STA305/1004 Class Notes - Week 10

Nathan Taback

March 23, 2016

1 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.

Suppose that we have observations independent observations $X_1, X_2, \ldots, X_n$ 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 $X_{(j)}$ vs. $\frac{j}{n+1}$ then if the underlying distribution is Unif[0,1] then the plot should be rooughly linear.

A continuous random variable with strictly increasing CDF F_X can be transformed to a Unif[0,1] by defining a new random variable $Y = F_X(X)$. This is also called the probability integral transformation.

This suggests the following procedure. Suppose that it's hypothesized that X follows a certain distribution function with CDF F. Given a sample $X_1, X_2, \ldots, X_n$ plot

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

or equivalently

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

$X_{(k)}$ can be thought of as empirical quantiles and $F^{-1}\left(\frac{k}{n+1}\right)$ as the hypothesized quantiles.

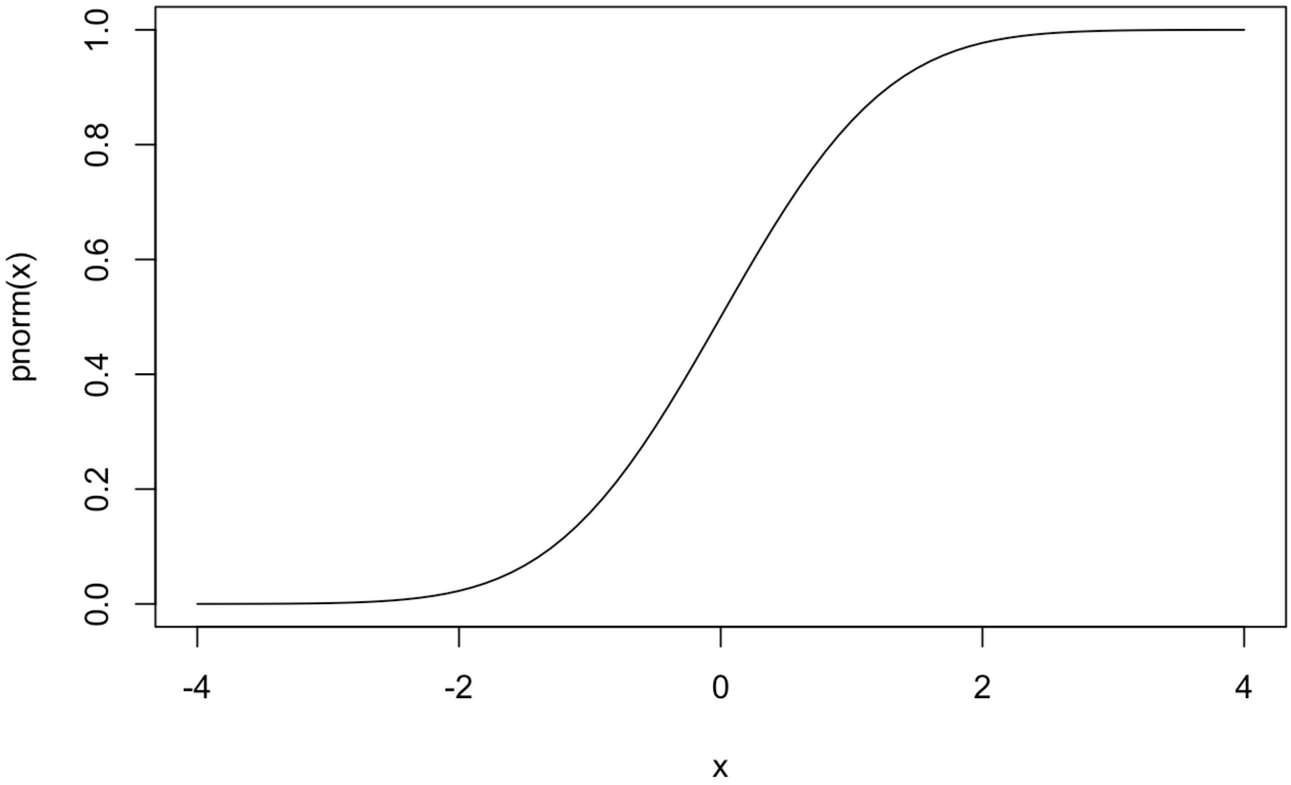
The quantile assigned to $X_{(k)}$ is not unique. Instead of assigning it $\frac{k}{n+1}$ it is often assigned $\frac{k-0.5}{n}$. In practice it makes little difference which definition is used.

2 Normal and Half Normal Plots

The normality of a set of data can be assessed by the following method. Let $r_{(1)} < \ldots < r_{(N)}$ denote the ordered values of $r_1, \ldots, r_N$. For example, $r_{(1)}$ is the minimum of $r_1, \ldots, r_N$, and $r_{(N)}$ is the maximum of $r_1, \ldots, r_N$. So, if the data is: -1, 2, -10, 20 then $r_{(1)} = -20, r_{(2)} = -1, r_{(3)} = 2, r_{(4)} = 20.$

The cumulative distribution function (CDF) of the $N(0, 1)$ has an S-shape.

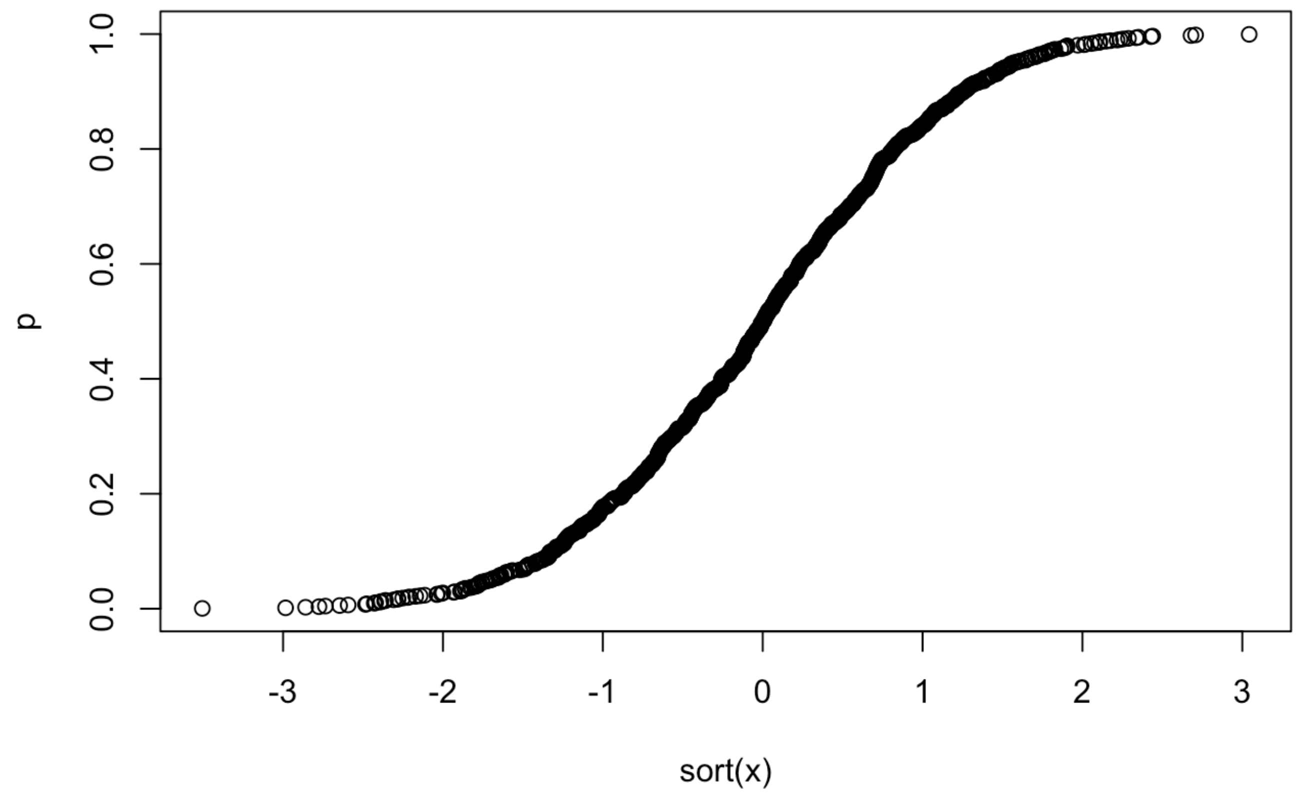
```
x <- seq(-4,4,by=0.1)
plot(x,pnorm(x),type="l")
```



So, a test of normality for a set of data is to plot the ordered values $r_{(i)}$ of the data versus $p_i = (i - 0.5)/N$. If the plot has the same S-shape as the normal CDF then this is evidence that the data come from a normal distribution.

Below is a plot of $r_{(i)}$ vs. $p_i = (i - 0.5)/N, i = 1, \ldots, N$ for a random sample of 1000 simulated from the plot

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



We can also construct a normal quantile-quantile plot. 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 ith order statistic from a uniform distribution over $[0, 1]$.

