Midterm test : one week today.

TA office hours : TBA today.

Prof. Tabuchi office hours : Cancelled today but will reschedule for Thursday.

STA305/1004-Class 16

Thursday's Class:

- Tutorial class
- ask any questions during class.

Nov. 12, 2019

Today's Class

Assessing significance in unreplicated factorial designs:

- Normal plots
- half-Normal plots
- Lenth's method

ANOVA:

- Multiple comparisons
- ~~Sample size for ANOVA~~

Assessing Significance in Unreplicated Factorial Designs

In replicated designs.

$\frac{\text{fact. est}}{\text{S.e(fact. est)}}$ Compare to a t-dist.

How can significance be assessed in unreplicated factorial designs?

But, in unreplicated designs it's not possible to compute S.e(fact. est.).

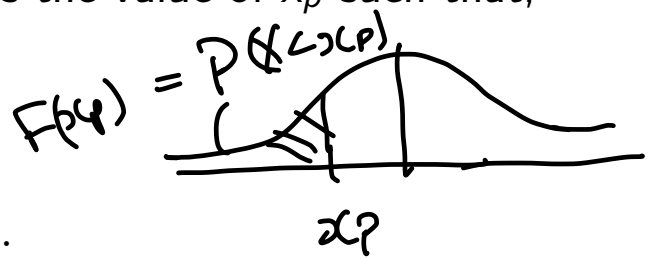
Quantile-Quantile Plots

- Quantile-quantile (Q-Q) plots are useful for comparing distribution functions.
- If X is a continuous random variable with strictly increasing distribution function $F(x)$ then the *pth* quantile of the distribution is the value of x_p such that,

$$F(x_p) = p$$

or

$$x_p = F^{-1}(p).$$



- In a Q-Q plot, the quantiles of one distribution are plotted against another distribution.
- Q-Q plots can be used to investigate if a set of numbers follows a certain distribution.

Quantile-Quantile Plots

- Suppose that we have independent observations $X_1, X_2, ..., X_n$ from a uniform distribution on $[0, 1]$ or Unif[0,1].
- The ordered sample values (also called the order statistics) are the values $X_{(j)}$ such that

$$X_{(1)} < X_{(2)} < \cdots < X_{(n)}$$

.

- It can be shown that

$$E\left(X_{(j)}\right) = \frac{j}{n+1}.$$

- This suggests that if we plot

$$X_{(j)} \quad \text{vs.} \quad \frac{j}{n+1}$$

then if the underlying distribution is Unif[0,1] then the plot should be roughly linear.

Quantile-Quantile Plots

- A continuous random variable with strictly increasing CDF F_X can be transformed to a Unif[0,1] by defining a new random variable $Y = F_X(X)$.
- Suppose that it's hypothesized that X follows a certain distribution function with CDF F.
- Given a sample $X_1, X_2, ..., X_n$ plot

$$F(X_{(k)}) \quad \text{vs.} \quad \frac{k}{n+1}$$

or equivalently

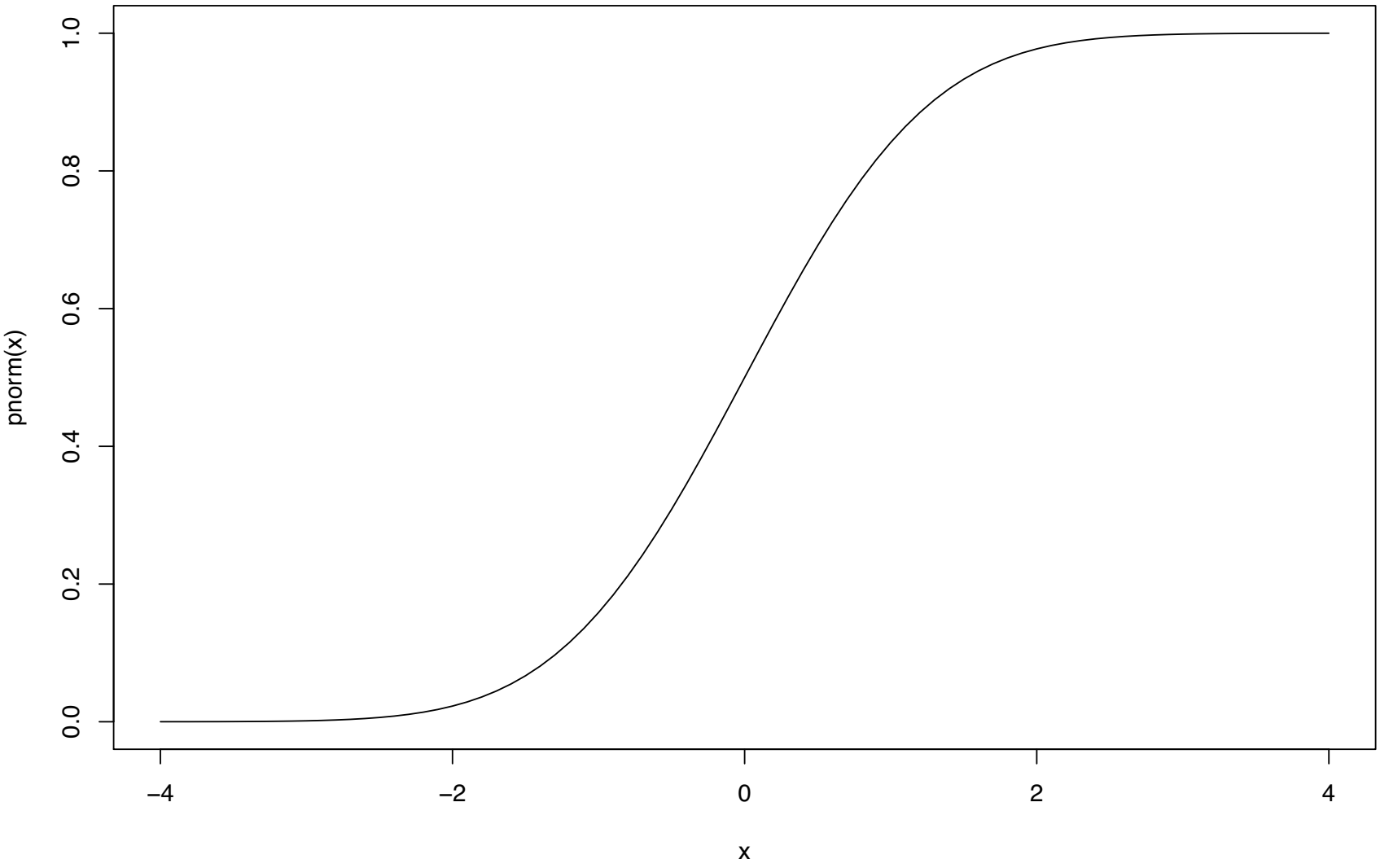
$$X_{(k)} \quad \text{vs.} \quad F^{-1}\left(\frac{k}{n+1}\right)$$

- $X_{(k)}$ can be thought of as empirical quantiles and $F^{-1}\left(\frac{k}{n+1}\right)$ as the hypothesized quantiles.
- The quantile assigned to $X_{(k)}$ is not unique.
- Instead of assigning it $\frac{k}{n+1}$ it is often assigned $\frac{k-0.5}{n}$. In practice it makes little difference which definition is used.

Normal Quantile-Quantile Plots

The cumulative distribution function (CDF) of the normal has an S-shape.

```
x <- seq(-4,4,by=0.1)
plot(x,pnorm(x),type="l")
```



Normal Quantile-Quantile Plots

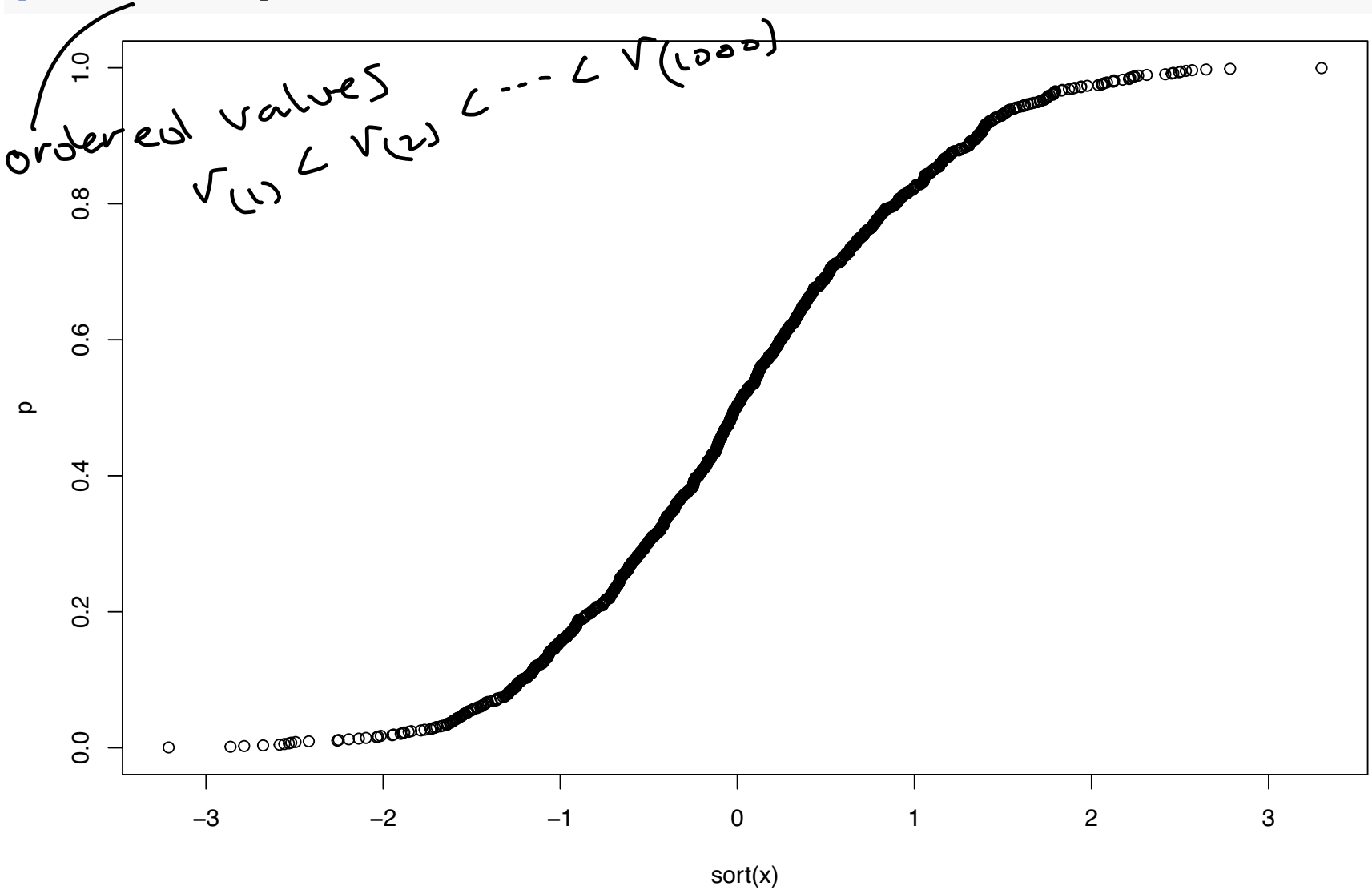
The normality of a set of data can be assessed by the following method.

