

# Statistical analysis

Note Title

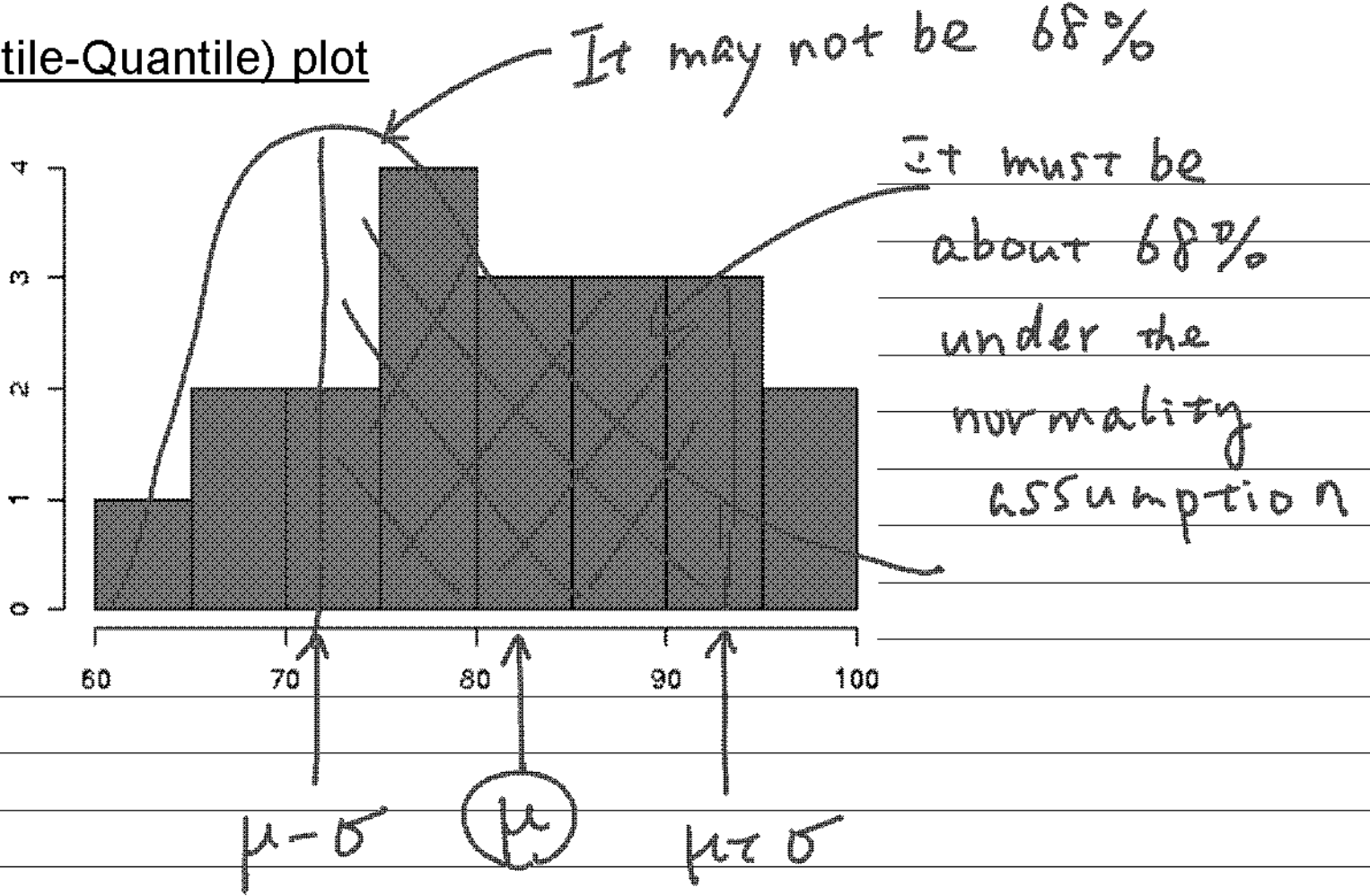
10/29/2008

1. State formal hypothesis: Null and alternative hypothesis.
2. Discuss the validity of normality assumption by using QQ plot.
3. Describe the result of hypothesis test by using p-value.

