NOTE Y: CONSTRUCTING A HYPOTHESIS

A STATEMENT OF THE HYPOTHESIS

Goodman (1999) states, "The process that we use to link underlying knowledge to the observed world is called *inferential reasoning,* of which there are two logical types: *deductive inference* and *inductive inference*.

DEDUCTIVE INFERENCE

In deductive inference, we start with a given hypothesis (a statement about how nature works) and predict what we should see if that hypothesis were true. Deduction is objective in the sense that the predictions about what we will see are always true if the hypotheses are true. Its problem is that we cannot use it to expand our knowledge beyond what is in the hypotheses." A Fisher based hypothesis is deductive in nature.

INDUCTIVE INFERENCE

"Inductive inference goes in the reverse direction: On the basis of what we see, we evaluate what hypothesis is most tenable. The concept of evidence is inductive; it is a measure that reflects back from observations to an underlying truth. … Its drawback is that we cannot be sure that what we conclude about nature is actually true, a conundrum known as the *problem of induction*." (Goodman, 1999)

"In their (Neiman and Pearson) hypothesis test, one poses *two* hypotheses about nature: a null hypothesis (usually a statement that there is a null effect) and an alternative hypothesis, which is usually the opposite of the null hypothesis (for example, that there is a nonzero effect)." The researcher here by induction, makes a decision (by acting or behaving) to either accept the null hypothesis or to reject it. The structure of the hypotheses allows for defacto acceptance of the null hypothesis if there is insufficient evidence to act.

"The outcome of a hypothesis test was to be a behavior, not an inference: to reject one hypothesis and accept the other, solely on the basis of the data. This puts the researcher at risk for two types of errors—behaving as though two therapies differ when they are actually the same (also known as a *false-positive result,* a *type I error,* or an α *error*) or concluding that they are the same when in fact they differ (also known as a *false-negative result,* a *type II error,* or a ß *error*)." (Goodman 1999)

THE CONSTRUCT OF THE HYPOTHESIS

THE NULL HYPOTHESIS (H_0):

This is a claim that involves ≤, =, or ≥ relationships between numbers. It involves equality relationships and is based on the Neyman-Pearson view of inductive behavior (Hubbard and Bayarri, 2003).

In regard to experimental data, the common approach is either:

1. A hypothesis that an experimental treatment had no effect (i.e. no difference between a base case and a treatment case, a test of equality)

2. A hypothesis that a treatment caused no adverse effect (adverse being $\leq$ or $\geq$ from a base case).
3. A hypothesis that there is no relationship between a response variable and a predictor variable.

For other situations (which include textbook problems)

1. A hypothesis that a sample comes from a specified population.
2. A hypothesis about numerical values calculated from a sample.
3. A hypothesis about a single observation in regard to a specified population.

Based on these, one constructs a hypothesis and an alternate hypothesis as statements about nature that is to be decided upon, based on tests.

The philosophy of science has emphasized that any test of a theory be a strong test. Because under "Modus Pollens", one cannot prove a hypothesis as being universally true. Any evidence of it being false, is a strong proof of the hypothesis being false, that negates any evidence of the hypothesis being true. This is why a hypothesis is contrary to what the researcher is attempting to achieve.

THE ALTERNATE HYPOTHESIS (H_A):

Here, one finds two possible approaches, the logical alternative, or another point hypothesis. The alternate hypothesis need not be the logical complement of the hypothesis (see). However the tools in Excel only support a logical alternative.

The logical alternative to the null hypothesis Involves $>$, $\neq$, or $<$ relationships between numbers (logical complement of H_0). In regard to experimental data, the alternate hypothesis is that the treatment had a true effect, and it was not due to chance.

PUTTING THE CLAIM INTO THE NULL OR THE ALTERNATE HYPOTHESIS

Claim Statement	Hypothesis	Symbol
m is above k	Alternate	$m > k$
m is an improvement over k (μ_d is +)	Alternate	$\mu_d > 0$
m is an improvement over k (μ_d is –)	Alternate	$\mu_d < 0$
m is at least k	Null	$m \geq k$
m is at most k	Null	$m \leq k$
m is below k	Alternate	$m < k$
m is different from k	Alternate	$m \neq k$
m is equal to k	Null	$m = k$
m is exactly k	Null	$m = k$
m is fewer than k	Alternate	$m < k$
m is greater than k	Alternate	$m > k$
m is greater than or equal to k	Null	$m \geq k$
m is k	Null	$m = k$
m is less than k	Alternate	$m < k$

Claim Statement	Hypothesis	Symbol
m is less than or equal to k	Null	$m \leq k$
m is more than k	Alternate	$m > k$
m is not equal to k	Alternate	$m \neq k$
m is not k	Alternate	$m \neq k$
m is not less than k	Null	$m \geq k$
m is not more than k	Null	$m \leq k$
m is smaller than k	Alternate	$m < k$

The hypothesis here is a statement being made about m and k, before any attempt is made to do a statistical test.

CONSTRUCTING SIMPLE HYPOTHESIS STATEMENTS

Population Mean
For H_a: $\mu > k$, right tail test
For H_a: $\mu < k$, left tail test
For H_a: $\mu \neq k$, two tail test

Population Median
For H_a: median > k, right tail test
For H_a: median < k, left tail test
For H_a: median ≠ k, two tail test

Difference in Population Means
For H_a: $\mu_d > 0$, right tail test
For H_a: $\mu_d < 0$, left tail test
For H_a: $\mu_d \neq 0$, two tail test

Population Variance
For H_a: $\sigma^2 > k$, right tail test
For H_a: $\sigma^2 < k$, left tail test
For H_a: $\sigma^2 \neq k$, two-tail test

Two Population Variances
For H_a: $\sigma^2{}_1 > \sigma^2{}_2$, right tail test
For H_a: $\sigma^2{}_1 < \sigma^2{}_2$, left tail test
For H_a: $\sigma^2{}_1 \neq \sigma^2{}_2$, two tail test, use only the right tail critical value

Correlation Coefficient

H_0	H_a	Correlation	Test
R >= 0	R < 0	Negative	Left Tailed
R <= 0	R > 0	Positive	Right Tailed
R = 0	R ≠ 0	None	Two Tailed

Observed Sample Frequencies (Grouped data)
For H_a: (observed frequencies) > (expected frequencies), right tail test
For H_a: (observed frequencies) < (expected frequencies), left tail test
For H_a: (observed frequencies) ≠ (expected frequencies), two tail test