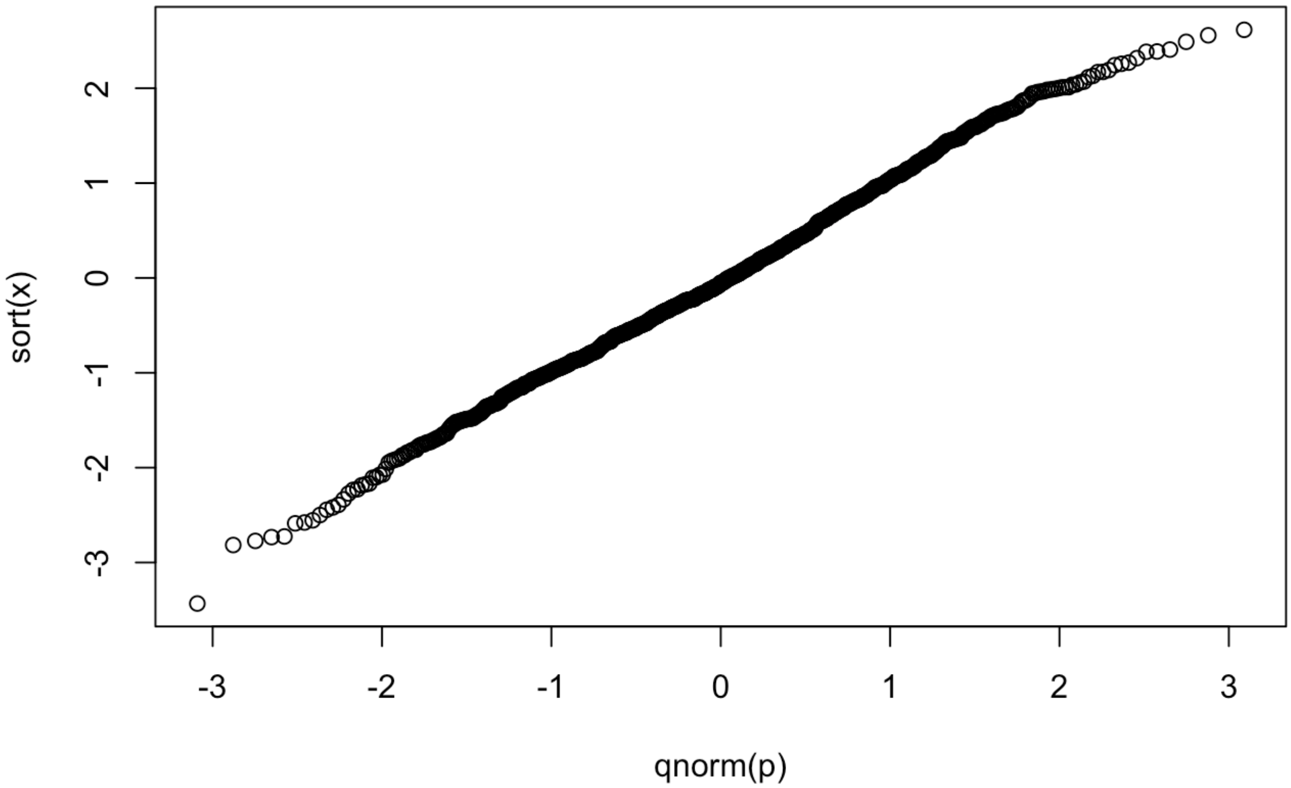
This implies that the N points $(p_i, \Phi(r_{(i)}))$ should fall on a straight line. Now apply the Φ^{-1} transformation to the horizontal and vertical scales. The N points

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

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

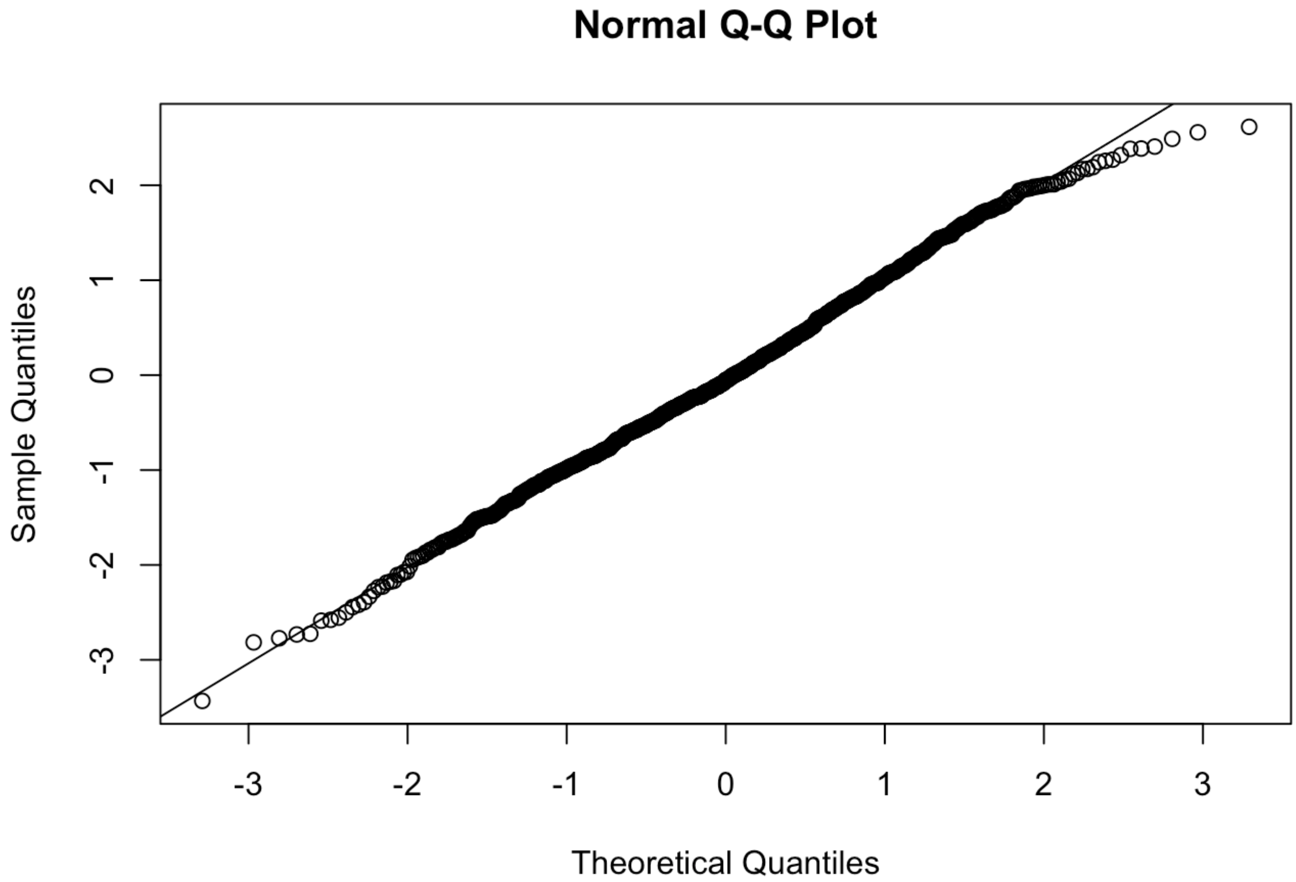
In R `qnorm()` is Φ^{-1}.

```
set.seed(2503)
N <- 1000
x <- rnorm(N)
p <- (1:N)/(N+1)
plot(qnorm(p),sort(x))
```



