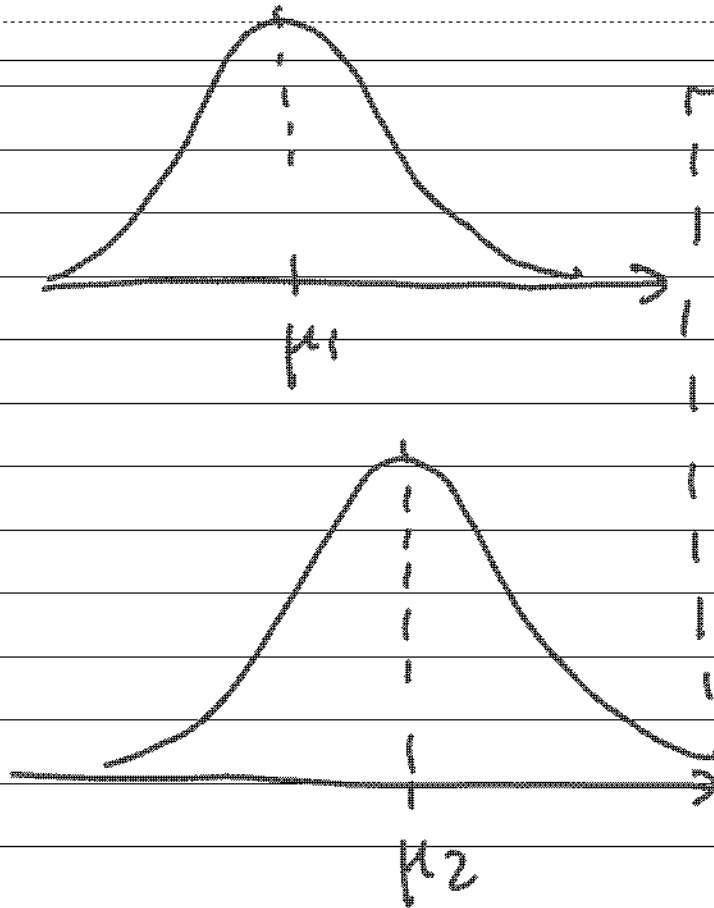
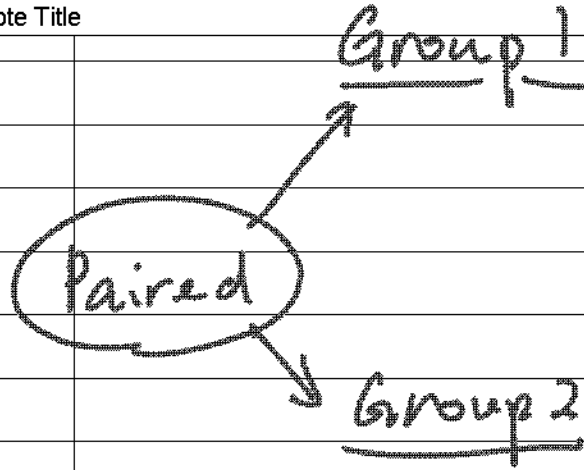


# Test for paired data

Note Title

11/21/2008



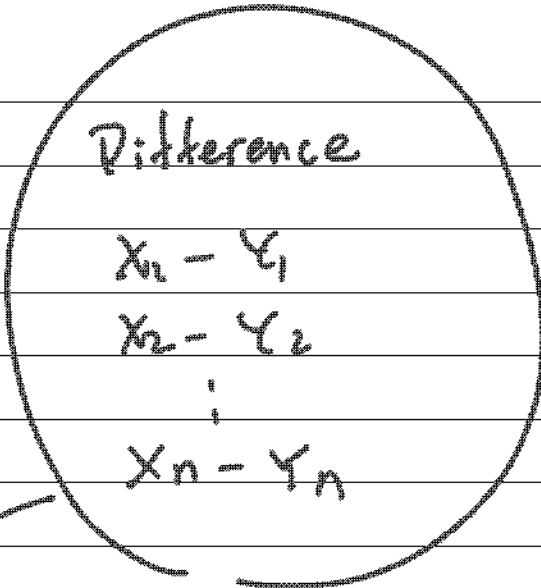
When the two variables are paired, we cannot use general or pooled t-test.

