Practice with Chance Errors in Sampling

1. Complete the following table for the tossing of an unbiased coin.

	Number of heads		*Percent of heads*	
Number of tosses	*Expected value*	*SE*	*Expected value*	*SE*
100	50	5	50%	5%
625				2%
2,500				
6,400				
10,000				

2. You are drawing at random from a large box of red and blue marbles. Fill in the blanks.

   (a) The expected value for the percentage of reds in the _____ equals the percentage of reds in the _____. *Options:* sample, population.

   (b) As the number of draws goes up, this SE for the _____ of reds in the sample goes up but the SE for the _____ of reds goes down. *Options:* number, percentage.

3. A polling organization takes a simple random sample of 2,000 registered voters from the 185 million registered voters in the US. In the sample 1,240 voters are for. Fill in the blanks, using the options below.Explain briefly.

   (a) The observed value of the _____ is 1,240.

   (b) The observed value of the _____ is 62%.

   (c) The expected value of the _____ is equal to the _____.

   Options:

   (i) number of registered voters in the sample who are for

   (ii) percentage of registered voters in the sample who are for

   (iii) percentage of registered voters in the whole country who are for

4. In a certain town in Utah there are 40,000 registered voters, of whom 36,000 are Republicans. A survey organization is about to take a simple random sample of 3,600 registered voters.

   (a) The expected value for the percentage of Republicans in the sample is _____. The SE for the percentage of Republicans in the sample is _____.

   (b) The percentage of Republicans in the sample is likely to be around _____, give or take _____ or so.

   (c) Find the chance that more than 89% of the registered voters in the sample are Republicans.

5. Five hundred draws will be made at random with replacement from the box

60,000 [0]'s	20,000 [1]'s

.

True or false, and explain:

(a) The expected value for the percentage of 1's among the draws is 25% exactly. There is no chance error in that value.

(b) The percentage of 1's among the draws will almost certainly be exactly 25%.

(c) The percentage of 1's among the draws is a number that comes from a random process. We can only state an approximate value. It is subject to chance error.

(d) The percentage of 1's in the box is exactly 25%. This number is known once the box is specified.

Answers follow on the next page.

Answers.

1.

	Number of heads		*Percent of heads*	
Number of tosses	*Expected value*	*SE*	*Expected value*	*SE*
100	50	5	50%	5%
625	312.5	12.5	50%	2%
2,500	1250	25	50%	1%
6,400	3,200	40	50%	0.625%
10,000	5,000	50	50%	0.5%

2. (a) sample, population
   (b) number, percentage

3. (a) (i) (b)(ii) (c) (ii) (iii)

   Comment. The expected value of the *sample percentage* equals the *population percentage*. This is so even after the sample is drawn—the expected value is sort of the average over all possible samples, not just the particular sample you happened to draw. In the frequency theory, it is a mistake to say that the expected value of the *population percentage* equals the *sample percentage*. See section 21.3.

4. The first thing to do it to set up a box model. There should be 40,000 tickets in the box, one for each registered voter; 36,000 are marked 1 (Republican) and 4,000 are marked 0. The number of Republicans in the sample is like the sum of 3,600 draws from the box. The fraction of 1's in the box is 0.9. The expected value for the sum is $3{,}600 \times 0.9 = 3(,)240$. The SD of the box is $\sqrt{0.9 \times 0.1} = 0.39$. The SE for the sum is $\sqrt{3600} \times 0.3 = 18$.

   (a) The expected value for the percent is 36,000 out of 40,000, or 90%. (Or simply observe that the expected value for the percentage of Republicans in the sample is equal to the percentage of Republicans in the town: $36{,}000/40{,}000 \times 100\% = 90\%$.)

   The SE for the percent is 18 out of 3,600 or 0.5%.

   The SE for the percent may also be calculated directly from the SD of the box and the sample size by

   $$\frac{\text{SD of the box}}{\sqrt{n}} \times 100\% = \frac{0.3}{\sqrt{3{,}600}} \times 100\% = 0.5\%.$$

   (No surprise about the expected value: 90% of the registered voters are Republicans.)

   (b) The percentage of Republicans in the sample will be about 90.0%, give or take 0.5% (one half of one percent) or so. Parts (a) and (b) require the same calculations; in (b) you have to interpret the results.

   (c) 89% is 2 SEs below average; the chance is about 98% (no continuity correction; percentages are already continuous).

5. (a) True. It is an "average" value of all possible sample percentages and it has an exact calculation.
   (b) False. It is subject to chance error.
   (c) True. That is the nature of random processes; chance error must be taken into account. It will be around 25%, give or take 2%.
   (d) True. It is a parameter that can be derived from the contents of the box. Once the box is specified, this number is known exactly.