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)
```



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.

3 Normal plots

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

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

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.

The rationale is as follows: 1. Assume that the estimated effects $\hat{\theta}_i$ are $N(\theta, \sigma)$ (estimated effects involve averaging of N observations and CLT ensures averages are nearly normal for N as small as 8). 2. If $H_0 : \theta_i = 0,\ i = 1, \ldots, N$ 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).

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

- For positive effects the estimated effects fall above the line and negative effects fall below the line.

3.1 Example - 2^4 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**, and 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.

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

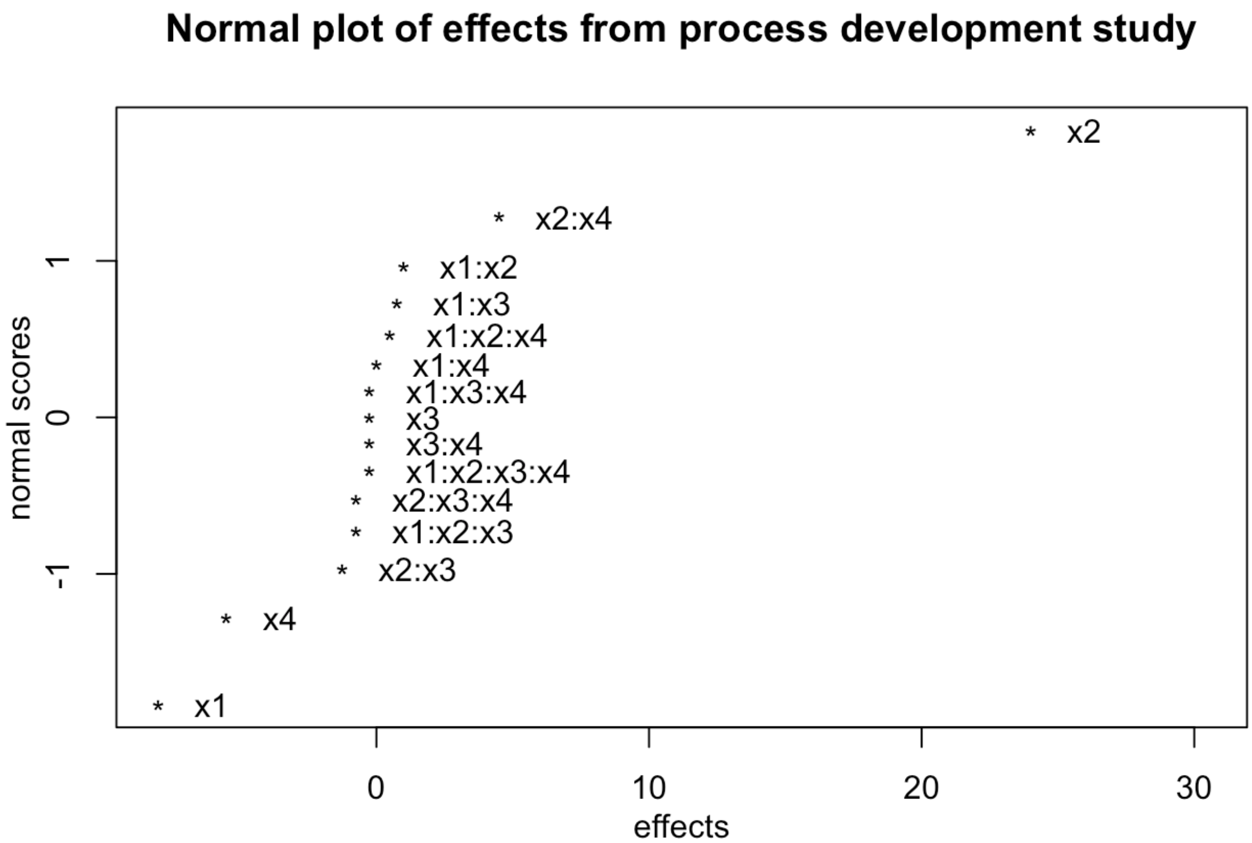
The design is not replicated so it's not possible to estimate the standard errors of the factorial effects.

```
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
```

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

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



The effects corresponding to `x1, x4, x2:x4, x2` do not fall along the straight line.

4 Half-Normal Plots

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.

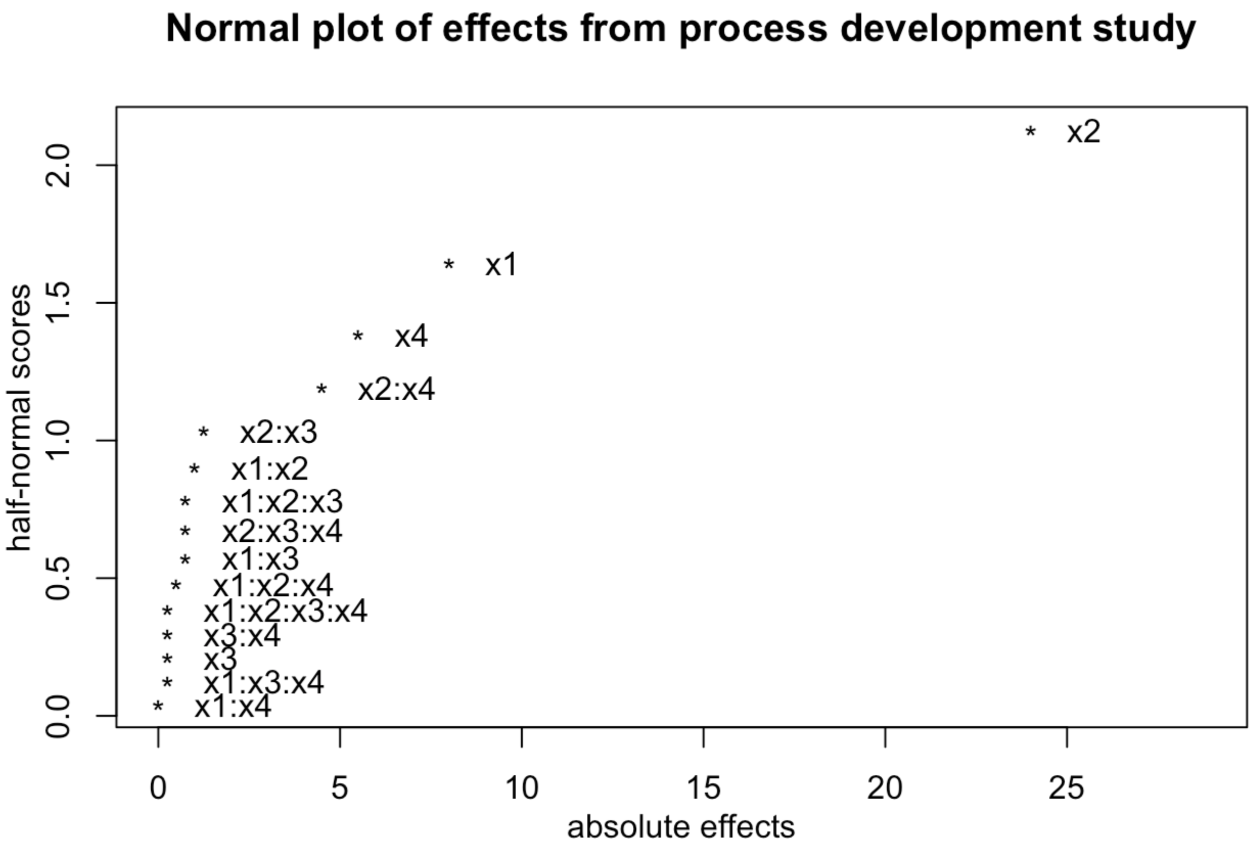
The half-normal probability plot consists of the points

$$\left|\hat{\theta}\right|_{(i)} \text{ vs. } \Phi^{-1}(0.5 + 0.5[i - 0.5]/N).\ i = 1, \ldots, N.$$

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`.

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



5 Lenth's method: testing significance for experiments without variance estimates

Half-normal and normal plots are informal graphical methods involving visual judgement. It's desirable to judge a deviation from a straight line quantitatively based on a formal test of significance. Lenth (1989) proposed a method that is simple to compute and performs well. (ref. pg. 205, Box, Hunter, and Hunter, 2005)

Let

$$\hat{\theta}_{(1)}, \ldots, \hat{\theta}_{(N)}$$

be estimated factorial effects of $\theta_1, \theta_2, \ldots, \theta_N$ In a 2^k design $N = 2^k - 1$. Assume that all the factorial effects have the same standard deviation.

The pseudo standard error (PSE) is defined as

$$PSE = 1.5 \cdot \text{median}_{\left|\hat{\theta}_i\right| < 2.5 s_0} \left|\hat{\theta}_i\right|,$$

where the median is computed among the $\left|\hat{\theta}_i\right|$ with $\left|\hat{\theta}_i\right| < 2.5 s_0$ and

$$s_0 = 1.5 \cdot \text{median} \left|\hat{\theta}_i\right|.$$

$1.5 \cdot s_0$ is a consistent estimator of the standard deviation of $\hat{\theta}$ when $\theta_i = 0$ and the underlying distribution is normal. The $P\left(|Z| > 2.57\right) = 0.01, Z \sim N(0, 1)$. So, $\left|\hat{\theta}_i\right| < 2.5 s_0$ trims approximately 1% of the $\hat{\theta}_i$ if $\theta_i = 0$. The trimming attempts to remove the $\hat{\theta}_i$ associated with non-zero (active) effects. By using the median in combination with the trimming means that PSE is not sensitive to the $\hat{\theta}_i$ associated with active effects.

By dividing $\hat{\theta}_i$ by PSE, t-like statistics are obtained:

$$t_{PSE,i} = \frac{\hat{\theta}_i}{PSE}.$$

(see Wu and Hamada, pg. 180)

Lenth's method declares an effect $\hat{\theta}_i$ significant if the value of $\left|t_{PSE,i}\right|$ value exceeds the critical value of the distribution. The critical values have been calculated by Wu and Hamada (2009).

To obtain a margin of error Lenth suggested multiplying the PSE by the $100 * (1 - \alpha)$ quantile of the t_d distribution, $t_{d,\alpha/2}$. The degrees of freedom is $d = N/3$. For example, the margin of error for a 95% confidence interval for θ_i is

$$ME = t_{d,.025} \times PSE.$$

All estimates greater than the ME may be viewed as "significant", but with so many estimates being considered simultaneously, some will be falsely identified.

A simultaneous margin of error that accounts for multiple testing can also be calculated,

$$SME = t_{d,\gamma} \times PSE,$$

where $\gamma = \left(1 + 0.95^{1/N}\right)/2$.

Let's calculate Lenth's method for the process development example. The estimated factorial effects are:

```
eff <- 2*fact1$coefficients
round(eff,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
```

The estimate of $s_0 = 1.5 \cdot \text{median} \left|\hat{\theta}_i\right|$ is

```
s0 <- 1.5*median(abs(eff))
s0
```

```
[1] 1.125
```

The trimming constant $2.5 s_0$ is

```
2.5*s0
```

```
[1] 2.8125
```

The effects $\hat{\theta}_i$ such that $\left|\hat{\theta}_i\right| \geq 2.5 s_0$ will be trimmed. Below it's the effects labelled `TRUE` (`x1,x2,x4,x2:x4`)

```
abs(eff)<2.5*s0
```

```
(Intercept)          x1          x2          x3          x4       x1:x2
      FALSE       FALSE       FALSE        TRUE       FALSE        TRUE
      x1:x3       x2:x3       x1:x4       x2:x4       x3:x4    x1:x2:x3
       TRUE        TRUE        TRUE       FALSE        TRUE        TRUE
   x1:x2:x4    x1:x3:x4    x2:x3:x4 x1:x2:x3:x4
       TRUE        TRUE        TRUE        TRUE
```

The PSE is then calculated as 1.5 times the median of these values.

```
PSE <- 1.5*median(abs(eff[abs(eff)<2.5*s0]))
PSE
```

```
[1] 0.75
```

The ME and SME are

```
ME <- PSE*qt(p = .975,df = (16-1)/3)
ME
```

```
[1] 1.927936
```

```
SME <- PSE*qt(p =(1+.95^{1/15})/2,df=(16-1)/3)
SME
```

```
[1] 3.913988
```

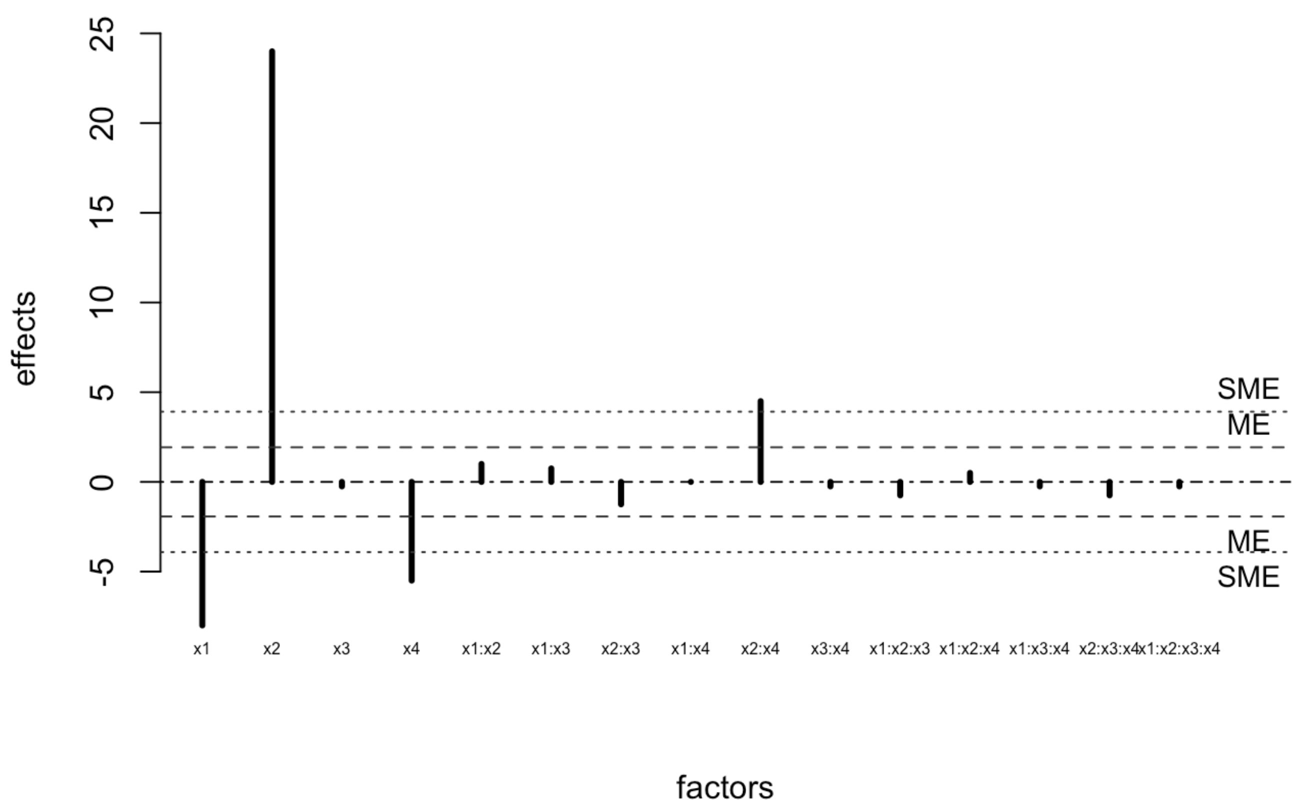
So, 95% confidence intervals for the effects are:

```
lower <- round(eff-ME,2)
upper <- round(eff+ME,2)
knitr::kable(cbind(eff,lower,upper))
```

	eff	lower	upper
(Intercept)	144.50	142.57	146.43
x1	-8.00	-9.93	-6.07
x2	24.00	22.07	25.93
x3	-0.25	-2.18	1.68
x4	-5.50	-7.43	-3.57
x1:x2	1.00	-0.93	2.93
x1:x3	0.75	-1.18	2.68
x2:x3	-1.25	-3.18	0.68
x1:x4	0.00	-1.93	1.93
x2:x4	4.50	2.57	6.43
x3:x4	-0.25	-2.18	1.68
x1:x2:x3	-0.75	-2.68	1.18
x1:x2:x4	0.50	-1.43	2.43
x1:x3:x4	-0.25	-2.18	1.68
x2:x3:x4	-0.75	-2.68	1.18
x1:x2:x3:x4	-0.25	-2.18	1.68

A plot of the effects with a ME and SME is usually called a Lenth plot. In R it can be implemented via the function `Lenthplot()` in the `BsMD` library. The values of PSE, ME, SME are part of the output. The spikes in the plot below are used to display factor effects.

```
library(BsMD)
LenthPlot(fact1,cex.fac = 0.5)
```




```
   alpha       PSE         ME        SME
0.050000 0.750000 1.927936 3.913988
```

The option `cex.fac=0.5` adjusts the size of the characters used for factor labels.

6 Blocking factorial designs

In a trial conducted using a 2^3 design it might be desirable to use the same batch of raw material to make all 8 runs. Suppose that batches of raw material were only large enough to make 4 runs. Then the concept of blocking could be used.

The following R code generates the design matrix for a 2^3 design.