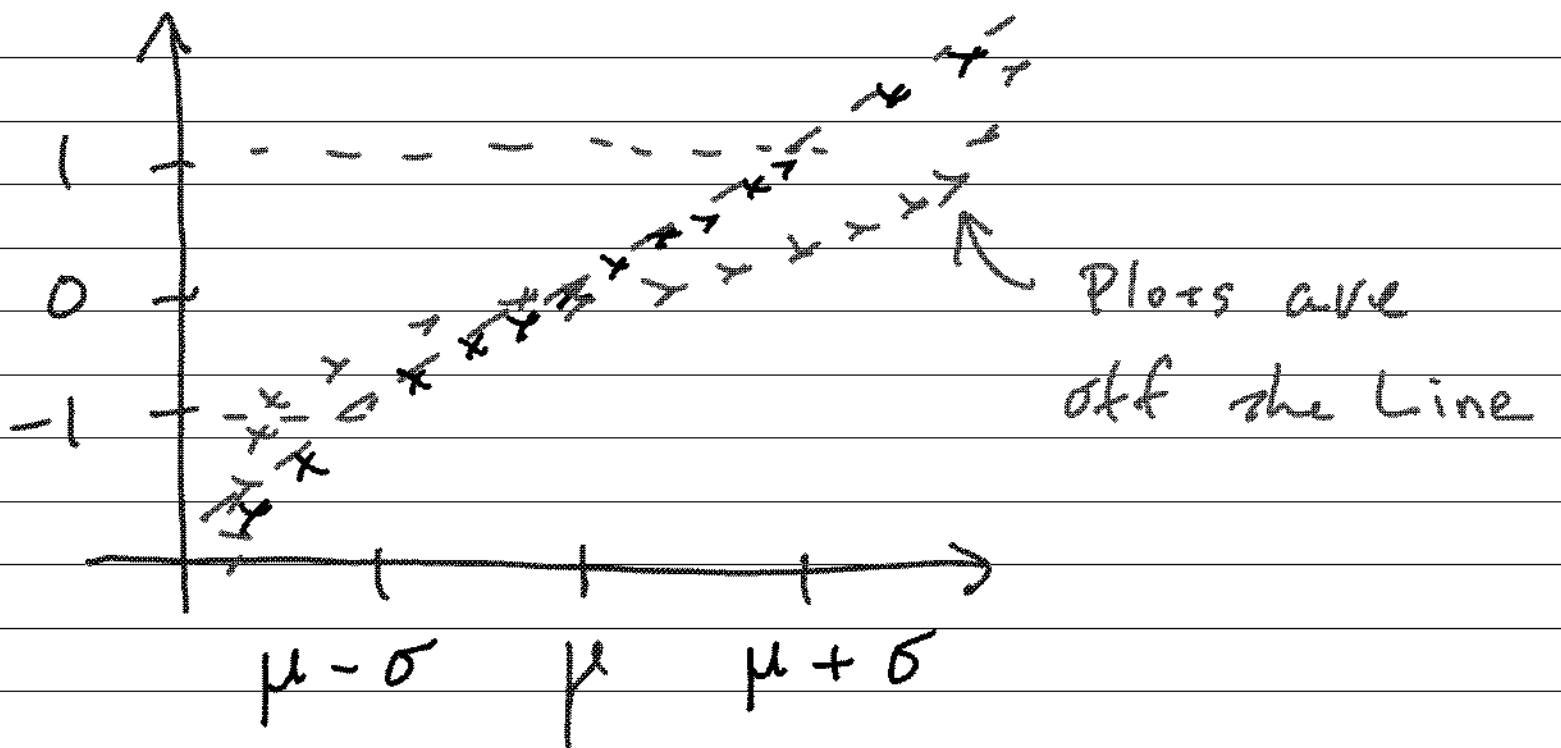
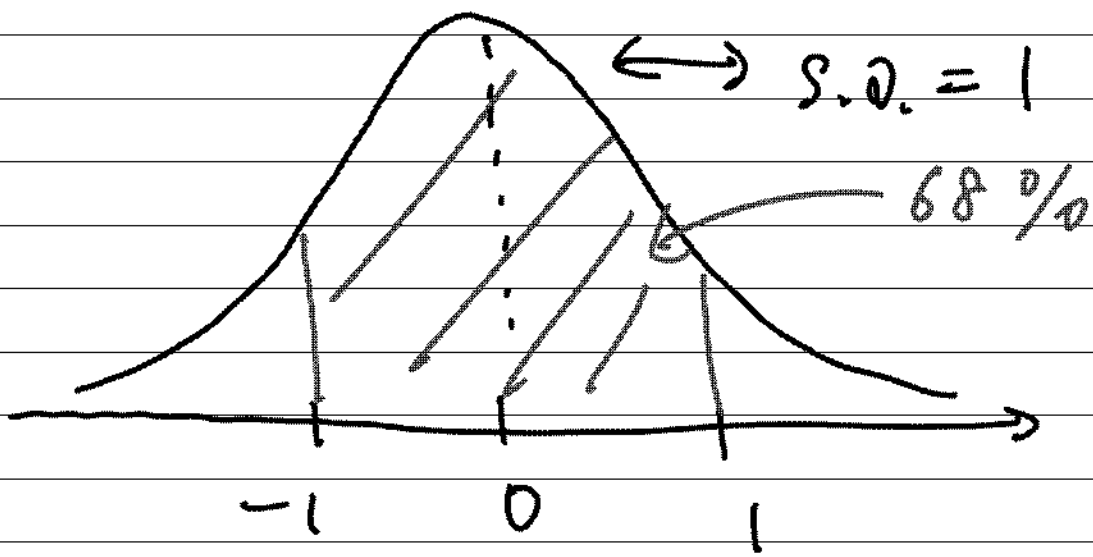
## Logic behind normality assumption

