PSYC 2317 - Lecture 4

In Lecture 3, we learned how to compare a single score,
to a distribution of scores

↳ two equivalent types of questions:

(1) what proportion of scores are greater / less than some given score?

(2) what is the probability of randomly selecting a score greater / less than some given score?

Recall from Lecture 1:

In research, we test models of "population" by measuring "samples"

↳ for example, take mean of the sample.

Given this sample mean, how does it compare to the distribution of all possible sample means?

Guiding example: Consider a test whose scores are normally distributed with mean $\mu = 16$ and standard deviation $\sigma = 5$.

What means would we expect if we were to take samples of size $N = 5$?

Lets construct the distribution of all possible sample means.

To learn about these "sampling distributions", we'll do two things:

(1) use the Distributions module of JASP (jasp-stats.org) to quickly see what kinds of means we get from such samples

(2) use an online Java applet to simulate what happens when we take thousands of these sample means.

http://onlinestatbook.com/stat_sim/sampling_dist

SCAN ME

What do we notice?

(1) the distribution of sample means appears normal

(2) the mean of this distribution is 16

↳ same as original distribution of scores.

(3) the SD of this distribution is smaller than original distribution of scores.

↳ hmm... what happens if we increase sample size to $N = 25$?

Ans: it gets even smaller!

Summary – the "central limit theorem"

Consider a (normal) distribution with given mean μ and standard deviation σ. Suppose we take samples of size N. Then the distribution of sample means has the following properties:

* it is (approximately) normal
* it has mean equal to μ
* it has standard deviation equal to $\dfrac{\sigma}{\sqrt{N}}$

Some notes:

(1) the $SD = \frac{\sigma}{\sqrt{N}}$ is called the "standard error" of measurement. We often abbreviate this to SE.

(2) even if the original distribution of scores is NOT normal, the distribution of sample means is approximately normal, if the sample size is large enough.

Ex: Ten students are randomly selected to take the ACT $(\mu = 21, \sigma = 6)$. What is the probability that their mean score is greater than 25?

Solution: we want to know $P(\overline{X} > 25)$.

Step 1: convert to z-score based on distribution of sample means. Remember that this distribution has std. deviation $SE = \frac{\sigma}{\sqrt{N}} = \frac{6}{\sqrt{10}} = 1.90$

$$z = \frac{\overline{X} - \text{mean}}{SE} = \frac{25 - 21}{1.90} = 2.11$$

Step 2: find $P(z > 2.11)$

↳ from online app, this is 0.0174.

Take home:

(1) if we want to answer questions about samples, we need to know something about the distribution of sample means

(2) Compared to original distribution, the Distribution of sample means has:

* same mean

* smaller standard deviation

$\hookrightarrow$ $SE = \dfrac{\sigma}{\sqrt{N}}$