```
x1 <- rep(c(-1,1),4)
x2 <- rep(c(-1,-1,1,1),2)
x3 <- rep(c(rep(-1,4),rep(1,4)))
x12 <- x1*x2
x13 <- x1*x3
x23 <- x2*x3
x123 <- x1*x2*x3
run <- 1:8
factnames <- c("Run","1","2","3","12","13","23","123")
knitr::kable(cbind(run,x1,x2,x3,x12,x13,x23,x123),col.names=factnames)
```

Run	1	2	3	12	13	23	123
1	-1	-1	-1	1	1	1	-1
2	1	-1	-1	-1	-1	1	1
3	-1	1	-1	-1	1	-1	1
4	1	1	-1	1	-1	-1	-1
5	-1	-1	1	1	-1	-1	1
6	1	-1	1	-1	1	-1	-1
7	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1

Suppose that we assign runs 1, 4, 6, 7 to block I which use the first batch of raw material and runs 2, 3, 5, 8 to block II which use the second batch of raw material. The design is blocked this way by placing all runs in which the 123 is minus in one block and all the other runs in which 123 is plus in the other block.

Any systematic differences between the two blocks of four runs will be eliminated from all the main effects and two factor interactions. What you gain is the elimination of systematic differences between blocks. But now the three factor interaction is confounded with any batch (block) difference. The ability to estimate the three factor interaction separately from the block effect is lost.

6.1 Effect hierarchy principle

1. Lower-order effects are more likely to be important than higher-order effects.
2. Effects of the same order are equally likely to be important.

This principle suggests that when resources are scare, priority should be given to the estimation of lower order effects. This is useful in screening experiments that have a large number of factors and relatively small number of runs.

One reason that many accept this principle is that higher order interactions are more difficult to interpret or justify physically. As a result investigators are less interested in estimating the magnitudes of these effects even when they are statistically significant.

Assigning a fraction of the 2^k treatment combinations to each block results in an incomplete blocking scheme as in the case of the balanced incomplete block design. The difference is that the factorial structure of a 2^k design allows a neater assignment of treatment combinations to blocks. The neater assignment is done by dividing the total combinations into various fractions and finding optimal assignments by exploiting combinatorial relationships.

6.2 Generation of Orthogonal Blocks

In the 2^3 example suppose that the block variable is given the identifying number 4.

Run	1	2	3	4=123
1	-1	-1	-1	-1
2	1	-1	-1	1
3	-1	1	-1	1
4	1	1	-1	-1
5	-1	-1	1	1
6	1	-1	1	-1
7	-1	1	1	-1
8	1	1	1	1

Then you could think of your experiment as containing four factors. The fourth factor will have the special property that it does not interact with other factors. If this new factor is introduced by having its levels coincide exactly with the plus and minus signs attributed to 123 then the blocking is said to be generated by the relationship 4=123. This idea can be used to derive more sophisticated blocking arrangements.

6.2.1 An example of how not to block

Suppose we would like to arrange the 2^3 design into four blocks.

Run	1	2	3	4=123	5=23	45=1
1	-1	-1	-1	-1	1	-1
2	1	-1	-1	1	1	1
3	-1	1	-1	1	-1	-1
4	1	1	-1	-1	-1	1
5	-1	-1	1	1	-1	-1

6	1	-1	1	-1	-1	1
7	-1	1	1	-1	1	-1
8	1	1	1	1	1	1

Consider two block factors called 4 and 5. 4 is associated with the three factor interaction and, say, 5 is associated with a the two factor interaction 23 which was deemed unimportant by the investigator. Runs are placed in different blocks depending on the signs of the block variables in columns 4 and 5. Runs for which the signs of 4 and 5 are – would go in one block, -+ in a second block, the +- in a third block, and the ++ runs in the fourth.

Block	**Run**
I	4,6
II	3,5
III	1,7
IV	2,8

Block variables 4 and 5 are confounded with interactions 123 and 23. But there are three degrees of freedom assoicated with four blocks. The third degree of freedom accomodates the 45 interaction. But, the 45 interaction has the same signs as the main effect 1. Therefore 45=1. Therefore, if we use 4 and 5 as blocking variables it will be confounded with block differences.

Main effects should not be confounded with block effects. Any blocking scheme that confounds main effects with blocks should not be used. This is based on the assumption: The block-by-treatment interactions are negligible.

This assumption states that treatment effects do not vary from block to block. Without this assumption estimability of the factorial effects will be very complicated.

For example, if $B_1 = 12$ then this implies two other relations:

$$1B_1 = 1 \times B_1 = 112 = 2 \text{ and } B_1 2 = B_1 \times 2 = 122 = 1.$$

If there is a significant interaction between the block effect B_1 and the main effect 1 then the main effect 2 is confounded with $B_1 1$. Similarily, if there is a significant interaction between the block effect B_1 and the main effect 2 then the main effect 1 is confounded with $B_1 2$.

It can be checked by plotting the residuals for all the treatments within each block. If the pattern varies from block to block then the assumption may be violated. A block-by-treatment interaction often suggests interesting information about the treatment and blocking variables.

6.3 Generators and Defining Relations

A simple calculus is available to show the consequences of any proposed blocking arrangement. If any column in a 2^k design are multiplied by themselves a column of plus signs is obtained. This is denoted by the symbol I. Thus you can write

$$I = 11 = 22 = 33 = 44 = 55,$$

where, for example, 22 means the product of the elements of column 2 with itself.

Any column multiplied by I leaves the elements unchanged. So, $I3 = 3$.

A general approach for arranging a 2^k design in 2^q blocks of size 2^{k-q} is as follows.

Let $B_1, B_2, \ldots, B_q$ be the block variables and the factorial effect v_i is confounded with B_i,

$$B_1 = v_1, B_2 = v_2, \ldots, B_q = v_q.$$

The block effects are obtained by multiplying the B_i's:

$$B_1B_2 = v_1v_2, B_1B_3 = v_1v_3, \ldots, B_1B_2 \cdots B_q = v_1v_2 \cdots v_q$$

There are $2^q - 1$ possible products of the B_i's and the I (whose components are +).

Example: A 2^5 design can be arranged in 8 blocks of size $2^{5-3} = 4$.

Consider two blocking schemes.

1. Define the blocks as

$$B_1 = 135, B_2 = 235, B_3 = 1234.$$

The remaining blocks are confounded with the following interactions:

$$B_1B_2 = 12, B_1B_3 = 245, B_2B_3 = 145, B_1B_2B_3 = 34$$

In this blocking scheme the seven block effects are confounded with the seven interactions

$$12, 34, 135, 145, 235, 245, 1234.$$

2. Define the blocks as:

$$B_1 = 12, B_2 = 13, B_3 = 45.$$

This blocking scheme confounds the following interactions.

$$12, 13, 23, 45, 1245, 1345, 2345.$$

Which is a better blocking scheme?

The second scheme confounds four two-factor interactions, while the first confounds only two two-factor interactions. Since two-factor interactions are more likely to be important than three- or four-factor interactions, the first scheme is superior.

7 Fractional factorial designs

A 2^k full factorial requires 2^k runs. Full factorials are seldom used in practice for large k (k>=7). For economic reasons fractional factorial designs, which consist of a fraction of full factorial designs are used. There are criteria to choose “optimal” fractions.

7.1 Example - Effect of five factors on six properties of film in eight runs

The following example is taken from Box, Hunter, and Hunter (2005).

Five factors were studied in 8 runs. The factors were:

- Catalyst concentration (A)
- Amount of additive (B)
- Amounts of three emulsifiers (C, D, E)

Polymer solutions were prepared and spread as a film on a microscope slide. Six different responses were recorded.

run	A	B	C	D	E	y1	y2	y3	y4	y5	y6
1	-1	-1	-1	1	-1	no	no	yes	no	slightly	yes
2	1	-1	-1	1	1	no	yes	yes	yes	slightly	yes
3	-1	1	-1	-1	1	no	no	no	yes	no	no
4	1	1	-1	-1	-1	no	yes	no	no	no	no
5	-1	-1	1	-1	1	yes	no	no	yes	no	slightly
6	1	-1	1	-1	-1	yes	yes	no	no	no	no
7	-1	1	1	1	-1	yes	no	yes	no	slightly	yes
8	1	1	1	1	1	yes	yes	yes	yes	slightly	yes

- The eight run design was constructed beginning with a standard table of signs for a 2^3 design in the factors A, B, C.
- The column of signs associated with the BC interaction was used to accommodate factor D, the ABC interaction column was used for factor E.
- A full factorial for the five factors A, B, C, D, E would have needed $2^5 = 32$ runs.
- Only 1/4 were run. This design is called a quarter fraction of the full 2^5 or a 2^{5-2} design (a two to the five minus two design). In general a 2^{k-p} design is a $\frac{1}{2^p}$ fraction of a 2^k design using 2^{k-p} runs.

7.2 Effect Aliasing and Design Resolution

A chemist in an industrial development lab was trying to formulate a household liquid product using a new process. The liquid had good properties but was unstable. The chemist wanted to synthesize the product in hope of hitting conditions that would give stability, but without success. The chemist identified four important influences: A (acid concentration), B (catalyst concentration), C (temperature), D (monomer concentration). His 8 run fractional factorial design is shown below.

test	A	B	C	D	y
1	-1	-1	-1	-1	20
2	1	-1	-1	1	14
3	-1	1	-1	1	17
4	1	1	-1	-1	10
5	-1	-1	1	1	19
6	1	-1	1	-1	13
7	-1	1	1	-1	14
8	1	1	1	1	10

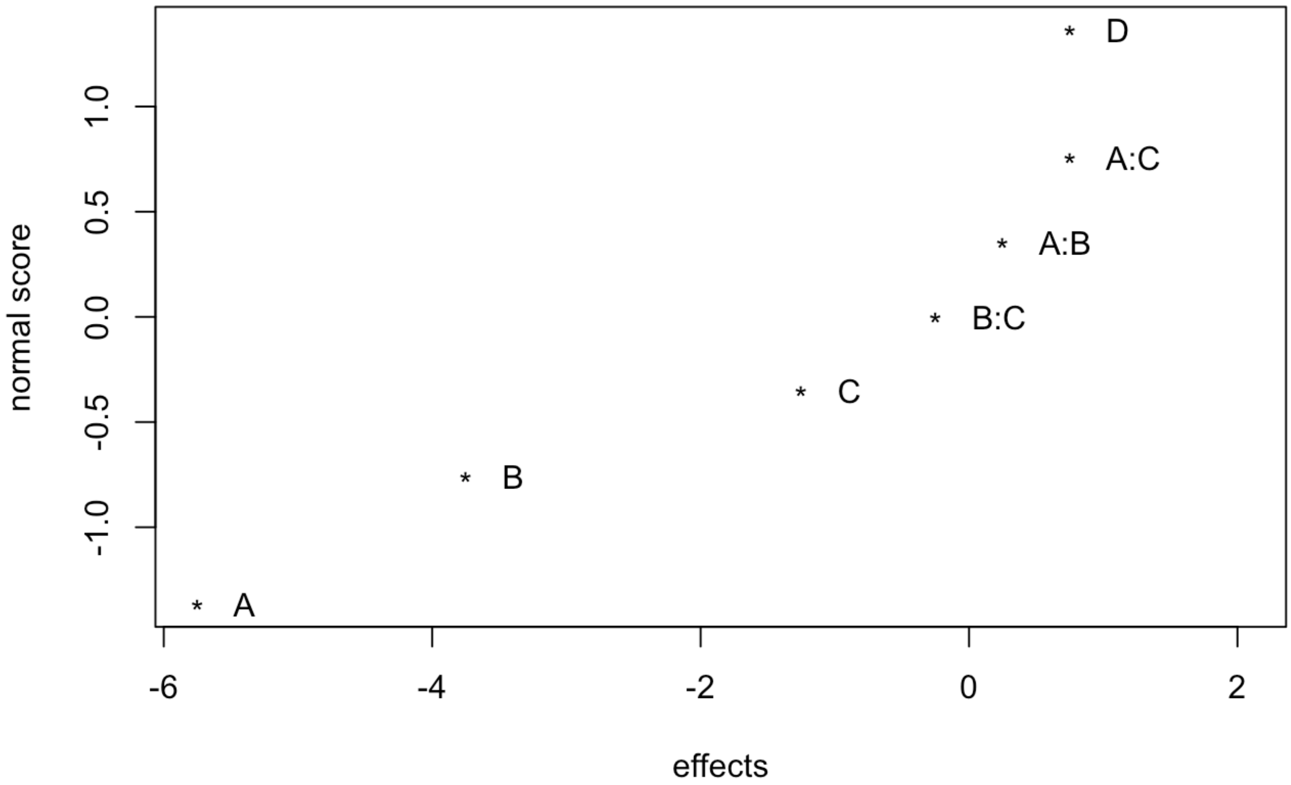
The signs of the ABC interaction is used to accommodate factor D. The tests were run in random order. He wanted to achieve a stability value of at least 25.

The factorial effects and Normal, half-Normal, and Lenth plots are below.

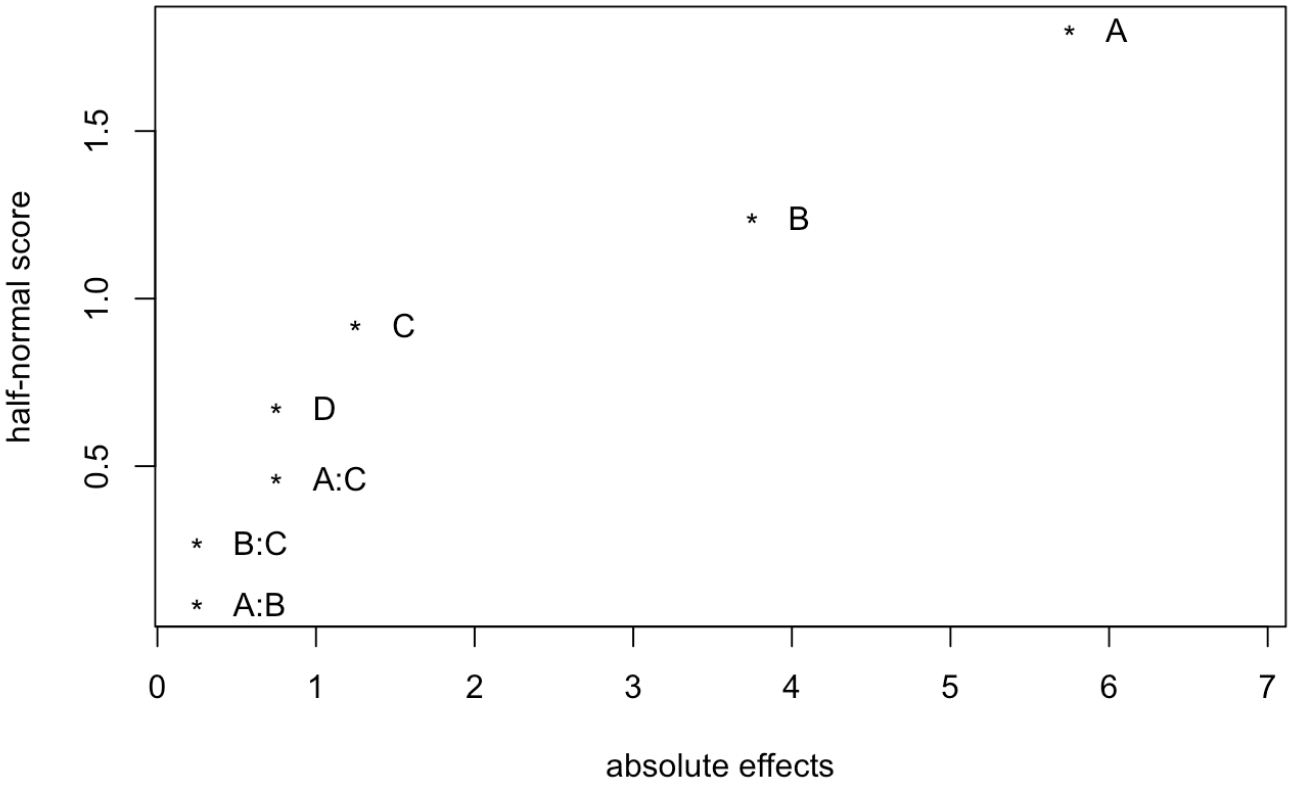
```
library(FrF2)
fact.prod <- lm(y~A*B*C*D,data=tab0602)
fact.prod1 <- aov(y~A*B*C*D,data=tab0602)
round(2*fact.prod$coefficients,2)
```

