*Sampling distributions:*

In Psychology we generally make inferences about ***populations*** **on the basis of limited** ***samples.***

**We therefore need to know what relationship exists between samples and populations.**

**A population of scores has**
**a *mean* (μ),**
**a *standard deviation* (δ) and**
**a *shape* (e.g., normal distribution).**

**If we take repeated samples of scores, each sample has a mean (X), a standard deviation (s) and a shape.**

**Due to random fluctuations, each sample is *different* - from other samples and from the parent population.**

**Fortunately, these differences are *predictable* - and hence we can still use samples in order to make inferences about their parent populations.**

***What are the properties of the distribution of Sample Means?***

**(a) The mean of the sample means is the *same* as the mean of the population of raw scores:**

$$\mu \bar{X} = \mu X$$

**(b) The standard deviation of the sample means is the "standard error". It is NOT the same as the standard deviation of the population of raw scores.**

$$\sigma \bar{X} \neq \sigma X$$

The standard error *underestimates* the population s.d.:

$$\sigma\bar{X} = \frac{\sigma X}{\sqrt{n}}$$

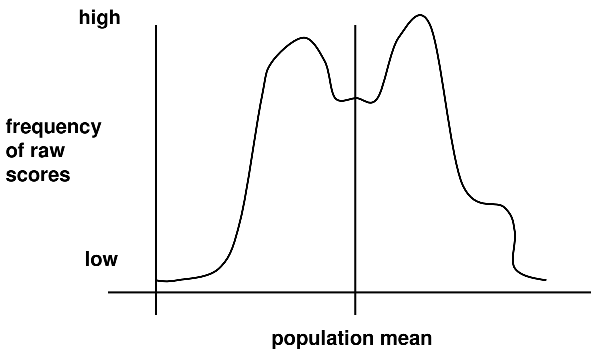
The bigger the sample size, the *smaller* the standard error.
i.e., variation between samples *decreases* as sample size *increases*.
(This is because producing a sample mean reduces the influence of any extreme raw scores in a sample).

(c) The distribution of sample means tends to be normally distributed, *no matter what the shape of the original distribution of raw scores in the population.*
This is called the Central Limit Theorem.
This effect is more marked the bigger the sample size.

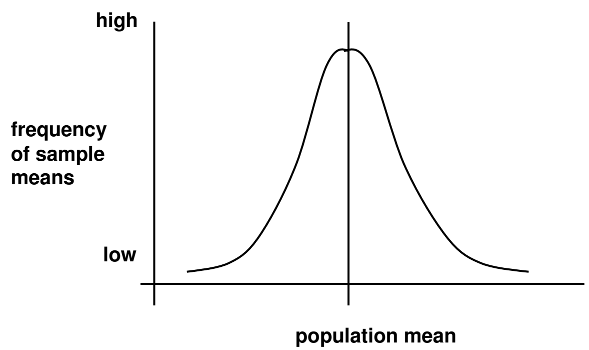
This means that we can use the normal distribution as a simple "model" for many different populations - even though we don't know their true characteristics.

*The Central Limit Theorem in action:*

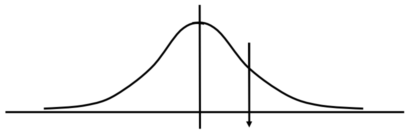
*(a) Frequency distribution of raw scores:*




*(b) Frequency distribution of means of samples taken from the previous population:*



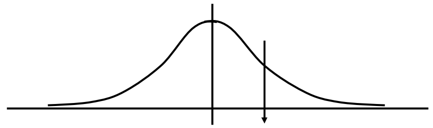
*Relationship of a raw score to the mean of the group of scores:*




**(a) Various proportions of scores fall within certain limits of the mean (i.e. 68% fall within the range of the mean +/- 1 s.d.; 95% within +/- 2 s.d., etc.).**

**(b) Using z-scores, we can represent a given score in terms of how *different* it is from the mean of the group of scores.**

*Relationship of a sample mean to the population mean:*




***We can do the same with sample means:***
**(a) we obtain a particular sample mean;**
**(b) we can represent this in terms of how *different* it is from the mean of its parent population.**

**Sample means vary randomly, but most will be reasonably close to the population mean. (e.g., 68% of sample means will fall within +/- 1 s.d. of the population mean, etc.).**

**If we obtain a sample mean that is *much* higher or lower than the population mean, there are two possible reasons:**

**(a) our sample mean is a rare "fluke" (a quirk of sampling variation);**

**(b) our sample has *not* come from the population we thought it did, but from some other, *different*, population.**

**The greater the difference between the sample and population means, the more plausible (b) becomes.**

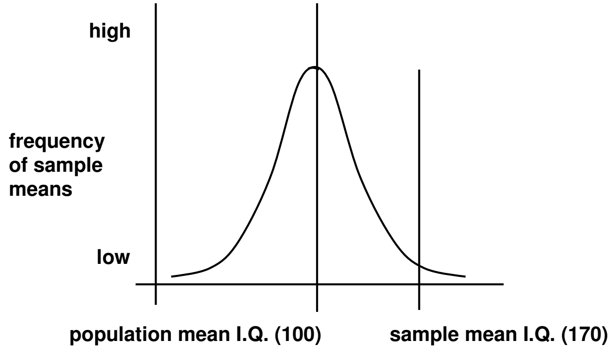
***A concrete example:***

**The human population I.Q. is 100.**
**A random sample of people has a mean I.Q. of 170.**

**There are two explanations:**
**(a) the sample is a fluke: by chance our random sample contained a large number of highly intelligent people.**