- Let $r_{(1)} < ... < r_{(N)}$ denote the ordered values of $r_1, ..., r_N$.
- A test of normality for a set of data is to plot the ordered values $r_{(i)}$ of the data versus $p_i = (i - 0.5)/N$. —— what position is the value $r_{(i)}$ in the Seq.
- If the plot has the same S-shape as the normal CDF then this is evidence that the data come from a normal distribution.

Normal Quantile-Quantile Plots

- A plot of $r_{(i)}$ vs. $p_i = (i - 0.5)/N, i = 1, ..., N$ for a random sample of 1000 simulated from a $N(0, 1)$.

```
N <- 1000;x <- rnorm(N);p <- ((1:N)-0.5)/N
plot(sort(x),p)
```



Normal Quantile-Quantile Plots

- It can be shown that $\Phi(r_i)$ has a uniform distribution on $[0, 1]$.
- This implies that $E(\Phi(r_{(i)})) = i/(N+1)$ (this is the expected value of the jth order statistic from a uniform distribution over $[0, 1]$.
- This implies that the N points $(p_i, \Phi(r_{(i)}))$ should fall on a straight line.
- Now apply the Φ^{-1} transformation to the horizontal and vertical scales. The N points

$$\left(\Phi^{-1}(p_i), r_{(i)}\right),$$

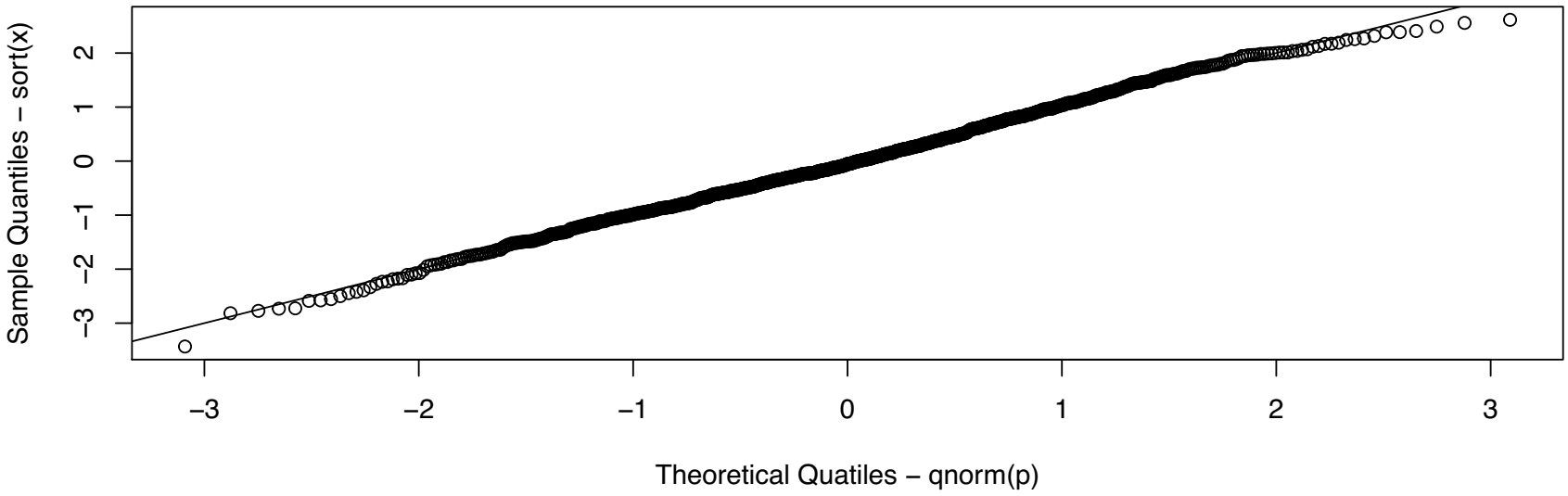
form the normal probability plot of $r_1, ..., r_N$.

- If $r_1, ..., r_N$ are generated from a normal distribution then a plot of the points $\left(\Phi^{-1}(p_i), r_{(i)}\right), i = 1, ..., N$ should be a straight line.

Normal Quantile-Quantile Plots

In R `qnorm()` is Φ^{-1}.

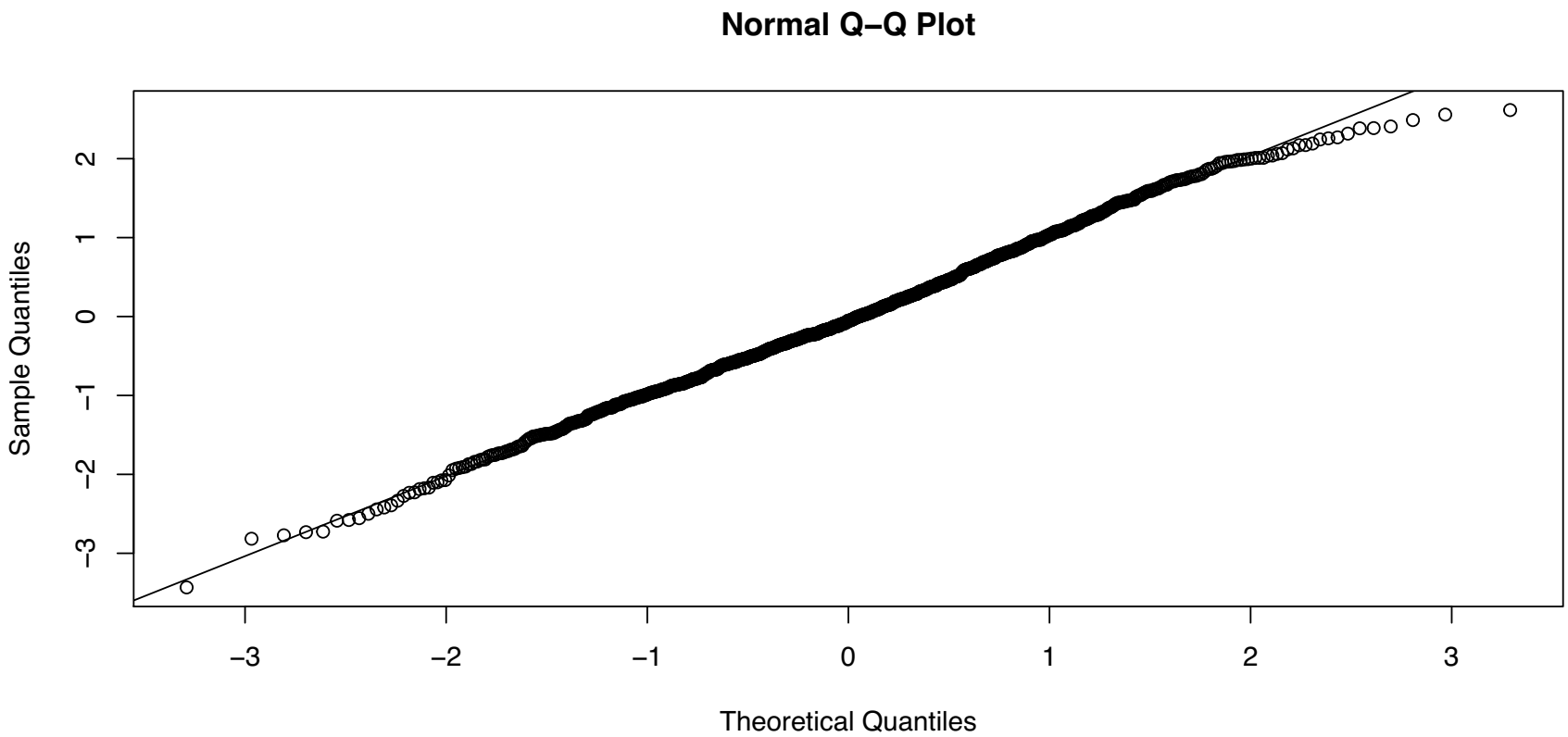
```
set.seed(2503)
N <- 1000
x <- rnorm(N)
p <- (1:N) / (N + 1) # the vector i/(N+1), i = 1,...,N.
plot(qnorm(p), sort(x),
     ylab = "Sample Quantiles - sort(x)",
     xlab = "Theoretical Quatiles - qnorm(p)")
abline(a = 0, b = 1) # add the line y=x
```



Normal Quantile-Quantile Plots

We usually use the built in function `qqnorm()` (and `qqline()` to add a straight line for comparison) to generate normal Q-Q plots. Note that R uses a slightly more general version of quantile ($p_i = (1-a)/(N+(1-a)-a)$, where $a = 3/8$, if $N \leq 10$, $a = 1/2$, if $N > 10$.

```
qqnorm(x);qqline(x)
```