```
(Intercept)           A           B           C           D         A:B
      29.25       -5.75       -3.75       -1.25        0.75        0.25
        A:C         B:C         A:D         B:D         C:D       A:B:C
       0.75       -0.25          NA          NA          NA          NA
      A:B:D       A:C:D       B:C:D     A:B:C:D
         NA          NA          NA          NA
```

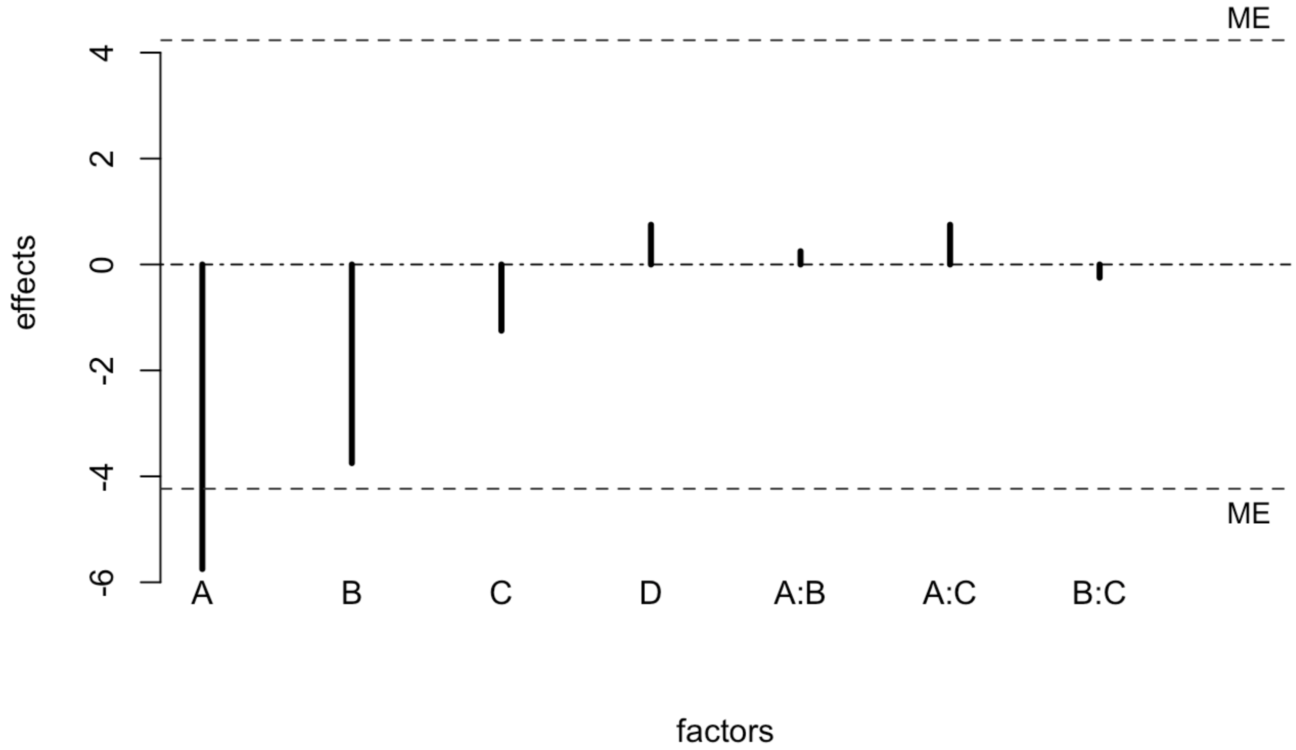
```
DanielPlot(fact.prod,half = F)
```