↓  
We need to use paired t-test

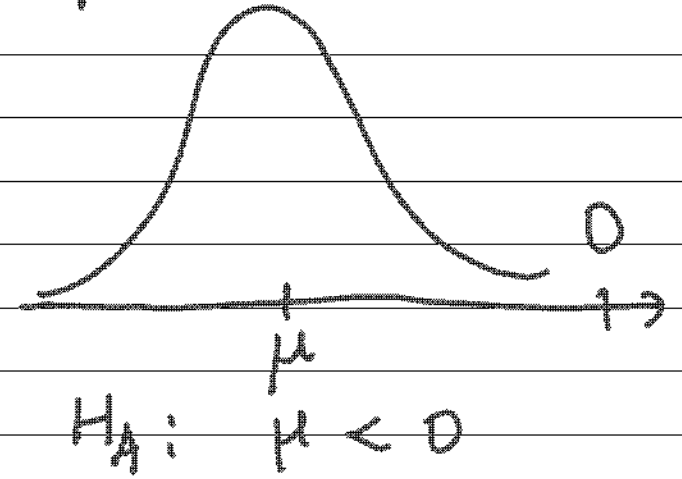
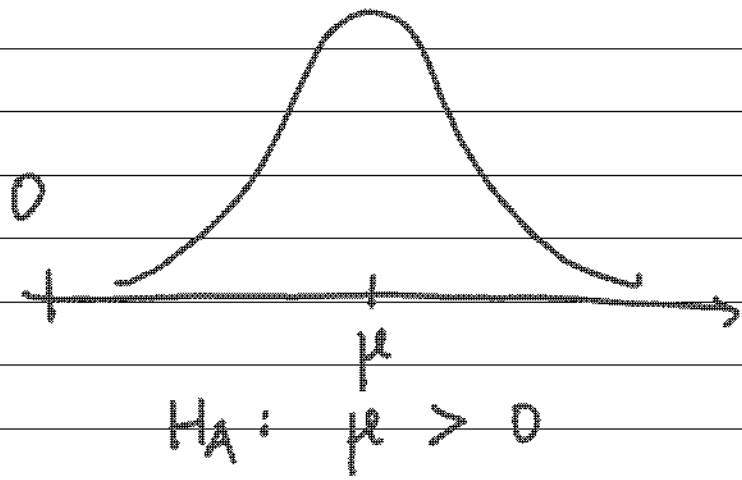
# Construction of difference variable

← Paired ↓

Variable 1	Variable 2
$X_1$	$Y_1$
$X_2$	$Y_2$
$\vdots$	$\vdots$
$X_n$	$Y_n$



Distribution of the difference  $\mu = \mu_1 - \mu_2$



3. Answer the following questions regarding the study in Exercise 6.28. (Data file: Chapter 6/ex6-28.csv)

(a) Present a short description of the study and the data, including summary statistics for each variable.

A study compares the yields for two new varieties of corn, and each variety was randomly assigned to a different 1-acre plot on each of seven farms. The 1-acre plots were planted, and the corn was harvested at maturity.

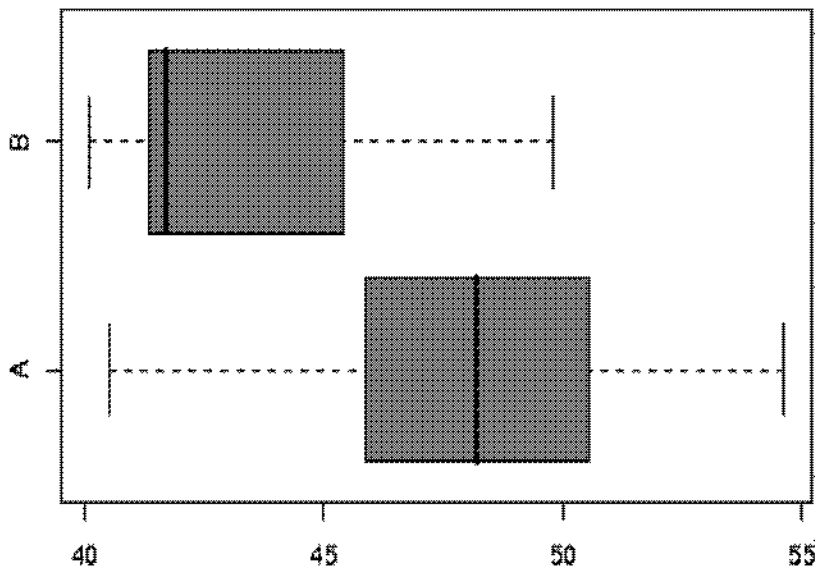
Variables:

A: The yield (bushels of corn) of variety A

B: The yield (bushels of corn) of variety B

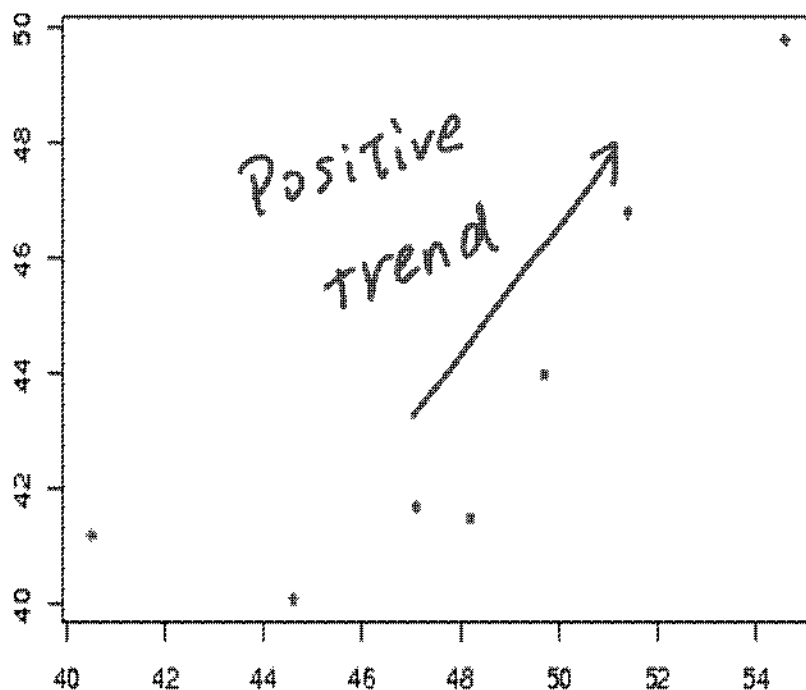
Variable	Mean	S.D	L. Quartile	Median	U. Quartile
A	48.01429	4.594717	45.85	48.2	50.55
B	43.58571	3.532435	41.35	41.7	45.4

(b) Present the comparison of the two groups in boxplot and scatter plot. Then comment on the data based on the visualization.



Variety A shows a higher yield compared to B.

Scatter plot for Variety A against B



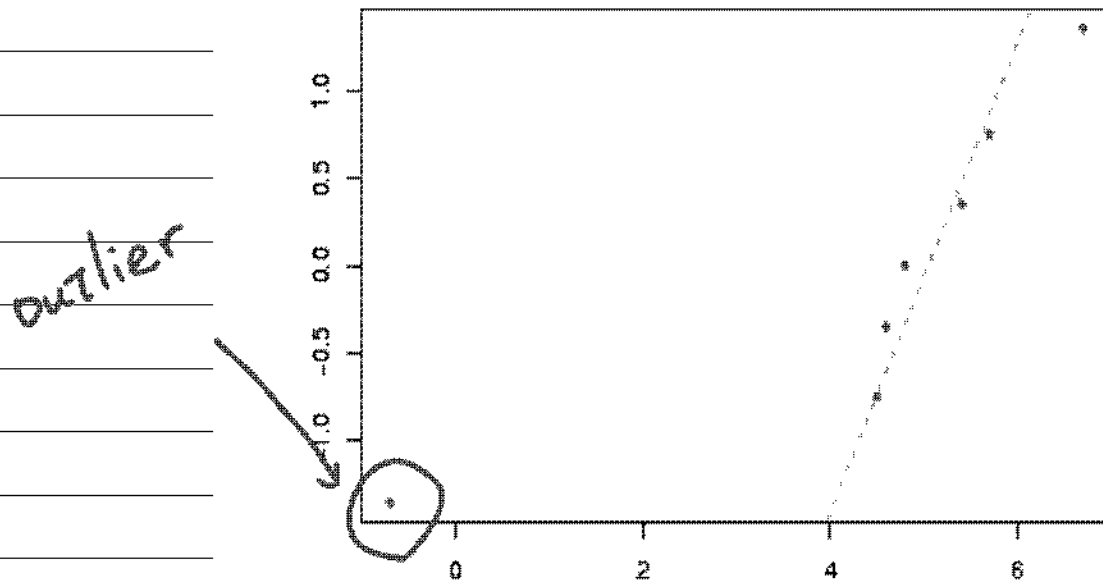
Scatter plot displays a positive trend. It indicates a correlation between two variables. Thus, they are dependent.

(c) Test whether there is a difference in the mean yields for the two varieties of corn.

Null hypothesis is that the mean difference is equal to zero.

Alternative hypothesis is that the mean difference is not equal to zero.

QQ plot for the difference



There is an outlier when we see the QQ plot for the normality.

The sample size is very small, and both parametric and nonparametric test are performed for the analysis,

Result of statistical tests

Procedure	p.value
t-test	0.002689
Wilcoxon	0.03125

(d) Estimate the difference in the mean yields of the two varieties.

The sample mean of difference is 4.428571, and the 95% confidence interval is (2.220717, 6.636426).

(e) State your conclusion of the study.

The result of t-test is highly significant, but the Wilcoxon signed-rank test shows moderate evidence for the difference. There is evidence for the difference. The yield from variety A is higher than variety B, and the average difference is 4.4 bushels per acre.