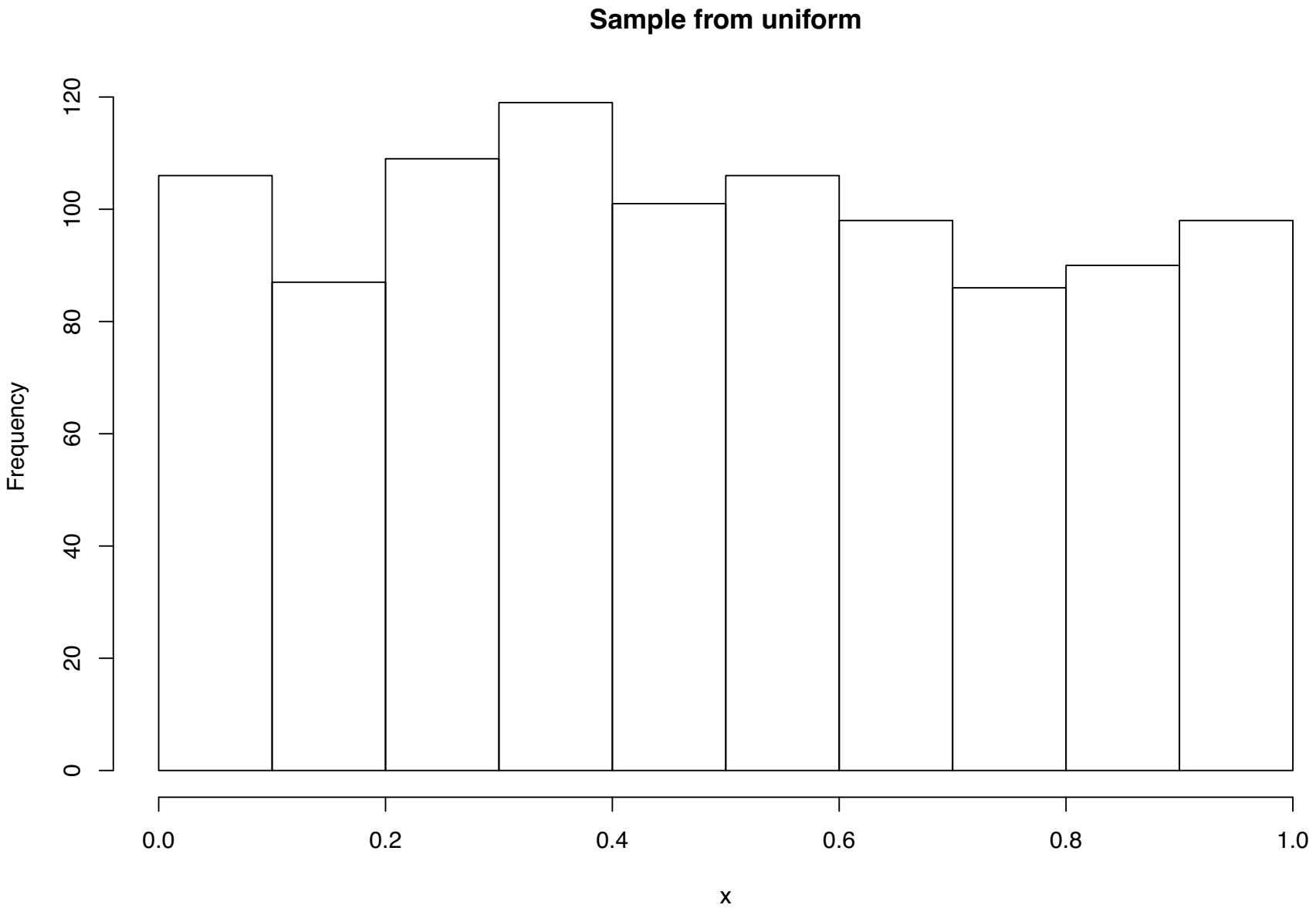
Normal Quantile-Quantile Plots

A marked (systematic) deviation of the plot from the straight line would indicate that:

1. The normality assumption does not hold.
2. The variance is not constant.

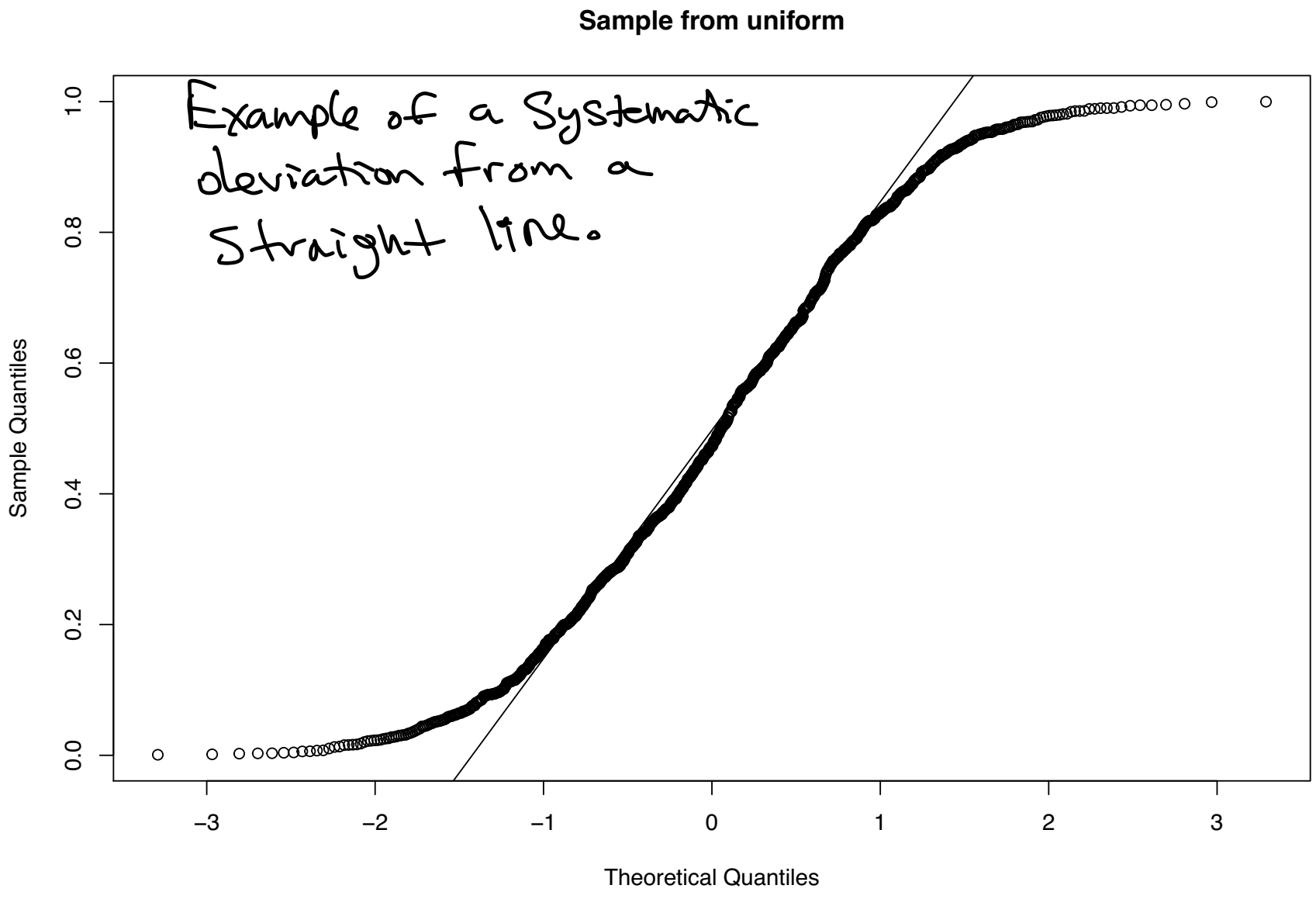
Normal Quantile-Quantile Plots

```
x <- runif(1000)
hist(x,main = "Sample from uniform")
```



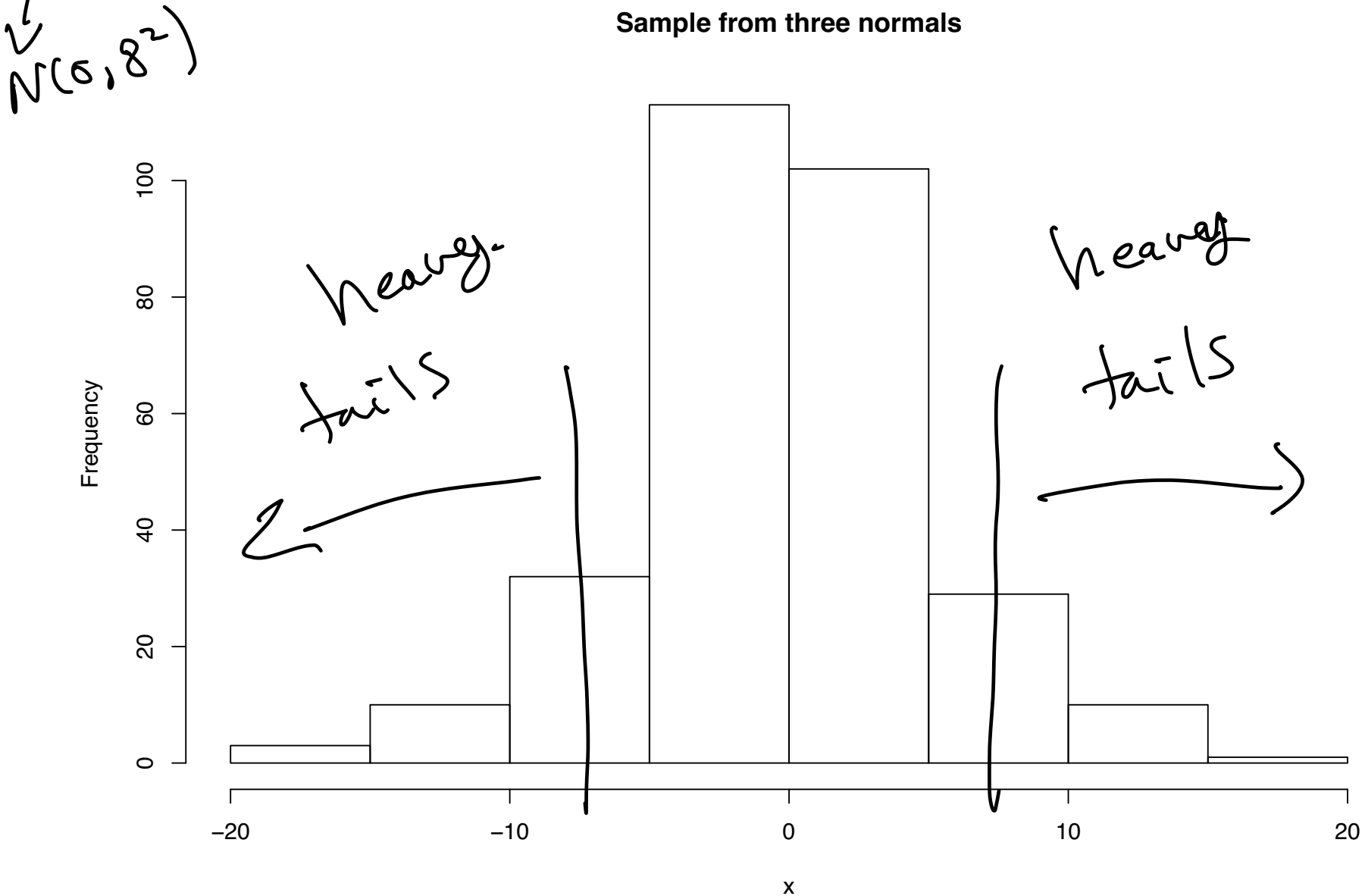
Normal Quantile-Quantile Plots

```
qqnorm(x,main = "Sample from uniform");qqline(x)
```



