To Do

Read Sections 7.1 – 7.2.

Assignment 5 is due Monday December 5.

See detailed information regarding Tutorial Test 3 (Wednesday November 30) posted on Learn.

Last Class

(1) Comparing the Means of Two Gaussian Populations

(2) Comparison of Two Means, Equal Variances

Today's Class

(1) Comparison of Two Means,
Unequal Variances and Large Sample Sizes

(2) Analysis of a Paired Experiment

(3) Pairing and Experimental Design

Comparison of Two Means, Unequal Variances

The analysis of last day depended on the assumptions:

$Y_{1i} \backsim G(\mu_1, \sigma_1)$ for $i = 1,2,\ldots,n_1$ independently

and independently

$Y_{2i} \backsim G(\mu_2, \sigma_2)$ for $i = 1,2,\ldots,n_2$ independently

and $\sigma_1 = \sigma_2 = \sigma$.

What if $\sigma_1 = \sigma_2 = \sigma$ is not a reasonable assumption?

Note: H_0: $\sigma_1 = \sigma_2$ could be tested using a likelihood ratio test.

Approximate Pivotal Quantity, Unequal Unknown Variances

If n_1 and n_2 are large (both at least 30) then the approximate pivotal quantity

$$\frac{\bar{Y}_1 - \bar{Y}_2 - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \sim G(0,1) \quad \textbf{approximately}$$

can be used to construct confidence intervals and test hypotheses for the mean difference μ_1 - μ_2.

Approximate Pivotal Quantity, Unequal Unknown Variances

For example, an approximate 95% confidence interval for μ_1 - μ_2 would be given by

$$\bar{y}_1 - \bar{y}_2 \pm 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Handspan Example

For the handspan example an approximate 95% confidence interval not assuming variances are equal is

$$\bar{y}_1 - \bar{y}_2 \pm 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= 2.14 \pm (1.96)\sqrt{\frac{3.4309}{78} + \frac{2.055}{64}}$$

$$= [1.60, 2.68]$$

compared with [1.58,2.70] assuming equal variances and using t distribution.

Smarties Experiment

Count the number of each colour in your box of Smarties.

Please also answer questions on the back.

Dominant versus Non-Dominant Bean Experiment: Data

Difference	-5	-4	-3	-2	-1	0
Frequency	1	1	2	1	4	21

Difference	1	2	3	4	5	6
Frequency	25	27	15	7	2	1

Dominant versus Non-Dominant Bean Experiment: Questions

Is there a difference in the mean number of beans moved in 15 seconds between the dominant and non-dominant hands?

How do we analyze these data?

What model should we assume?

Is this a two sample problem?

Dominant versus Non-Dominant Bean Experiment

Assumptions for two sample model are:

$Y_{1i} \backsim G(\mu_1, \sigma)$ **for** $i = 1,2,\ldots,n_1$ **independently**

and independently

$Y_{2i} \backsim G(\mu_2, \sigma)$ **for** $i = 1,2,\ldots,n_2$ **independently**

Let Y_{1i} **= the number of beans moved using the dominant hand and let** Y_{2i} **= the number of beans moved using the non-dominant hand for the ith student.**

Dominant versus Non-Dominant Bean Experiment: Question

Does it seem reasonable to assume that the Y_{1i} 's and, i =1,2,…,n are independent of the Y_{2i} 's, i = 1,2,…,n?

We would expect the observations on the i'th student (Y_{1i} , Y_{2i}) to be positively correlated, that is, we would expect Cov(Y_{1i} ,Y_{2i}) > 0.

Paired Experiment

In fact the observations have been deliberately paired to eliminate some factors (finger size, agility, competitive spirit, etc.) which might otherwise affect conclusions about the parameter of interest which is the mean difference

$\mu_1 - \mu_2$.

The experiment is an example of a what is called a paired experiment.

Paired Experiment

For a paired experiment (can you show this?):

$$Var\left(\overline{Y}_1 - \overline{Y}_2\right) = \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n} - 2Cov\left(Y_{1i}, Y_{2i}\right)$$

If Cov(Y_{1i}, Y_{2i}) > 0, then

$$Var\left(\overline{Y}_1 - \overline{Y}_2\right)$$

is smaller than for an unpaired experiment.