**(b) the sample does not come from the population we thought they did: our sample was actually from a different population - e.g., aliens masquerading as humans.**

*Relationship between population mean and sample mean:*



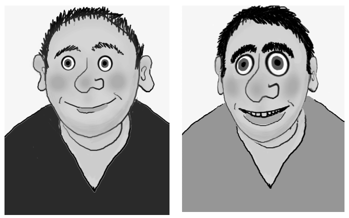

This logic can be extended to the ***difference between two samples from the same population***:

***A common experimental design:***
**We compare two groups of people:**
**an experimental group and a control group.**
**Experimental group get a "wolfman" drug.**
**Control group get a harmless placebo.**

**D.V.: number of dog-biscuits consumed.**

**At the start of the experiment, they are two samples from the same population ("humans").**

**At the end of the experiment, are they:**
**(a) still two samples from the *same* population (i.e., still two samples of "humans" - our experimental treatment has left them unchanged);**
**OR**
**(b) now samples from two *different* populations - one from the "population of humans" and one from the "population of wolfmen"?**



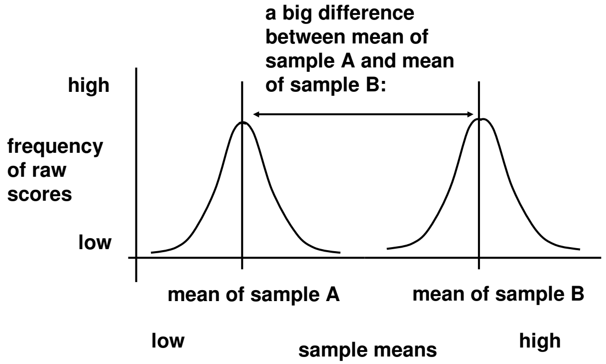
***We can decide between these alternatives as follows:***

**The differences between any two sample means from the same population are normally distributed, around a mean difference of zero.**

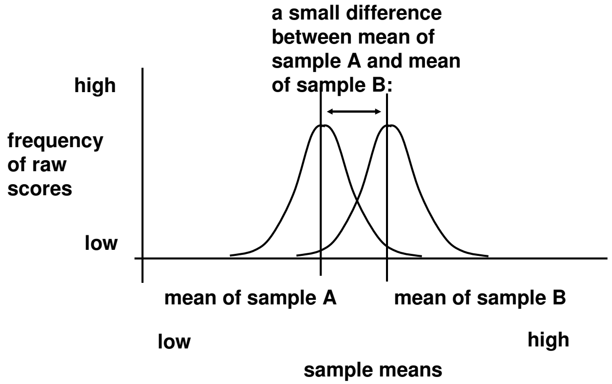
**Most differences will be relatively small, since the Central Limit Theorem tells us that most samples will have similar means to the population mean (and hence similar means to each other).**

**If we obtain a *very* large difference between our sample means, it could have occurred by chance, but this is very unlikely - it is more likely that the two samples come from different populations.**

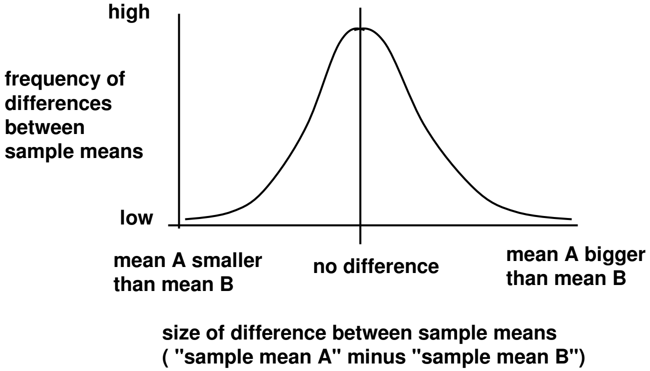
### *Possible differences between two sample means:*



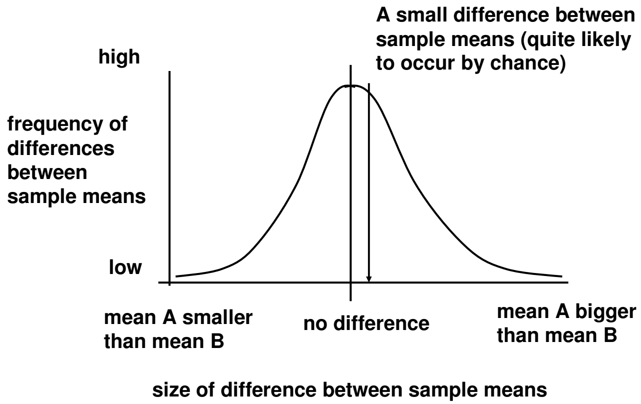

### *Possible differences between two sample means (cont.):*




### *Frequency distribution of differences between sample means:*




### *Frequency distribution of differences between sample means:*



***Frequency distribution of differences between sample means:***