Pooled t-test

Note Title 11/17/2008

If we can assume that

$\boxed{\sigma_1^2 = \sigma_2^2}$ (σ_1^2 and σ_2^2 are variances)

we may use pooled t-test.

Advantage of pooled t-test

Pooled t-test provides a higher power for the test.

Test for comparison of two groups

Primary objective

$H_0: \mu_1 = \mu_2$ vs. $\boxed{H_A: \mu_1 \neq \mu_2}$

In general it performs better (that is, a higher power) than general t-test if $\boxed{\sigma_1^2 = \sigma_2^2}$.

It will become a secondary objective

3. Answer the following questions regarding the study in Exercise 6.16. (Data file: Chapter 6/ex6-16.csv)

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

The end-of-year bonuses for 24 female and 36 male managers were studied, expressed as a percentage of yearly salary.

Variables:

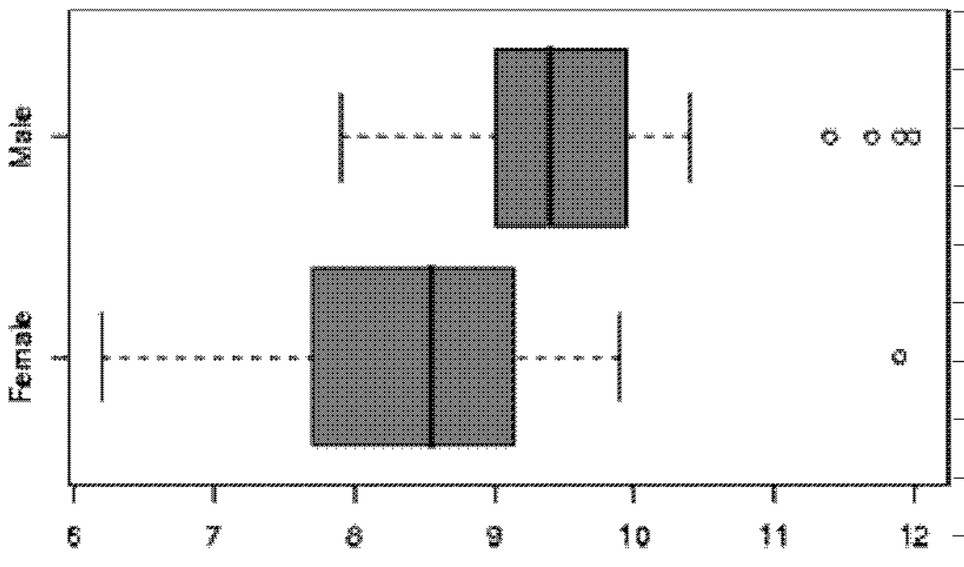
Female: The end-of-year bonuses for female managers (in percentage).

Male: The end-of-year bonuses for male managers (in percentage).

Summary statistics:

Variable	Mean	S.D	L.Quartile	Median	U.Quartile
Female	8.533333	1.188959	7.7	8.55	9.125
Male	9.683333	1.003805	9	9.4	9.925

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



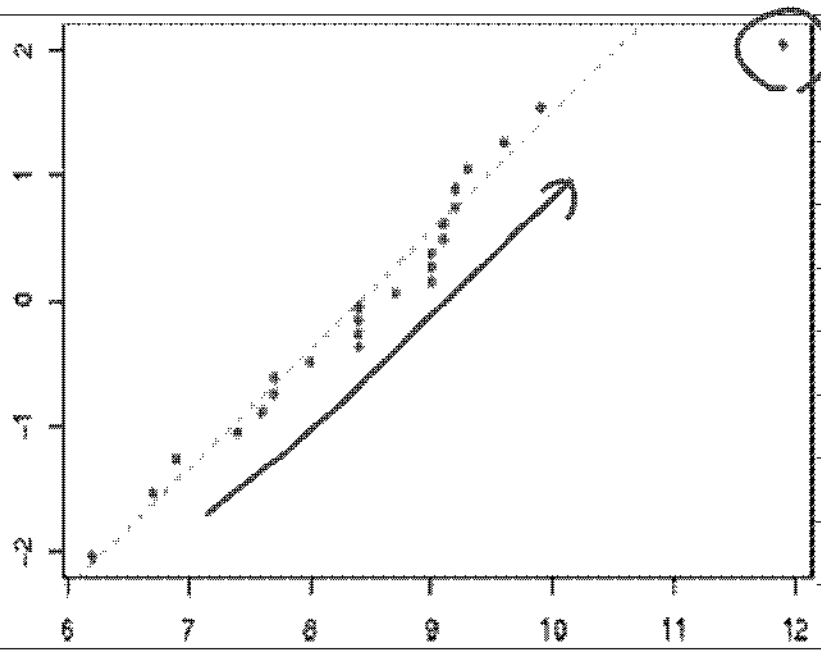
(c) Do the date indicate any difference between female and male managers in the bonus percentage of yearly salary? Construct the null and the alternative hypothesis for the test.