Paired Experiment

To make inferences about $\mu = \mu_1 - \mu_2$, we analyze the within-pair differences

$Y_i = Y_{1i} - Y_{2i}$, $\quad i = 1,2,\ldots,n$

by assuming

$Y_i = Y_{1i} - Y_{2i}, \sim G(\mu_1 - \mu_2, \sigma) \quad i = 1,2,\ldots,n$

independently.

We then use the one sample analysis that we used previously for analysing a random sample from a $G(\mu,\sigma)$ distribution with $\mu = \mu_1 - \mu_2$.

Bean Experiment

In the Bean experiment

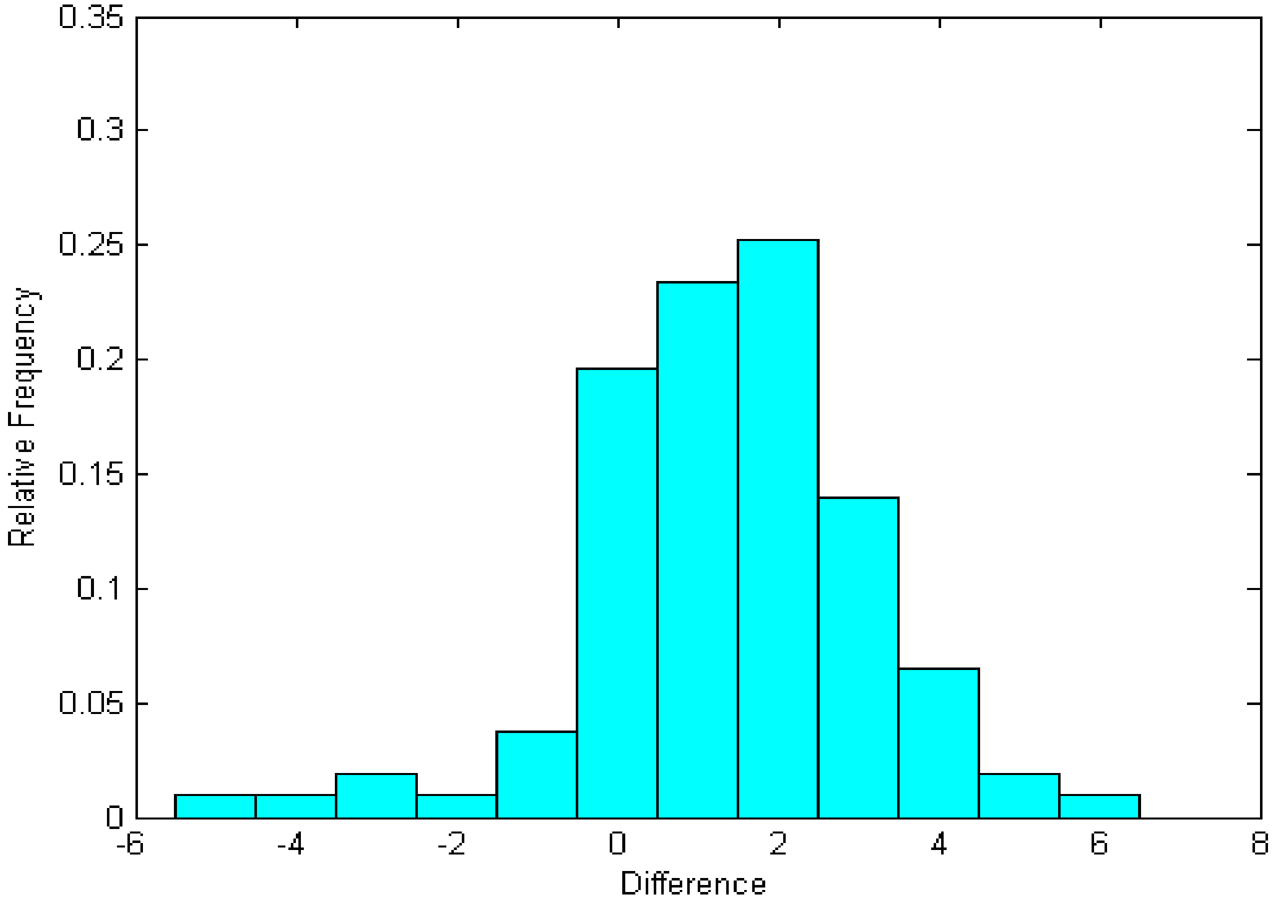
Y_i = the number of beans moved by dominant hand
– number of beans moved by non-dominant hand