```
DanielPlot(fact.prod,half = T)
```




```
LenthPlot(fact.prod1)
```




```
   alpha       PSE         ME       SME
0.050000  1.125000  4.234638 10.134346
```

Even though the stability never reached the desired level of 25, two important factors, A and B, were identified. This Normal and half-Normal plots indicate the importance of these factors, although factor B is not significant according to the Lenth plot.

What information could have been obtained if a full 2^5 design had been used?

Factors	Number of effects
Main	5
2-factor	10
3-factor	10
4-factor	5
5-factor	1

There are 31 degrees of freedom in a 32 run design. But, are 16 used for estimating three factor interactions or higher. Is it practical to commit half the degrees of freedom to estimate such effects? According to effect hierarchy principle three-factor and higher not usually important. Thus, using full factorial wasteful. It's more economical to use a fraction of full factorial design that allows lower order effects to be estimated.

Consider a design that studies five factors in 16 run. A half fraction of a 2^5 or 2^{5-1}.

Run	B	C	D	E	Q
1	-1	1	1	-1	-1
2	1	1	1	1	-1
3	-1	-1	1	1	-1
4	1	-1	1	-1	-1
5	-1	1	-1	1	-1
6	1	1	-1	-1	-1
7	-1	-1	-1	-1	-1
8	1	-1	-1	1	-1
9	-1	1	1	-1	1
10	1	1	1	1	1
11	-1	-1	1	1	1
12	1	-1	1	-1	1
13	-1	1	-1	1	1
14	1	1	-1	-1	1
15	-1	-1	-1	-1	1
16	1	-1	-1	1	1

The factor E is assigned to the column BCD. But, the column for E is used to estimate the main effect of E and also for BCD. So, this design cannot distinguish between E and BCD. The main factor E is said to be **aliased** with the BCD interaction.

This aliasing relation is denoted by

$$E = BCD \, or \, I = BCDE,$$

where I denotes the column of all +'s.

This uses same mathematical definition as the confounding of a block effect with a factorial effect. Aliasing of the effects is a price one must pay for choosing a smaller design.

The 2^{5-1} design has only 15 degrees of freedom for estimating factorial effects, it cannot estimate all 31 factorial effects among the factors B, C, D, E, Q.

The equation $I = BCDE$ is called the **defining relation** of the 2^{5-1} design. The design is said to have resolution IV because the defining relation consists of the "word" BCDE, which has "length" 4.

Multiplying both sides of $I = BCDE$ by column B

$$B = B \times I = B \times BCDE = CDE$$

, the relation $B = CDE$ is obtained. B is aliased with the CDE interaction. Following the same method all 15 aliasing relations can be obtained.

$$B = CDE, C = BDE, D = BCE, E = BCD,$$
$$BC = DE, BD = CE, BE = CD,$$
$$Q = BCDEQ, BQ = CDEQ, CQ = BDEQ, DQ = BCEQ,$$
$$EQ = BCDQ, BCQ = DEQ, BDQ = CEQ, BEQ = CDQ$$

Each of the four main effects B, C, D, E is respectively aliased with CDE, BDE, BCE, BCD. Therefore, the main effects of B, C, D, E are estimable only if the aforementioned three-factor interactions are negligible. The other factorial effects have analogous aliasing properties.

7.3 Example - Leaf spring experiment

The following example is from Wu and Hamada (2009). An experiment to improve a heat treatment process on truck leaf springs. The heat treatment that forms the camber in leaf springs consists of heating in a high temperature furnace, processing by forming a machine , and quenching in an oil bath. The free height of an unloaded spring has a target value around 8in. The goal of the experiment is to make the variation about the target as small as possible.

Five factors were studied in this 2^{5-1} design.

Factor	**Level**
B. Temperature	1840 (-), 1880 (+)
C. Heating time	23 (-), 25 (+)
D. Transfer time	10 (-), 12 (+)
E. Hold down time	2 (-), 3 (+)
Q. Quench oil temperature	130-150 (-), 150-170 (+)

B	**C**	**D**	**E**	**Q**	**y**
-1	1	1	-1	-1	7.7900
1	1	1	1	-1	8.0700
-1	-1	1	1	-1	7.5200
1	-1	1	-1	-1	7.6333
-1	1	-1	1	-1	7.9400
1	1	-1	-1	-1	7.9467
-1	-1	-1	-1	-1	7.5400
1	-1	-1	1	-1	7.6867
-1	1	1	-1	1	7.2900

1	1	1	1	1	7.7333
-1	-1	1	1	1	7.5200
1	-1	1	-1	1	7.6467
-1	1	-1	1	1	7.4000
1	1	-1	-1	1	7.6233
-1	-1	-1	-1	1	7.2033
1	-1	-1	1	1	7.6333

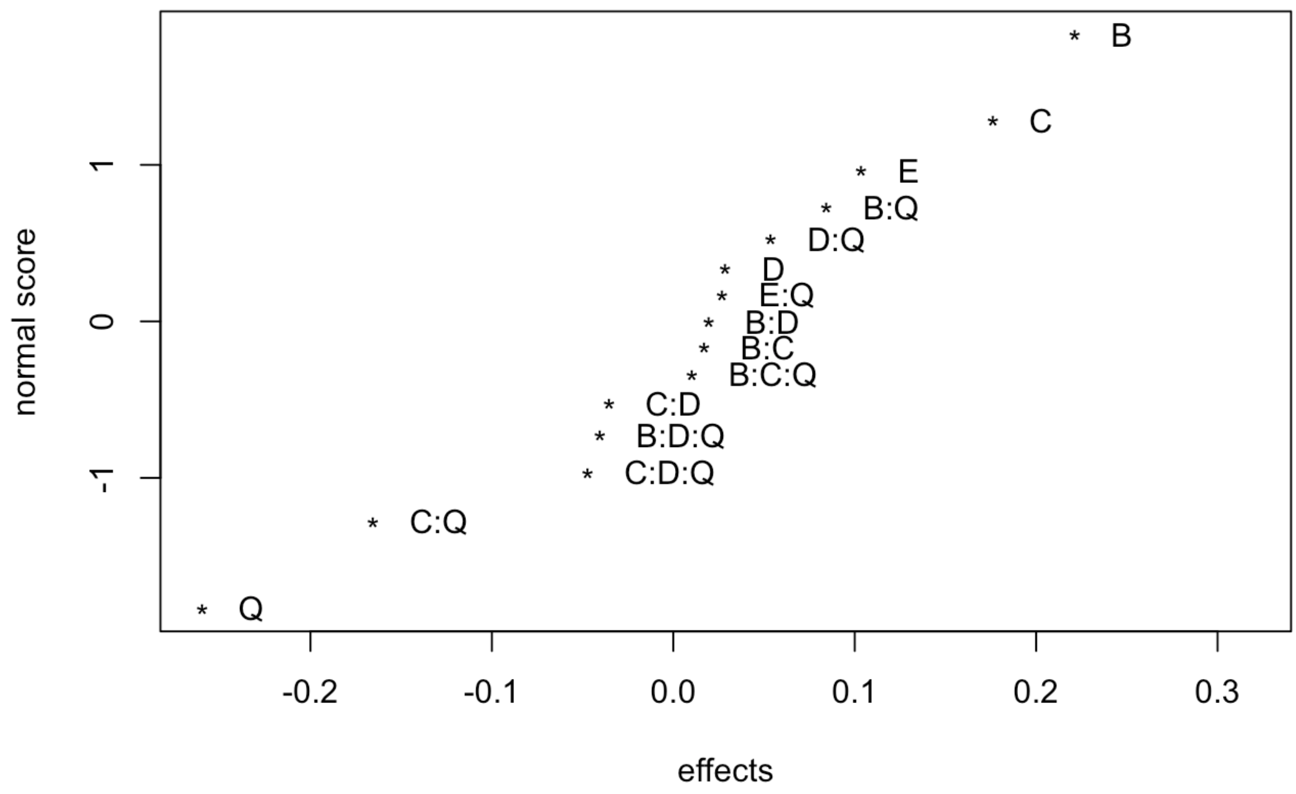
The factorial effects are estimated as before.

