Wednesday, Mar 29

Dependent Samples

With two **dependent samples**, the observations in the two samples are somehow "linked" between the samples by design (e.g., same unit observed twice, twins). A more technical definition is that the *distribution* of an observation in one sample would depend on the value of an observation in the other sample. Matched-pairs designs create dependent samples.

Example: Consider the following data from a study of the effect of in-synchrony versus out-of-sychrony speech on infant attention.[1]

	Synchrony		
Infant	Out	In	Diff
DC	50.4	20.3	30.1
MK	87.0	17.0	70.0
BH	25.1	6.5	18.6
JM	28.5	25.0	3.5
SB	26.9	5.4	21.5
MM	36.6	29.2	7.4
RH	1.0	2.9	-1.9
DJ	43.8	6.6	37.2
JD	44.2	15.8	28.4
ZC	10.4	8.3	2.1
CW	29.9	34.0	-4.1
AF	27.7	8.0	19.7

The observations are of the percent time attending to the person speaking. The average difference is about 19.4% and the standard deviation of the differences is about 20.8%.

[1]Dodd, B. (1979). Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology*, *11*, 478–484.

Independent Samples

With two **independent samples**, the observations in the two samples are *not* dependent.

Example: Had the study of the effect of in-synchrony versus out-of-synchrony speech used a *randomized design* in which each infant was assigned to *either* the in-synchrony or out-of-sychrony condition, it would produce independent samples that might look like this.

Infant	Synchrony	Attention
DC	In	20.3
MK	Out	87.0
BH	Out	25.1
JM	In	25.0
SB	In	5.4
MM	Out	36.6
RH	In	2.9
DJ	In	6.6
JD	Out	44.2
ZC	Out	10.4
CW	In	34.0
AF	Out	27.7

The mean, standard deviation, and sample sizes for the two sample are given below.

Synchrony	$\bar{x}$	s	n
In	15.7	12.6	6
Out	38.5	26.4	6

Example: A large-scale randomized experiment investigated the effect of regular aspirin use on myocardial infarctions (i.e., heart attacks).[2]

Subject	Treatment	Heart Attack?
1	aspirin	no
2	aspirin	no
3	control	yes
4	aspirin	no
5	aspirin	no
6	control	yes
7	control	yes
8	aspirin	no
9	control	no
10	aspirin	yes
$\vdots$	$\vdots$	$\vdots$
22071	aspirin	no

The data can be summarized as follows.

	Heart Attack?		
Group	yes	no	Total
aspirin	104	10933	11037
control	189	10845	11034

[2]Steering Committee of the Physicians' Health Study Research Group. (1989). Final report on the aspirin component of the ongoing Physicians' Health Study. *New England Journal of Medicine*, *321*, 129–135.

Sampling Distribution of $\hat{p}_1 - \hat{p}_2$

What do we know about the sampling distribution of $\hat{p}_1 - \hat{p}_2$?

1. The *mean* of the sampling distribution is $p_1 - p_2$.
2. The *standard deviation* (i.e., standard error) is

$$\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}},$$

   assuming *independent* samples. Why? Because the variance of the *sum* or *difference* of two *indepedent* random variables is equal to the *sum of their variances.*

3. The *shape* of the sampling distribution is approximately that of a normal probability distribution by an application of the central limit theorem.

Confidence Interval for $p_1 - p_2$

A general "recipe" for a confidence interval is

$$\text{point estimate} \pm \text{margin of error},$$

where

$$\text{margin of error} = \text{multiplier} \times \text{standard error}.$$

How then might we compute the point estimate, margin of error, and confidence interval for $p_1 - p_2$, and how might we apply that to the results of the aspirin study?

	Heart Attack?		
Group	yes	no	Total
aspirin	104	10933	11037
control	189	10845	11034

Statistical Test Concerning $p_1 - p_2$

A general "recipe" for a test statistic is

$$\text{test statistic} = \frac{\text{point estimate} - \text{hypothesized mean}}{\text{standard error}}.$$

How might we conduct a statistical test concerning $p_1 - p_2$, and how might we apply that to the results of the aspirin study?

	Heart Attack?		
Group	yes	no	Total
aspirin	104	10933	11037
control	189	10845	11034

Example: A study published in *The New England Journal of Medicine* reported the results of a randomized experiment with 128 children and adolescents to investigate the effectiveness of the drug fluvoxamine in the treatment of anxiety disorders in young people.[3] The study found that 48 out of 63 subjects that were given the drug showed a reduction in anxiety, in comparison to only 19 out of 65 subjects that were not given the drug. What can we infer about the effect of the fluvoxamine on anxiety reduction?

[3] Walkup, J. T. et al. (2001). Fluvoxamine for the treatment of anxiety disorders in children and adolescents. *The New England Journal of Medicine*, 344, 1279–1285.