Dominant versus Non-Dominant Bean Experiment: Data

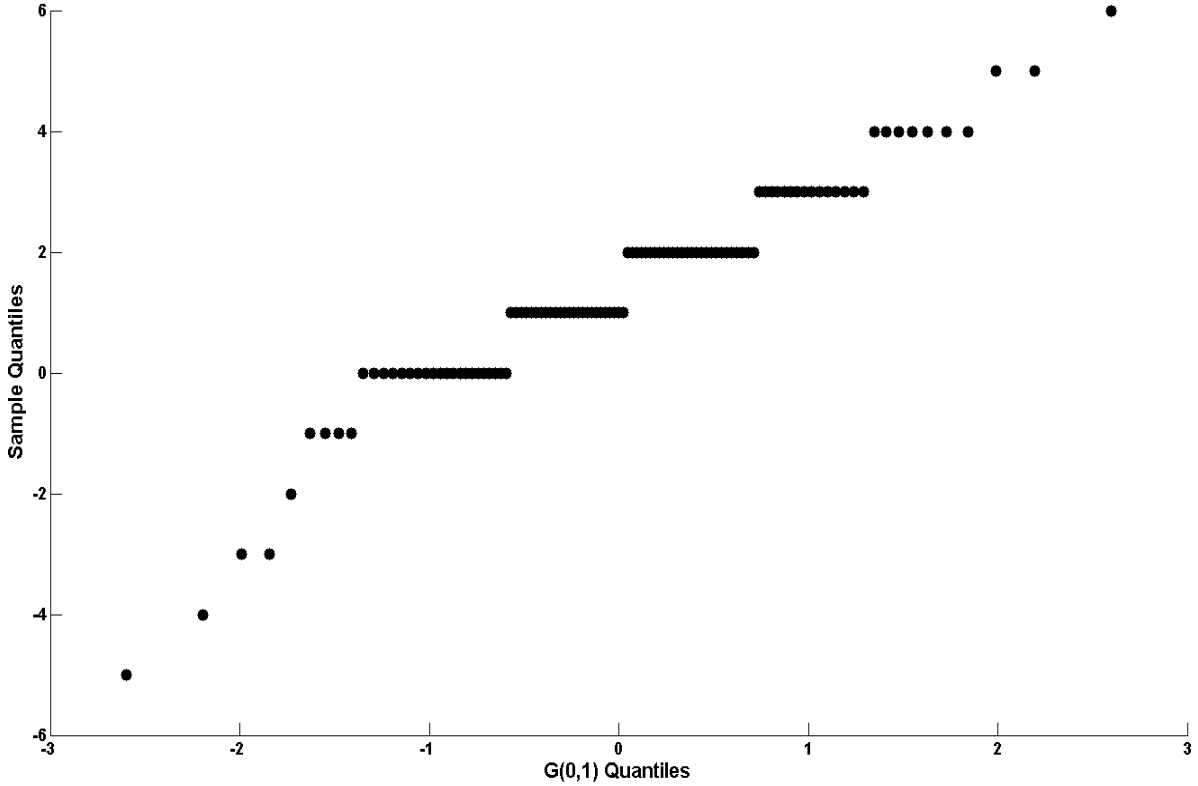
y_i	-5	-4	-3	-2	-1	0
Frequency	1	1	2	1	4	21

y_i	1	2	3	4	5	6
Frequency	25	27	15	7	2	1

Histogram of Differences



Qqplot of Differences



Model

Does it seem reasonable to use the model

$Y_i = Y_{1i} - Y_{2i}$, $\sim G(\mu_1 - \mu_2, \sigma)$ $\quad i = 1,2,\ldots,107$

independently?