If the size  $n$  of sample data is small ( $n < 30$ ), the t-test procedure requires that the data be approximately normally distributed; otherwise, the result of t-test may not be reliable.

QQ (Quantile-Quantile) plot



Standard normal distribution



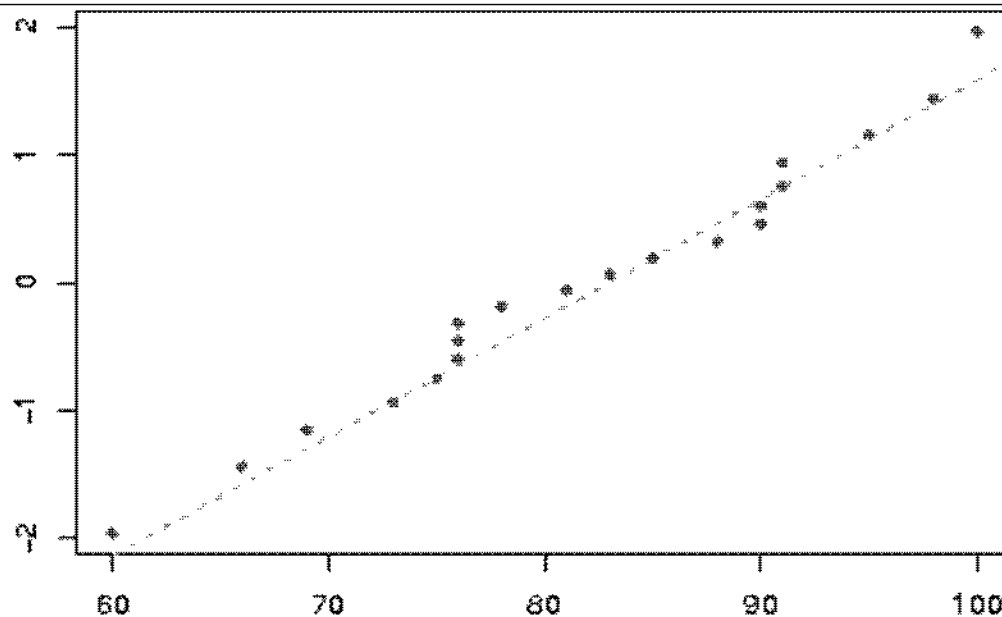
1. Answer the following questions regarding the study in Exercise 5.52, and report your findings. (Data file ex5-52.csv)

(c) Researchers want to find statistical evidence that the mean comprehension for all fourth graders is greater than 80? Construct the null and the alternative hypothesis for the test.

