THE EXACT DISTRIBUTION OF the SAMPLE PROPORTION $\hat{p}$ (Ex 17B)

1) SAMPLING FROM A SMALL POPULATION

If we sample from a small population, then each time we select an object to observe, we reduce the population size and change the probability of observing the attribute we are interested in for the next object we select.

EXAMPLE: A bag contains 6 red balls and 4 blue balls. If we take a sample of size 4 from the bag, what is the probability distribution of the sample proportion $\hat{p}$, where $\hat{p}$ represents the proportion of red balls in the sample?

X = no. of red balls in sample	0	1	2	3	4
$\hat{p}$ = proportion of red balls in sample					
$\Pr(\hat{P} = \hat{p})$					

a. Find $\Pr(\hat{P} = \frac{1}{2})$

b. Find $\Pr(\frac{1}{2} < \hat{P} \leq 1)$

We can create a probability distribution for the random variable $\hat{P}$

$\hat{p}$ = proportion of red balls in sample					
$\Pr(\hat{P} = \hat{p})$					

This is called a sampling distribution, because $\hat{P}$ is a **sample statistic**

2) SAMPLING FROM A LARGE POPULATION

Generally, a sample is selected from a large population which is too difficult to enumerate or even count. An example would be: the population of all adults over the age of 18 in Australia, or all the lamas in Mongolia, or all the mobile phones produced by a company in 1 year of production.

When a population is very large, we make the assumption that the probability of observing the attribute we are interested in remains constant with each selection, regardless of what we have selected before.

This means that the result of each selection (whether we observe the attribute or not) is independent of the previous selection, so we have a **Binomial** situation.

EXAMPLE: Suppose we know that 70% of all 17 year olds in Australia attend school. (This means that the **population proportion** is p = 0.7). Suppose we take a random sample of four 17 year olds and we are interested in the sample proportion of those who attend school.

X = no. of 17 year olds who attend school in sample	0	1	2	3	4
$\hat{p}$ = proportion of 17 year olds who attend school in sample					
$\Pr(\hat{P} = \hat{p})$					

Once again we can summarize the sampling distribution:

$\hat{p}$ = proportion of 17 year olds who attend school in sample					
$\Pr(\hat{P} = \hat{p})$					

a. Calculate the mean, $E(\hat{P})$

b. Find the probability that in a random sample of 4 Australian 17 year olds, the proportion attending school is less than 50%.

MEAN and STANDARD DEVIATION of the SAMPLE PROPORTION when SAMPLING FROM a LARGE POPULATION

As we saw above, when we are sampling from a large population, we have a Binomial situation. This means that we can use our knowledge of the Binomial distribution to find a formula for the mean of the sampling proportion and the standard deviation of the population proportion.

MEAN	STANDARD DEVIATION

If we are selecting a random sample of size n from a large population, then the mean and the standard deviation of the sample proportion are given by:

$E(\hat{P}) = p$

$\sigma(\hat{P}) = \sqrt{\frac{p(1-p)}{n}}$

Suppose that 70% of 17 year olds in Australia attend school. If a random sample of size 20 is chosen from this population, find:

a) The probability that the sample proportion is equal to the population proportion (0.7)
b) The probability that the sample proportion lies within one standard deviation of the mean.
c) The probability that the sample proportion lies within two standard deviations of the mean.