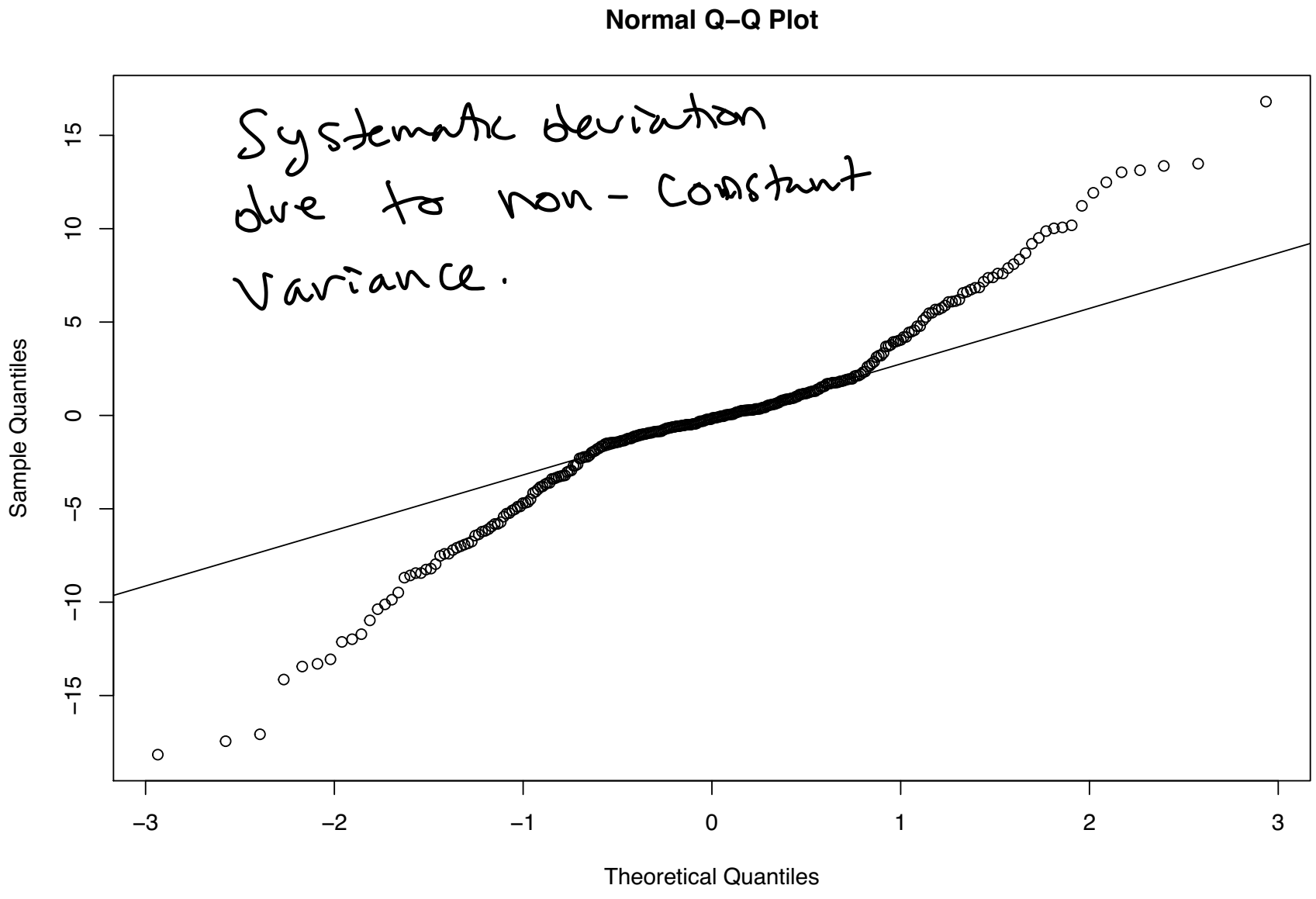
Normal Quantile-Quantile Plots

```
x1 <- rnorm(100,mean = 0,sd = 1);x2 <- rnorm(100,mean = 0,sd = 5)
x3 <- rnorm(100,mean = 0,sd = 8); x <- c(x1,x2,x3)
hist(x,main = "Sample from three normals")
```



Normal Quantile-Quantile Plots

```
qqnorm(x);qqline(x)
```



Normal plots in factorial experiments

2^2 factorial design

Factorial effects:

$$A: \left(\frac{y_2+y_4}{2}\right) - \left(\frac{y_1+y_3}{2}\right) = \hat{\theta}_1$$

$$B: \left(\frac{y_3+y_4}{2}\right) - \left(\frac{y_1+y_2}{2}\right) = \hat{\theta}_2$$

a	b	ab	
−	−	+	y_1
+	−	−	y_2
−	+	−	y_3
+	+	+	y_4

unreplicated

- A major application is in factorial designs where the $r(i)$ are replaced by ordered factorial effects.
- Let $\hat{\theta}_{(1)} < \hat{\theta}_{(2)} < \cdots < \hat{\theta}_{(N)}$ be N ordered factorial estimates.
- If we plot

$$\hat{\theta}_{(i)} \text{ vs. } \Phi^{-1}(p_i).\ i = 1, ..., N.$$

then factorial effects $\hat{\theta}_i$ that are close to 0 will fall along a straight line. Therefore, points that fall off the straight line will be declared significant.

$$AB: \left(\frac{y_1+y_4}{2}\right) - \left(\frac{y_2+y_3}{2}\right) = \hat{\theta}_3$$

Normal plots in factorial experiments

The rationale is as follows:

1. Assume that the estimated effects $\hat{\theta}_i$ are $N(\theta, \sigma^2)$ (estimated effects involve averaging of N observations and CLT ensures averages are nearly normal for N as small as 8).
2. If $H_0 : \theta_i = 0$, $i = 1, ..., N$ is true then all the estimated effects will be zero.
3. The resulting normal probability plot of the estimated effects will be a straight line.
4. Therefore, the normal probability plot is testing whether all of the estimated effects have the same distribution (i.e. same means). $N(0, \sigma^2)$

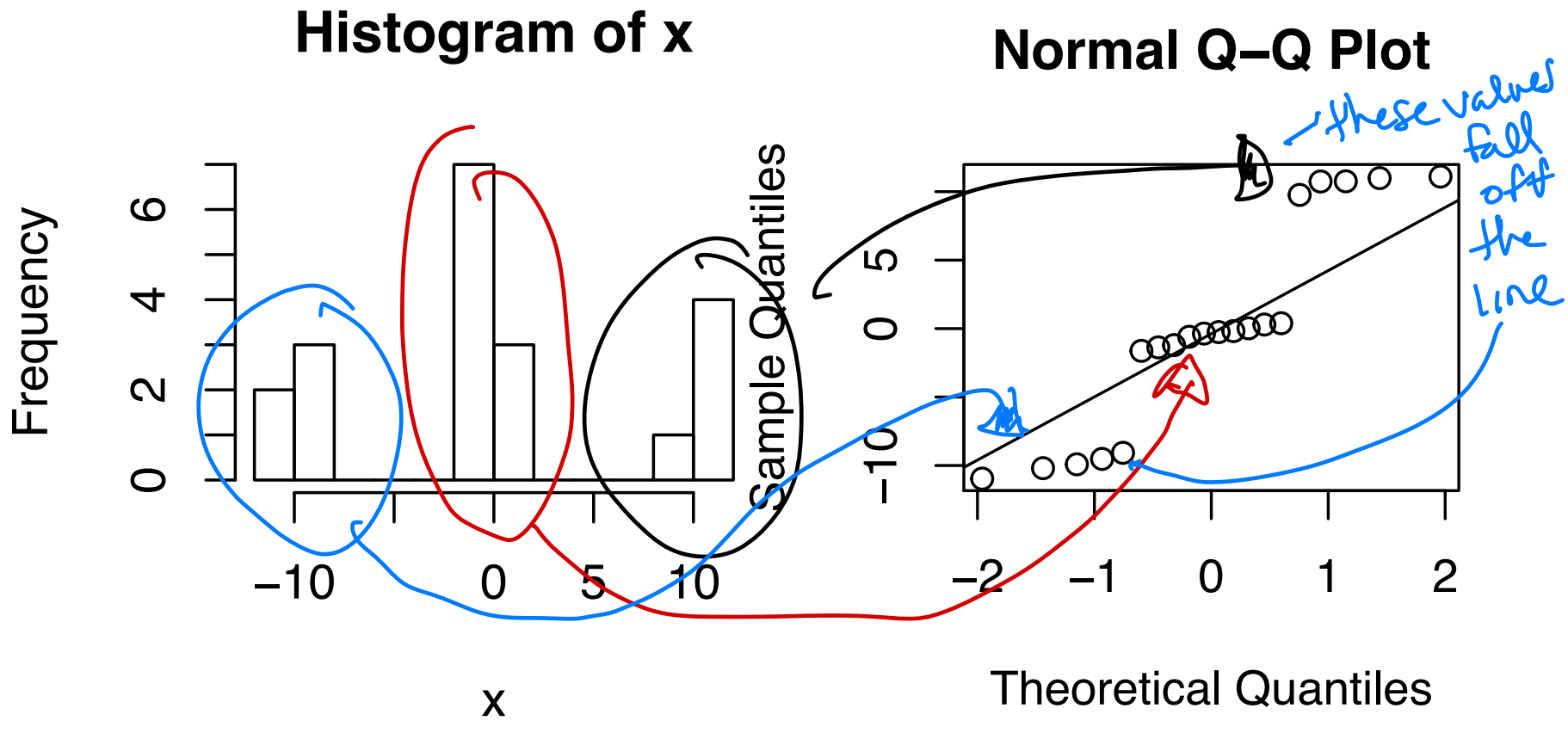
   - When some of the effects are nonzero the corresponding estimated effects will tend to be larger and fall off the straight line.

factorial effects

Normal plots in factorial experiments

Positive effects fall above the line and negative effects fall below the line.

```
set.seed(10)
x1 <- rnorm(10,0,1); x2 <- rnorm(5,10,1); x3 <- rnorm(5,-10,1)
x <- c(x1,x2,x3)
hist(x, breaks = 10)
qqnorm(x); qqline(x)
```



Example - 2^3 design for studying a chemical reaction

A process development experiment studied four factors in a 2^4 factorial design.