Null hypothesis is that the bonus percentage of yearly salary for female and male managers are equal.

Alternative hypothesis is that the bonus percentage of yearly salary for female and male managers are different.

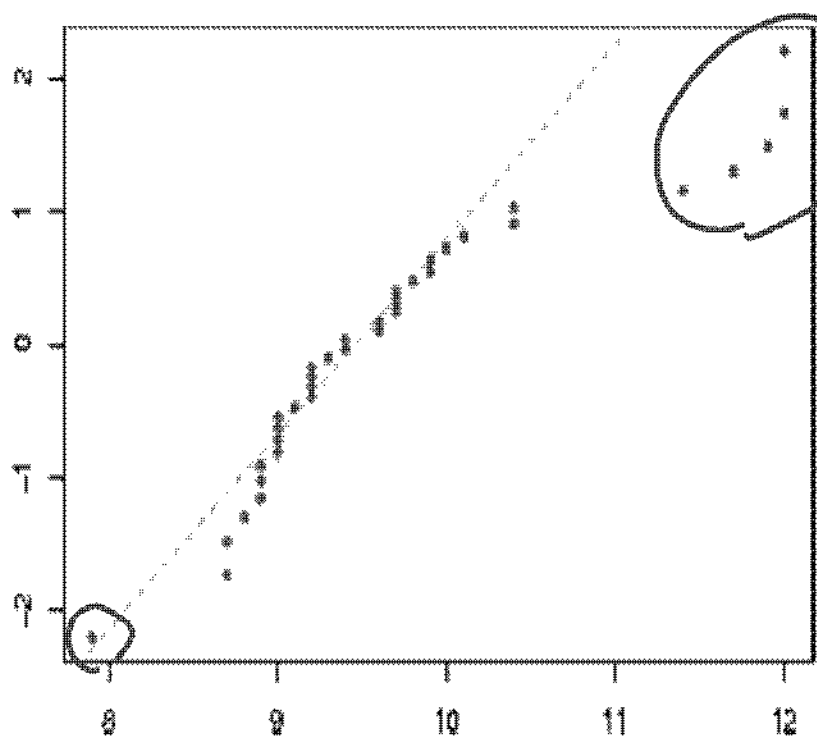
(d) Construct QQ normal plot for each group. Is there reason to think the normality assumption for the t-test have been violated?

QQ plot for the group of female managers



potential outlier

QQ plot for the group of male managers



They are off the line

Data do not follow a normal distribution.

However, we have the sample size 36, And we do not need the normality

How to assess the equal variances

σ_1 = standard deviation of group 1

σ_2 = " of group 2

σ_1^2 = variance of group 1

σ_2^2 = " of group 2

Claim: $\sigma_1^2 = \sigma_2^2$

IDEA: Use F-test to compare variances:

$H_0: \sigma_1^2 = \sigma_2^2$ vs. $H_A: \sigma_1^2 \neq \sigma_2^2$

In this context we "reluctantly" accept H_0 when we fail to reject H_0. Thus, we are justified to use pooled t-test.

Other informal way to check equal variances

Look at the ratio of sample variances

If $\frac{1}{2} < \frac{s_1^2}{s_2^2} < 2$ then we can assume $\sigma_1 = \sigma_2$

(e) Discuss the choice of general and pooled t-test.

We can assume that the two variances are equal by means of another test. So that we can choose the pooled t-test.

(f) Does the conclusion depend on which test is used?

test	t.statistic	p.value
general	-3.90126	3.27E-04
pooled	-4.03675	1.61E-04

$3.27 \times 10^{-4} = 0.000327$

In both tests the result is highly significant. There is evidence to show a significant difference between female and male managers' bonuses. Note that the p-value in the pooled test is smaller.

99% CI for the difference is (-1.908725, -0.3912754). It indicates that the bonus percentage for female managers is significantly smaller than that of male managers.