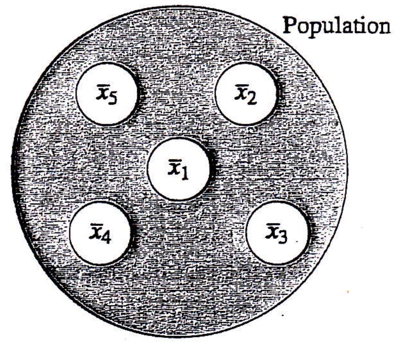
Sampling distribution of the mean (distributions of the sample means)

Suppose a large number of different random samples, each of the same size, n, are selected independently from a population with mean μ and standard deviation σ.
Each of these samples will have its own mean, $\bar{x}$, and standard deviation, s. The set of these different sample means, $\{\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots\}$ are called the **sample means**.

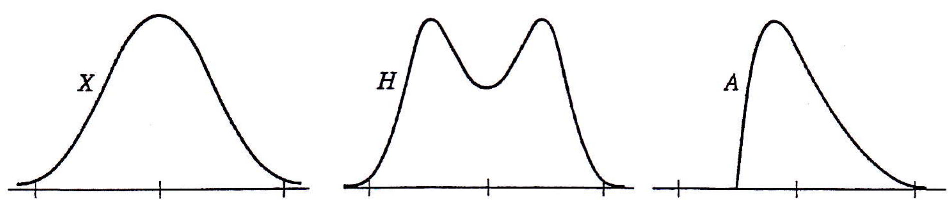


If these sample means are represented with a curve, they have a distribution with the following properties, called the **central limit theorem**.

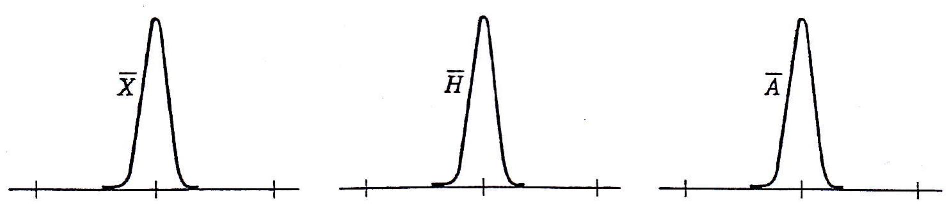
Central limit theorem

1. The distribution of the sample means is always normal (for $n \geq 30$).
2. $$\mu_{\bar{x}} = \mu$$
   Mean of the sample means = Means of the population
3. $$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$
   Standard deviation of the sample means = (standard deviation of the population) $\div \sqrt{n}$

What is remarkable about the central limit theorem is that regardless of the shape of the original distribution, taking averages of samples results in a normal curve. To find the distribution of $\bar{x}$, the sample means, we need to know only the original population mean and standard deviation.



The three probability densities above all have the same mean and standard deviation. Despite their different shapes, when $n = 10$ (or more), the sampling distributions of the mean, $\bar{x}$, are nearly identical and in the shape of a normal curve.



$\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ = standard deviation of the sample proportions

$\dfrac{\sigma}{\sqrt{n}}$ = standard deviation of the sample means