- amount of catalyst charge **1**,
- temperature **2**,
- pressure **3**,
- concentration of one of the reactants **4**.
- The response y is the percent conversion at each of the 16 run conditions. The design is shown below.

Example - 2^4 design for studying a chemical reaction

x1	x2	x3	x4	conversion
-1	-1	-1	-1	70
1	-1	-1	-1	60
-1	1	-1	-1	89
1	1	-1	-1	81
-1	-1	1	-1	69
1	-1	1	-1	62
-1	1	1	-1	88
1	1	1	-1	81
-1	-1	-1	1	60
1	-1	-1	1	49
-1	1	-1	1	88
1	1	-1	1	82
-1	-1	1	1	60
1	-1	1	1	52
-1	1	1	1	86
1	1	1	1	79

The design is not replicated so it's not possible to estimate the standard errors of the factorial effects.

Example - 2^4 design for studying a chemical reaction

Remember that $2 \times \hat{\beta}_i$ = factorial effect that β_i corresponds to.

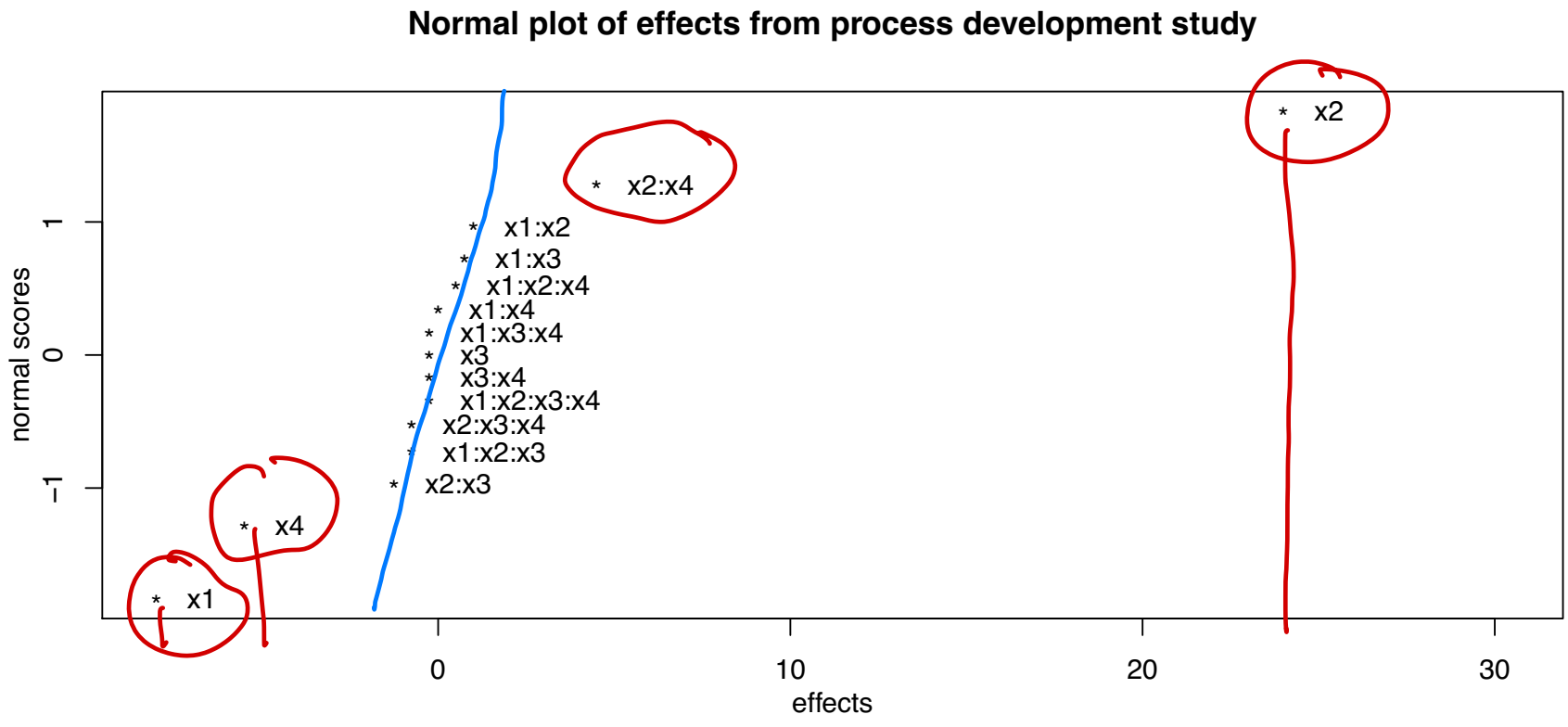
```
fact1 <- lm(conversion~x1*x2*x3*x4,data=tab0510a)
round(2*fact1$coefficients,2)
```

```
(Intercept)          x1          x2          x3          x4       x1:x2
     144.50       -8.00       24.00       -0.25       -5.50        1.00
      x1:x3       x2:x3       x1:x4       x2:x4       x3:x4    x1:x2:x3
       0.75       -1.25        0.00        4.50       -0.25       -0.75
   x1:x2:x4    x1:x3:x4    x2:x3:x4 x1:x2:x3:x4
       0.50       -0.25       -0.75       -0.25
```

Example - 2^4 design for studying a chemical reaction

A normal plot of the factorial effects is obtained by using the function `DanielPlot()` in the `FrF2` library.

```
library(FrF2)
DanielPlot(fact1, autolab=F,
           main="Normal plot of effects from process development study")
```



Example 1

Which effects are not explained by chance?

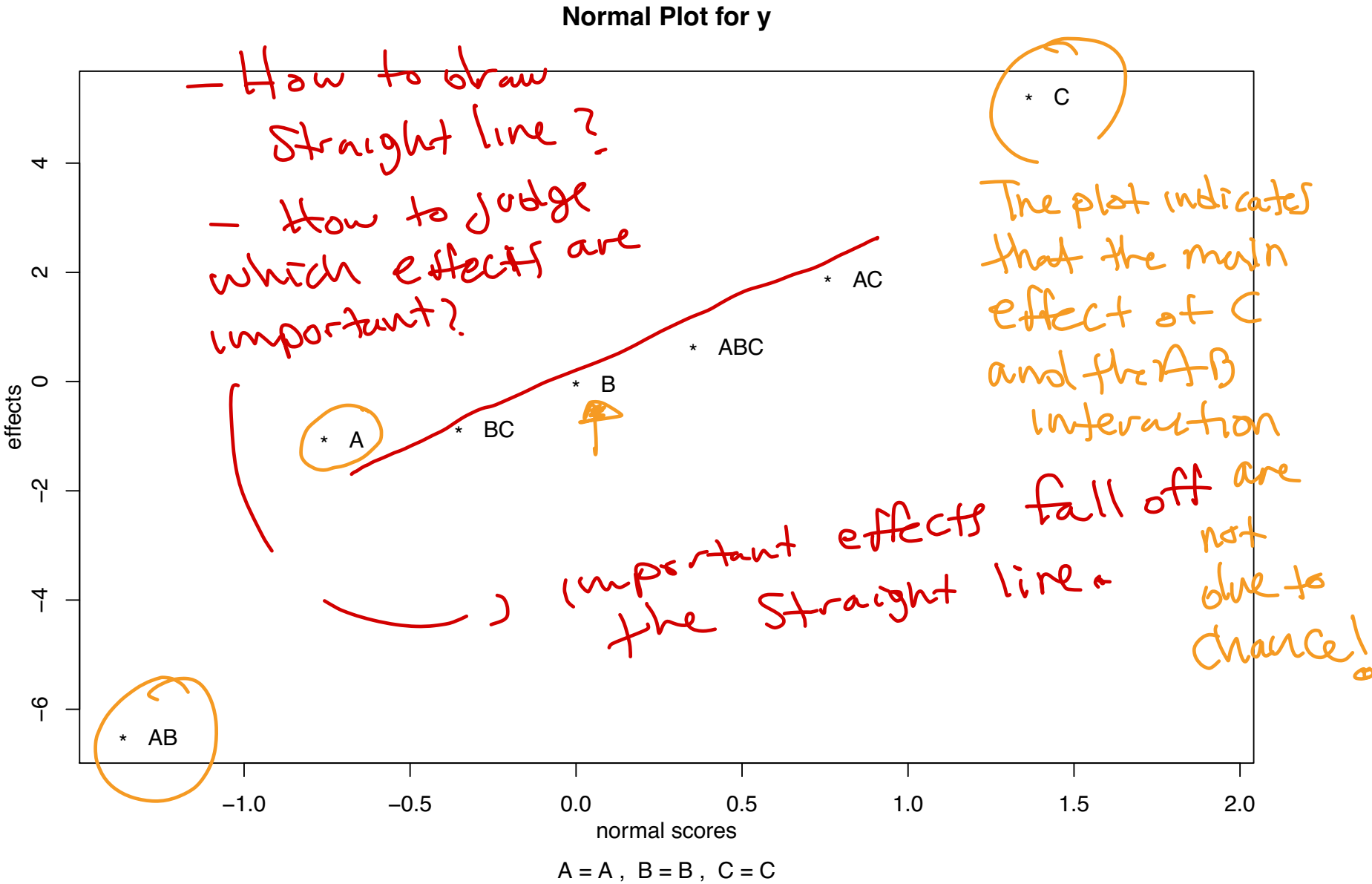
2^3 design. – 3 main effects.
– 3 2-way interactions.
– 1 3 way interaction.

```
##
## Call:
## lm.default(formula = y ~ A * B * C, data = dat)
##
## Residuals:
## ALL 8 residuals are 0: no residual degrees of freedom!
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.90361         NA      NA       NA
## A1          -0.52770         NA      NA       NA
## B1          -0.01836         NA      NA       NA
## C1           2.60717         NA      NA       NA
## A1:B1       -3.25821         NA      NA       NA
## A1:C1        0.93739         NA      NA       NA
## B1:C1       -0.43695         NA      NA       NA
## A1:B1:C1     0.31787         NA      NA       NA
##
## Residual standard error: NaN on 0 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:    NaN
## F-statistic:   NaN on 7 and 0 DF,  p-value: NA
```