Test of H_0: $\mu = 0$

For the bean data

$$\bar{y} = 1.3738, \quad s = 1.7672$$

To test H_0:μ=0 we use the test statistic

$$D = \frac{\left|\bar{Y} - 0\right|}{S / \sqrt{n}}$$

with observed value

$$d = \frac{\left|\bar{y} - 0\right|}{s / \sqrt{n}} = \frac{\left|1.3738 - 0\right|}{1.7672 / \sqrt{107}} = 8.0414$$

p-value

p-value = $2[1 - P(T \leq 8.0414)] \approx 0$

where $T \backsim t(106)$

There is strong evidence against H_0: $\mu = 0$ based on the observed data.

The difference is statistically significant.

Is the difference of practical significance?

Examples in which the parameter of interest is the mean difference $\mu_1 - \mu_2$

(1) Test whether one numerical algorithm for nonlinear optimization is faster than another on a large population of potential test functions.

(2) Test for a difference in the error rates of two algorithms designed for image resolution or character/speech recognition on many different scenarios/problems.

(3) Artificial intelligence: test whether one learning algorithm learns a task faster than another.

DESIGN 1

Generate or select n_1 random "problems" (for example data sets to be sorted, functions to be minimized, images to be resolved,...) and compute the mean execution time Y_1 of algorithm A.

Generate or select another n_2 ($n_1 = n_1$ possibly) random "problems" and compute the mean execution time Y_2 of algorithm B.

DESIGN 1

Estimate the difference as

$$\overline{Y}_1 - \overline{Y}_2$$

In this case $\overline{Y}_1, \ \overline{Y}_2$

are independent random variables with

$$E\left(\overline{Y}_1 - \overline{Y}_2\right) = \mu_1 - \mu_2$$

and

$$Var\left(\overline{Y}_1 - \overline{Y}_2\right) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

DESIGN 2

Generate or select n random “problems” (for example data sets to be sorted, functions to be minimized, images to be resolved,...) and compute the mean execution time Y_1 of algorithm A.

Compute the mean execution time Y_2 of algorithm B on the same set of n problems.

DESIGN 2

Estimate the difference as

$$\overline{Y}_1 - \overline{Y}_2$$

In this case $\overline{Y}_1, \ \overline{Y}_2$

are dependent random variables with

$$E\left(\overline{Y}_1 - \overline{Y}_2\right) = \mu_1 - \mu_2$$

and

$$Var\left(\overline{Y}_1 - \overline{Y}_2\right) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} - 2Cov\left(Y_{1i}, Y_{2i}\right)$$

DESIGN 2

$$Var\left(\overline{Y}_1 - \overline{Y}_2\right) = \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n} - 2Cov\left(Y_{1i}, Y_{2i}\right)$$

Therefore if Cov(Y_{1i} ,Y_{2i}) > 0, then

$$Var\left(\overline{Y}_1 - \overline{Y}_2\right)$$

is smaller than for Design 1.

Pairing as a Design Choice

We expect that the covariance between the execution times of algorithms on the same problem to be positively correlated (harder problems have longer execution times).

A sample of dependent pairs (Y_{1i}, Y_{2i}) is better than two independent random samples for estimating $\mu_1 - \mu_2$ since the difference $\mu_1 - \mu_2$, can be estimated more accurately (shorter confidence intervals) if Cov(Y_{1i}, Y_{2i}) > 0.

Note: If Cov(Y_{1i}, Y_{2i}) < 0, then pairing is a bad idea.

Pairing as a Design Choice

In a paired experiment, Y_{1i} and Y_{2i} are not independent random variables.

We do, however, assume the differences

$Y_i = Y_{1i} - Y_{2i}$ **,** $i = 1,2,\ldots,n$

are independent random variables (all different problems).

See Problem 6.19 - Example on Sorting Algorithms

Paired versus Unpaired

When you see data from a comparative study (i.e. one whose objective is to compare two distributions, often through their means), you have to determine whether it involves paired data or not.

Of course, a sample of Y_{1i} 's and Y_{2i} 's cannot be from a paired study unless there are equal numbers of each, but if there are equal numbers the study might be either “paired” or “unpaired”.

Examples in Course Notes

Example 6.4.3
Heights of males versus females

Example 6.4.4
Comparison of car fuels

Example 6.4.5
Fibre in diet and cholesterol level