The null hypothesis is that the mean comprehension for all fourth graders is less than or equal to 80.

The alternative hypothesis is that the mean comprehension for all fourth graders is greater than 80.

(d) Construct a QQ normal plot for the comprehension scores, and assess the normality of the sample distribution.



QQ plots lie closer to the straight line, indicating that the data come from a distribution close to normal.

## Significance level and critical region

1. Under the null hypothesis, it is "unlikely" that the  $t$ -statistic lies in the *critical region*.
2. If so, it suggests significant evidence against the null hypothesis

" $T > c$ "  $\Rightarrow$  We reject  $H_0$ ; otherwise, we don't reject.  
 $\uparrow$

Significance level  $\alpha$  = Probability of  
type 2 error

## P-value and decision regarding null hypothesis

P-value is the smallest choice of significance level so that we can reject  $H_0$ .

$$p\text{-value} = 0.205$$

Idea: You have your choice of significance level  $\alpha^*$  ( $0.05$ )

Use the p-value to determine whether  $H_0$  is rejected or not.

If  $p\text{-value} > \alpha^*$ , do not reject  $H_0$

If  $\alpha^* \geq p\text{-value}$ , reject  $H_0$ .

(e) Calculate the test statistic and the p-value of the test.

Choose significance level  $\alpha = 0.05$

t. statistic	p. value
0.8426653	0.204949

$$H_0: \mu \leq 80$$

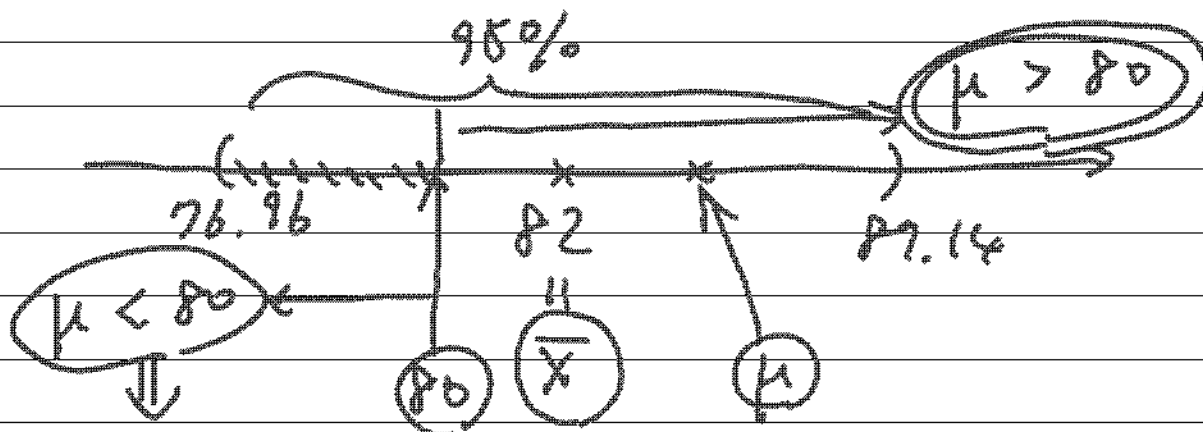
$$H_A: \mu > 80$$

$\alpha = 0.05$  is smaller than p-value 0.205

$H_0$  cannot be rejected because  $p\text{-value} > \alpha$ .

(f) Construct a confidence interval for the average comprehension score.

95% C.I. is (76.96, 87.14)



$H_0$  cannot be rejected  $\rightarrow$  We cannot find evidence to support  $\mu > 80$ .

(g) Choose the significance level for the test, and report your findings.

Conclusion: we choose  $\alpha = 0.05$

So that  $p\text{-value} = 0.205 > \alpha = 0.05$  and

we cannot reject  $H_0$ . There is not sufficient evidence to support that the average score is larger than 80.

## Interpretation of the result and p-value

Choose your own  $\alpha = 0.01, 0.05$  or  $0.1$

Obtain the p-value

If the p-value is smaller than  $\alpha$  (p-value  $< \alpha$ )  
then  $H_0$  is rejected.  $\Rightarrow$  There is evidence

If p-value  $> \alpha$ , then  $H_0$  cannot be rejected.  
 $\Rightarrow$  There is no evidence