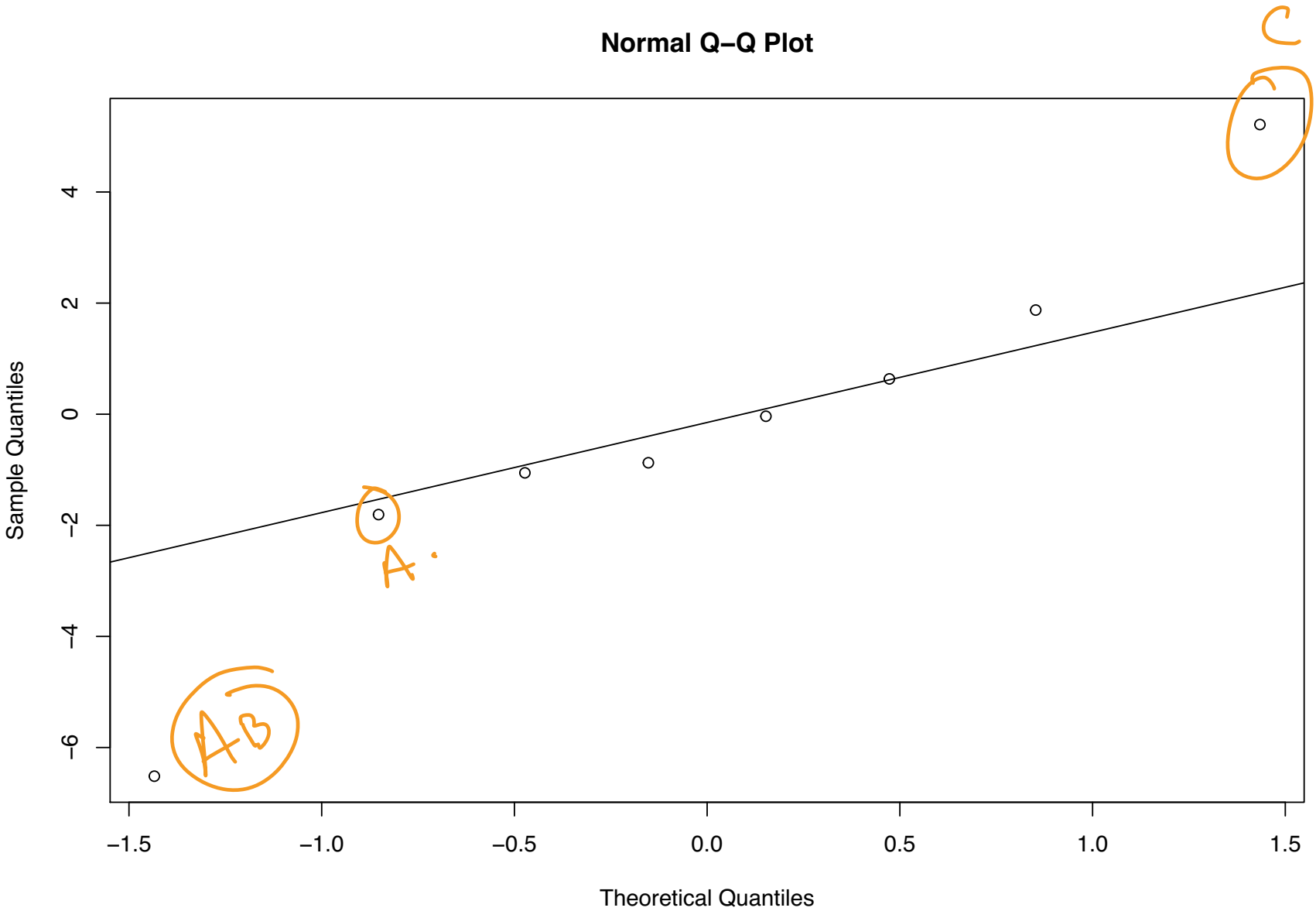
Example 1

Which effects are not explained by chance according to the normal plot?

```
FrF2::DanielPlot(mod1,code=TRUE,autolab=F,datax=F)
```



Example 1


Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles
4
2
0
−2
−4
−6
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

Example 2

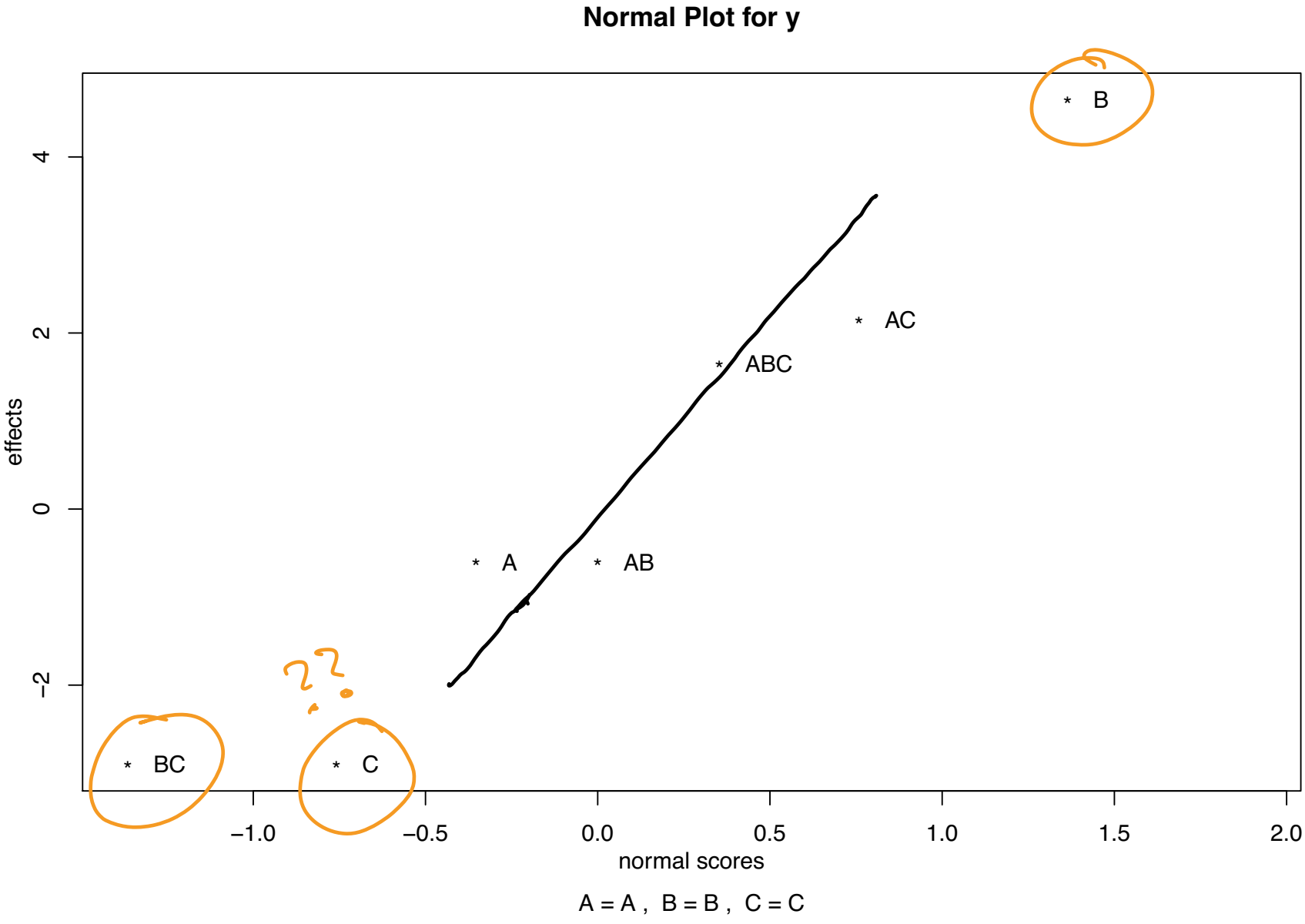
Unreplicated 2^3 design.

Which effects are not explained by chance?

