,25,1)
> sample(members,5,replace=F)
[1]  8  7  2 24 23
```

This random numbers correspond to the SRS: Ferri, Epstein, **Binet**, Yerkes, Zimmer. The number of females in this sample is 1.

Now, we have to do this 19 more times! Instead of showing you each sample like I did above, I will just show you the number of females selected in each SRS I selected:

1 2 2 3 3 3 2 1 3 2 1 2 2 0 3 3 1 3 1 0

Here is a histogram of these values:



I used the following R code to make the histogram (last page):

```
no.females = c(1,2,2,3,3,3,2,1,3,2,1,2,2,0,3,3,1,3,1,0)
bins = seq(-0.5,5.5,1)
hist(no.females,breaks=bins,xlab="Number of females",ylab="Count",
    main="",col="lightblue")
```

The average number of females is 1.9. I used R to calculate this:

```
> mean(no.females)
[1] 1.9
```

(b) Two out of the 20 SRS's that I took above did result in "zero females being chosen." So, we know it is possible to get no females when a simple random sample is taken. However, it is more likely that the sample will contain at least one female (18 out of the 20 samples did so).

3.18. Figure 3.2 on pp 40 (textbook) shows simulated values of $\widehat{p}$ when $p = 0.5$ and $n = 1015$. From the histogram, we can see $\widehat{p} = 0.54$ would not be too likely if p really is 0.5. We certainly would not expect to see $\widehat{p} = 0.42$. This statistic is even further away from the truth about the population ($p = 0.5$).

Histograms like those in Figure 3.2, and those in the notes, show us which values of the sample proportion $\widehat{p}$ we would expect to see. Values far away from the center and outside the range of those in the histogram *could* be observed (it is possible). However, the further away from the center they are, the less likely we would see them.

It's important to remember that our conclusions here are based on the assumption that $p = 0.5$ (i.e., 50% of the population feel childhood vaccination is extremely important). If this number was different, then our characterization of "likely" and "not likely" would change depending on what the truth is.

3.19. (a) The sample proportion is

$$\widehat{p} = \frac{627}{1028} \approx 0.61 \text{ (or 61\%)}.$$

Recall this is a statistic because it is calculated from the sample (the 1028 adults). The population here is "all American adults." Therefore,

$$p = \text{population proportion of American adults who are satisfied with the total cost they pay for health care.}$$

This is a parameter because it describes the population.

(b) To write the confidence statement, we need the margin of error. The problem gives the margin of error to be 0.04 (or 4%). Here is the confidence statement:

- "We are 95% confident that the proportion of American adults who are satisfied with the total cost they pay for health care is between 0.57 and 0.65 (i.e., between 57% and 65%)."

3.20. This exercise deals with Figure 3.5 on pp 54 (textbook).
(a) There is high bias. Almost all of the sample statistics are below the parameter, which suggests underestimation on average. There is also high variability.
(b) There is low bias. The sample statistics cluster around the parameter (in the center). There is also low variability.
(c) There is low bias. The sample statistics cluster around the parameter (in the center). However, there is high variability.
(d) There is high bias. All of the sample statistics are above the parameter, which suggests overestimation on average. There is low variability.

3.28. The sample is the 1028 American adults contacted. The population is all American adults.
(a) The sample proportion is

$$\widehat{p} = 0.55 \text{ (or } 55\%).$$

Fifty-five percent of the total sample size is

$$0.55 \times 1028 = 565.4.$$

Rounding off (the 55% statistic is probably rounded), this means about 565 adults in the sample are in favor of the death penalty for a person convicted of murder.
(b) The sample proportion $\widehat{p} = 0.55$ is an estimate of the population proportion p. The margin of error is a number that describes how variable (precise) this estimate is. We can use the margin of error (0.04 or 4%) to write the following confidence statement:

- "We are 95% confident that the proportion of American adults who are in favor of the death penalty for a person convicted of murder is between 0.51 and 0.59 (i.e., between 51% and 59%)."