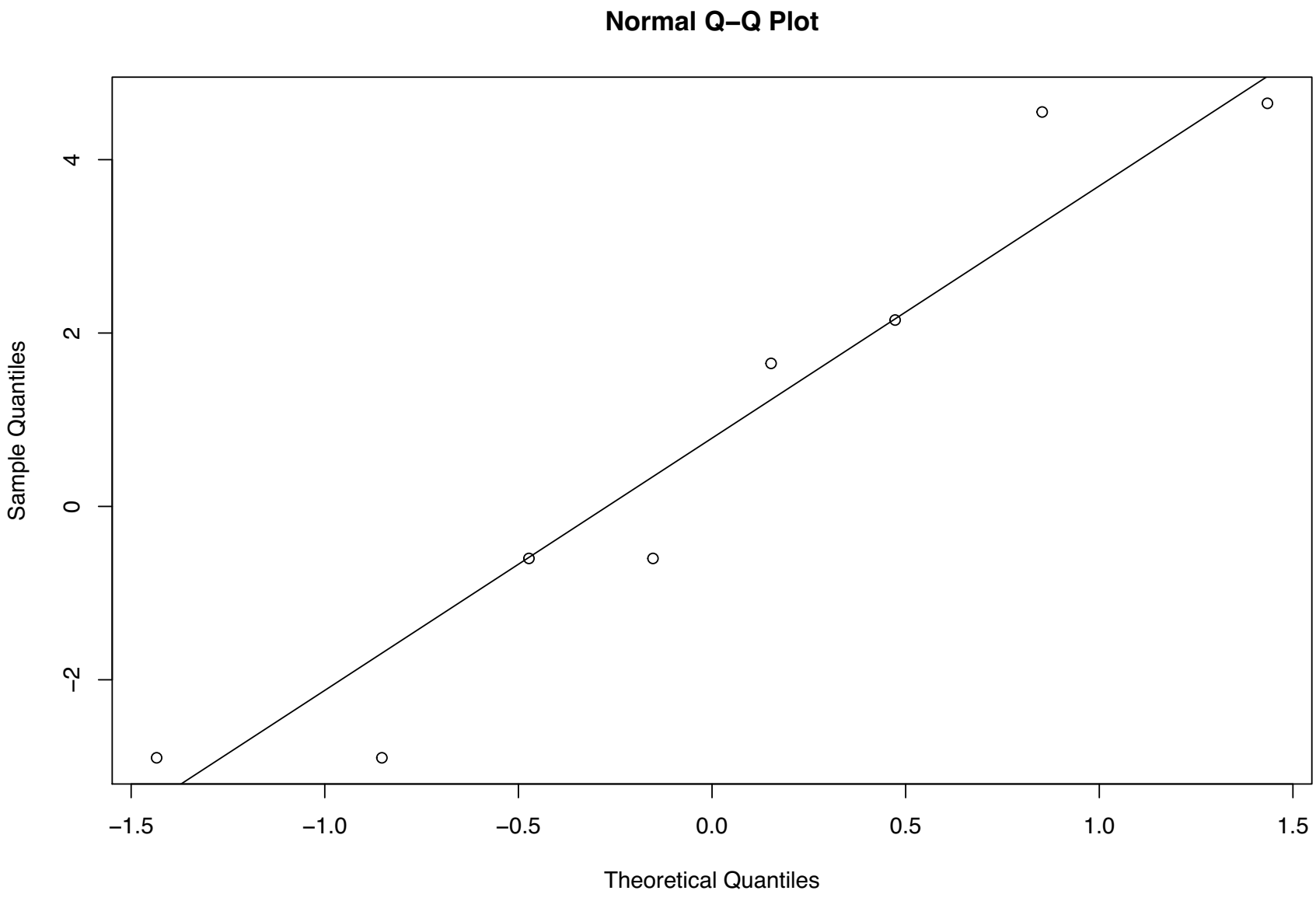
```
##
## Call:
## lm.default(formula = y ~ A * B * C, data = dat)
##
## Residuals:
## ALL 8 residuals are 0: no residual degrees of freedom!
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.275         NA      NA       NA
## A1            -2.150         NA      NA       NA
## B1             0.300         NA      NA       NA
## C1            -0.950         NA      NA       NA
## A1:B1         -0.125         NA      NA       NA
## A1:C1          1.125         NA      NA       NA
## B1:C1         -1.575         NA      NA       NA
## A1:B1:C1       1.500         NA      NA       NA
##
## Residual standard error: NaN on 0 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:    NaN
## F-statistic:   NaN on 7 and 0 DF,  p-value: NA
```

Example 2

Which effects are not explained by chance according to the normal plot?



Example 2


Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles
−2
0
2
4
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

Half-Normal Plots

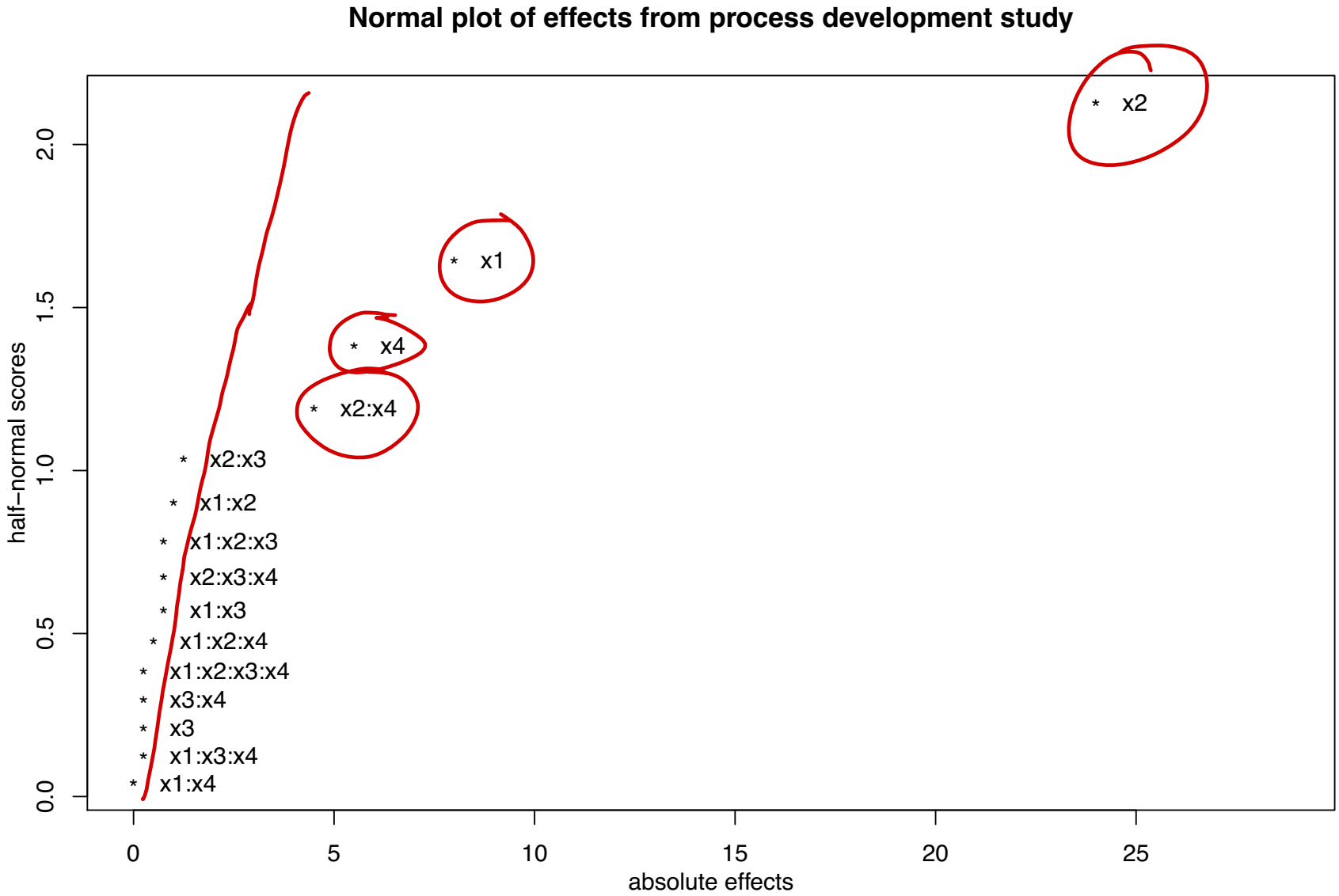
- A related graphical method is called the half-normal probability plot.
- Let
$$\left|\hat{\theta}\right|_{(1)} < \left|\hat{\theta}\right|_{(2)} < \cdots < \left|\hat{\theta}\right|_{(N)}.$$
denote the ordered values of the unsigned factorial effect estimates.
- Plot them against the coordinates based on the half-normal distribution - the absolute value of a normal random variable has a half-normal distribution.
- The half-normal probability plot consists of the points
$$\left|\hat{\theta}\right|_{(i)} \text{ vs. } \Phi^{-1}(0.5 + 0.5[i - 0.5]/N).\ i = 1, ..., N.$$

Half-Normal Plots

- An advantage of this plot is that all the large estimated effects appear in the upper right hand corner and fall above the line.
- The half-normal plot for the effects in the process development example is can be obtained with `DanielPlot()` with the option `half=TRUE`.

Half-Normal Plots - 2^4 design for studying a chemical reaction

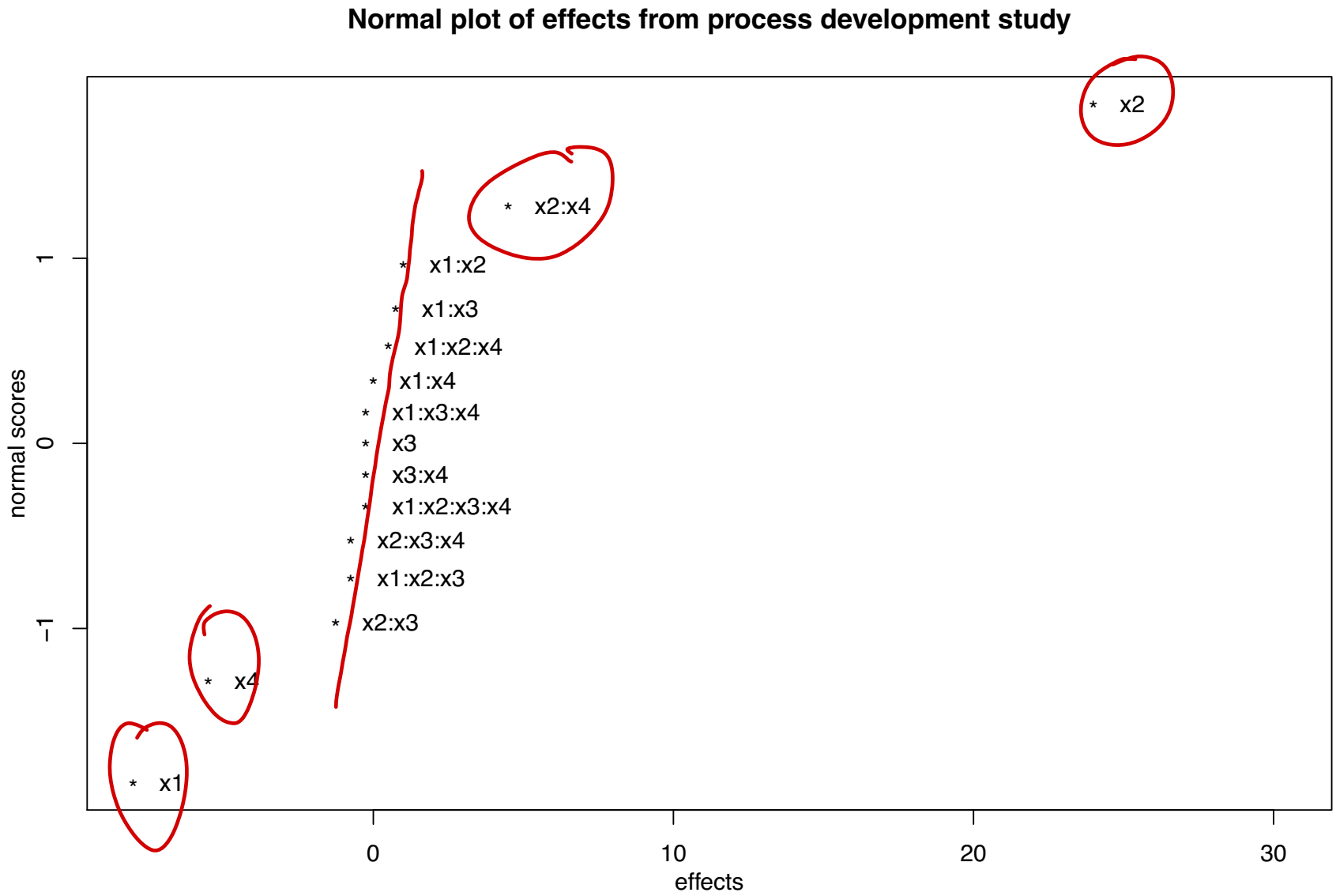
```
library(FrF2)
DanielPlot(fact1,half=TRUE,autolab=F,
           main="Normal plot of effects from process development study")
```



Half-Normal Plots - 2^4 design for studying a chemical reaction

Compare with full Normal plot.

```
library(FrF2)
DanielPlot(fact1,half=F,autolab=F,
           main="Normal plot of effects from process development study")
```



Multiple Comparisons

Suppose F test from ANOVA is Significant ∴ we would reject H_0.

Suppose that experimental units were randomly assigned to three treatment groups. The hypothesis of intrest is:

$$H_0 : \mu_1 = \mu_2 = \mu_3 \text{ vs. } H_1 : \mu_i \neq \mu_j.$$

Now, suppose that we reject H_0 at level α. Which pairs of means are significantly different from each other at level α? There are $\binom{3}{2} = 3$ possibilites.

1. $\mu_1 \neq \mu_2$
2. $\mu_1 \neq \mu_3$
3. $\mu_2 \neq \mu_3$

$H_{0_1}: \mu_1 = \mu_2$ $\quad$ $H_{0_2}: \mu_1 = \mu_3$

$H_{1_1}: \mu_1 \neq \mu_2$ $\quad$ $H_{1_2}: \mu_1 \neq \mu_3$

$H_{0_3}: \mu_2 = \mu_3$

$H_{1_3}: \mu_2 \neq \mu_3$.

Multiple Comparisons

Suppose that $k = 3$ separate (independent) hypothesis tests at level α tests are conducted:

By defn.

$$H_{0_k} : \mu_i = \mu_j \text{ vs. } H_{1_k} : \mu_i \neq \mu_j,$$

When H_0 is true, $P(\text{reject } H_0) = \alpha \Rightarrow 1 - P(\text{do not reject } H_0) = 1 - (1 - \alpha)$.

So, if H_0 is true then

A_i = {do not reject H_{0_i}}

$$P\left(\text{reject at least one } H_{0_k}\right) = 1 - P\left(\text{do not reject any } H_{0_k}\right)$$

A_i are indep.

Same event.

This is the same as

$$1 - P(\text{do not reject } H_{0_1} \text{ and do not reject } H_{0_2} \text{ and do not reject } H_{0_3})$$

$1 - P(A_1 \cap A_2 \cap A_3)$

or since the hypotheses are independent

If A_1, A_2, A_3 are indep. then $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$

$$1 - P(\text{do not reject } H_{0_1})\, P(\text{do not reject } H_{0_2})\, P(\text{do not reject } H_{0_3}) = 1 - (1 - \alpha)^3$$

$(1-\alpha)$ $(1-\alpha)$ $(1-\alpha)$

If $\alpha = 0.05$ then the probability that at least one H_0 will be falsely rejected is $1 - (1 - .05)^3 = 0.14$, which is almost three times the type I error rate.

of 0.05.

Multiple Comparisons

ANOVA Source SS df MS F P
Trt
error
XX
.05

A clinical trial comparing four treatment means using an ANOVA model at the 5% level found a significant F test. If all pairs of treatment means are compared then the probability of falsely declaring that at least one pair of treatment means is significantly different is:

Respond at **PollEv.com/nathantaback**

Text **NATHANTABACK** to **37607** once to join, then **A or B**

50% less than or equal to 0.05 **A** 44%

50% greater than 0.05 **B** 56%

ANOVA

Poll.

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

H_a: $\mu_i \neq \mu_j$ Some $i \neq j$.

	H_0:	H_A:
1	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$
2	$\mu_1 = \mu_3$	$\mu_1 \neq \mu_3$
3	$\mu_1 = \mu_4$	$\mu_1 \neq \mu_4$
4	$\mu_2 = \mu_3$	$\mu_2 \neq \mu_3$
5	$\mu_2 = \mu_4$	$\mu_2 \neq \mu_4$
6	$\mu_3 = \mu_4$	$\mu_3 \neq \mu_4$.

family-wise error rate.

$1 - .95^6 = 0.26$

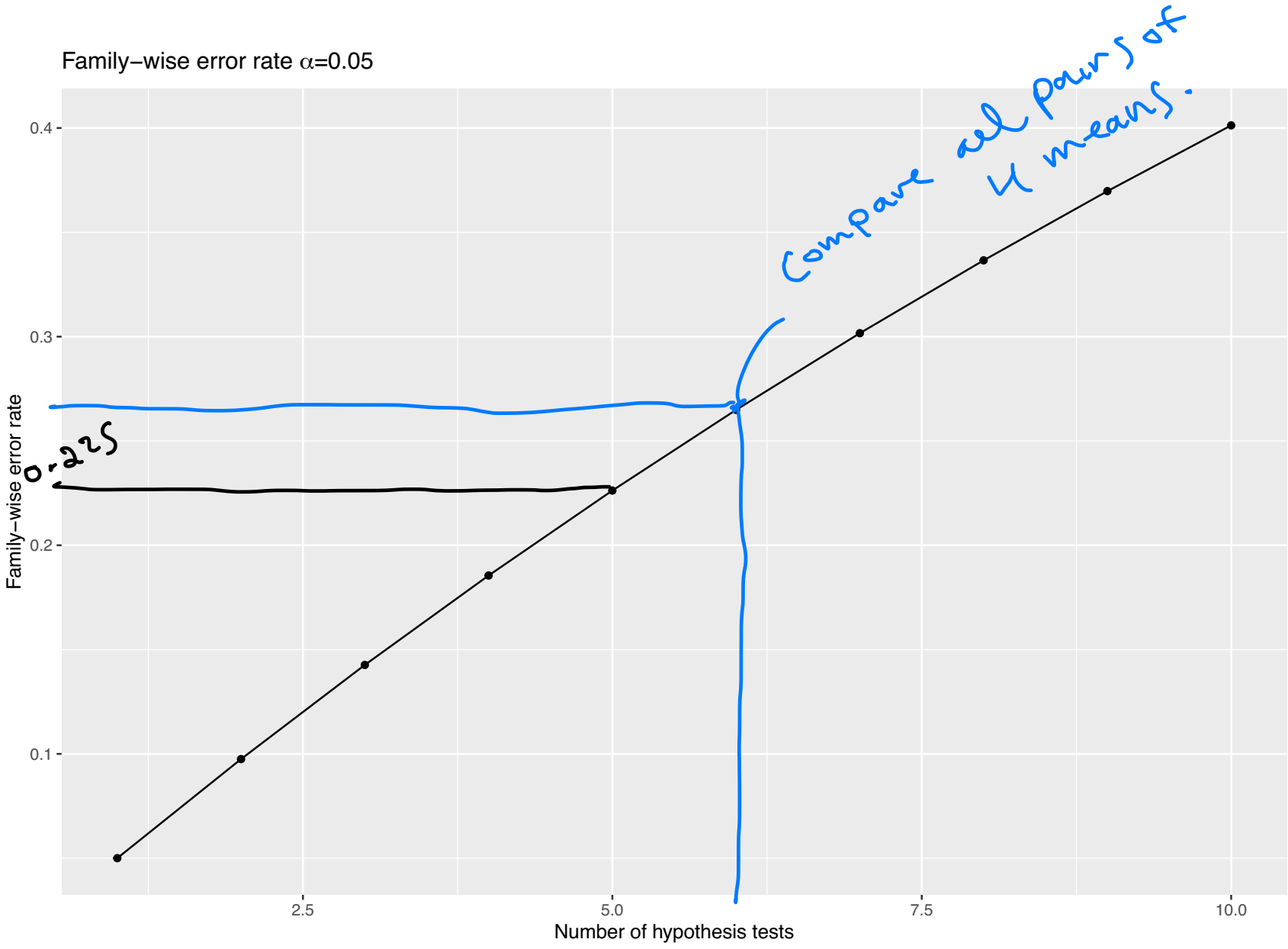
$1 - (1 - .05)^6 > .05$.

P(reject at least one H_0 when H_0 is true)

ANOVA gives information about if a significant difference exists, but doesn't tell you which differences are significant.

To tell which differences are significant use several two sample t-tests. But now the problem is that the type I error rate increases from original type I error rate.

Multiple Comparisons



Multiple Comparisons

In general if

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \text{ vs. } H_1 : \mu_i \neq \mu_j.$$ for some $i \neq j$.

If c independent hypotheses are conducted then the probability

$$P\left(\text{reject at least one } H_{0_k}\right) = 1 - (1-\alpha)^c$$ assuming that H_0 is true.

is called the **family-wise error rate**.

The **pairwise error rate** is $P\left(\text{reject } H_{0_k}\right) = \alpha$ for any c.

e.g., if $k = 3$

$c = 3$: $\mu_1 \neq \mu_2$, $\mu_1 \neq \mu_3$, $\mu_2 \neq \mu_3$

The Multiple Comparisons Problem

- The multiple comparison problem is that multiple hypotheses are tested level α which increases the probability that at least one of the hypotheses will be falsely rejected (family-wise error rate). $1 - .95^6$
- If treatment means are significantly different from the ANOVA F test then researchers will usually want to explore where the differences lie.
- Is it appropriate to test for differences looking at all pairwise comparisons?
- Testing all possible pairs increases the type I error rate.
- This means the chance that there is a higher probability, beyond the pre-stated type I error rate (e.g. 0.05), that that a significant difference is detected when the truth is that no difference exists.

Stop for Test #2.