```
library(FrF2)
fact.leaf <- lm(y~B*C*D*E*Q,data=leafspring)
fact.leaf2 <- aov(y~B*C*D*E*Q,data=leafspring)
round(2*fact.leaf$coefficients,2)
```

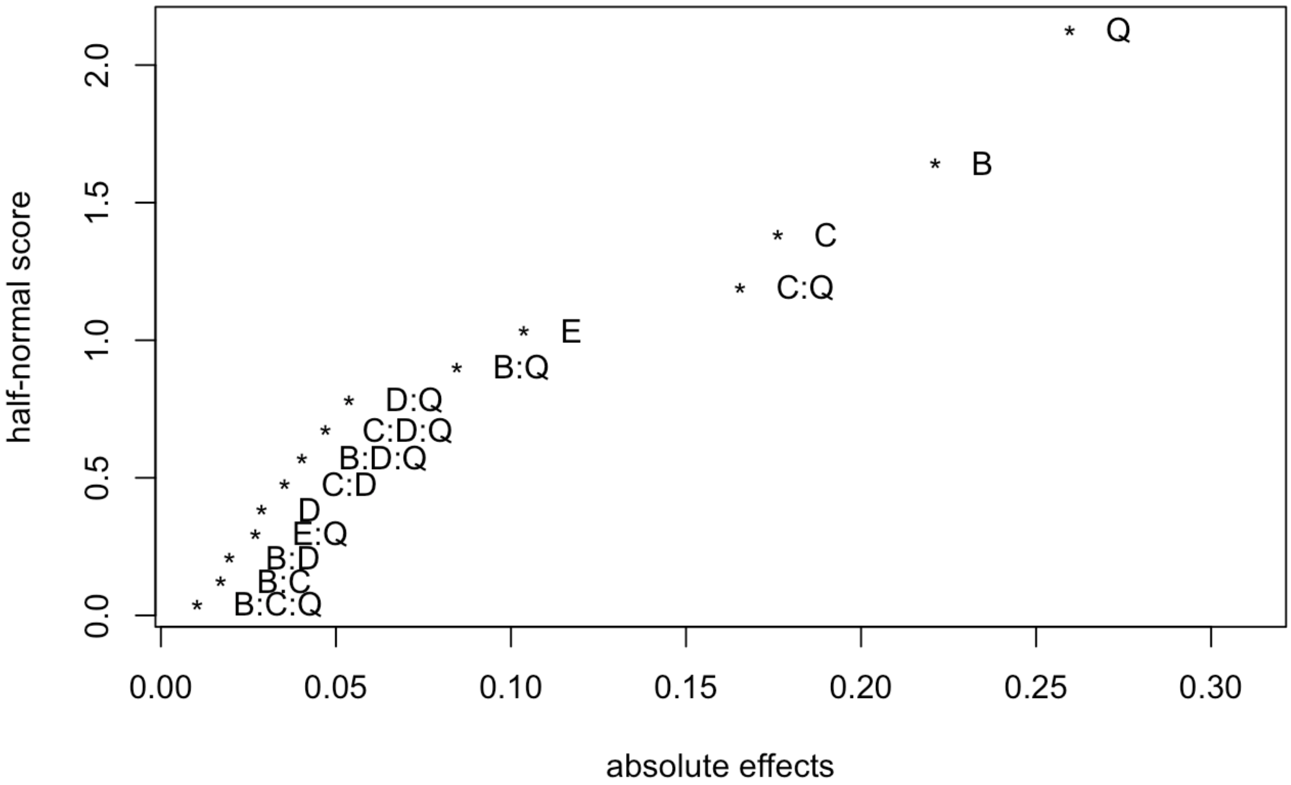
```
(Intercept)           B           C           D           E           Q 
      15.27        0.22        0.18        0.03        0.10       -0.26 
        B:C         B:D         C:D         B:E         C:E         D:E 
       0.02        0.02       -0.04          NA          NA          NA 
        B:Q         C:Q         D:Q         E:Q       B:C:D       B:C:E 
       0.08       -0.17        0.05        0.03          NA          NA 
      B:D:E       C:D:E       B:C:Q       B:D:Q       C:D:Q       B:E:Q 
         NA          NA        0.01       -0.04       -0.05          NA 
      C:E:Q       D:E:Q     B:C:D:E     B:C:D:Q     B:C:E:Q     B:D:E:Q 
         NA          NA          NA          NA          NA          NA 
    C:D:E:Q   B:C:D:E:Q 
         NA          NA 
```

Notice that the factorial effects are missing for effects that are aliased. The Normal, half-Normal, and Lenth plots are below.

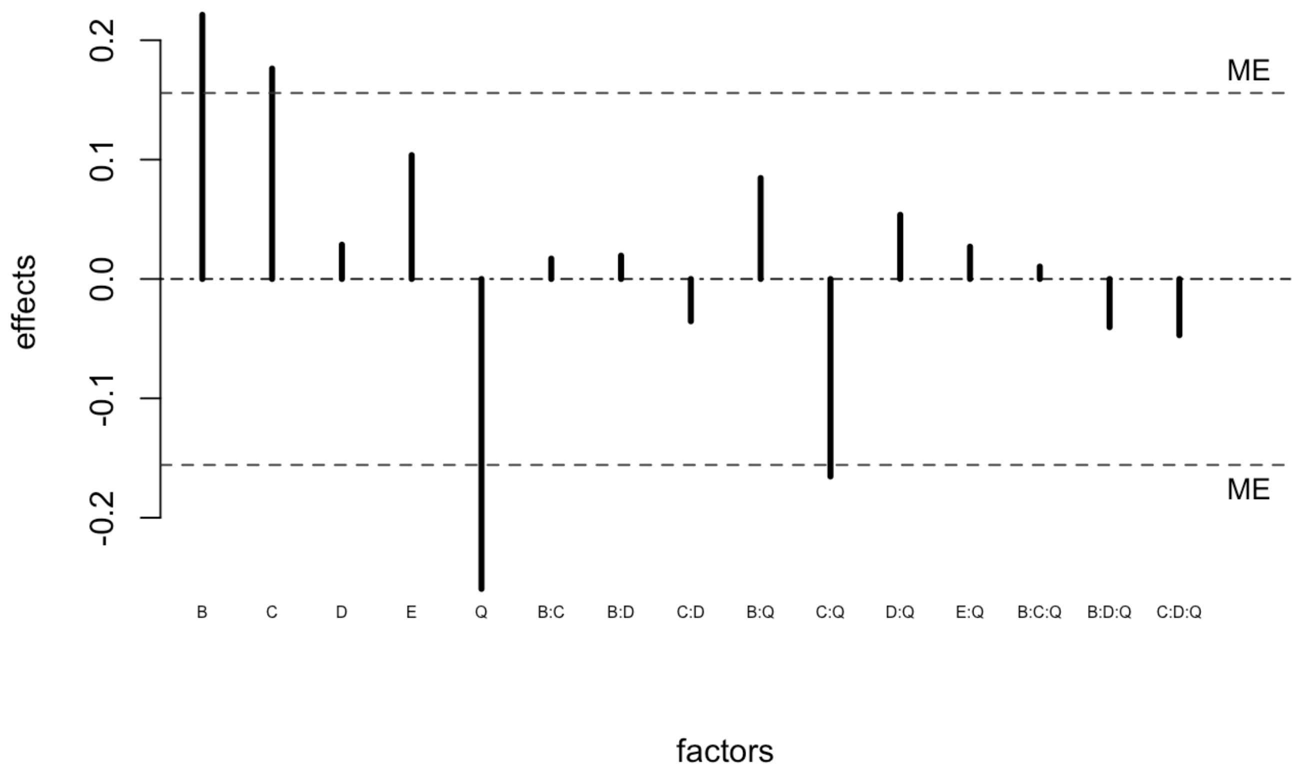
```
DanielPlot(fact.leaf,half = F)
```




```
DanielPlot(fact.leaf,half = T)
```




```
LenthPlot(fact.leaf2,cex.fac = 0.5)
```




```
    alpha       PSE        ME        SME
0.0500000 0.0606000 0.1557773 0.3162503
```

7.4 Example - Baking cookies

Consider a factorial design to study the effects of the amounts of three factors on the taste of chocolate chip cookies.

Factor	Amount
Butter	10g (-1), 15g (+1)
Sugar	1/2 cup (-1), 3/4 cup (+1)
Baking powder	1/2 teaspoon (-), 1 teaspoon (+)

Taste will be measured on a scale of 1 (poor) to 10 (excellent). A full factorial will require $2^3 = 8$ runs.

The 8 runs of the full factorial design tell the experimenter how to set the levels of the different ingrediants (factors).

Run	butter	sugar	powder
1	-1	-1	-1

2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1

In the first run the experimenter will bake cholate chip cookies with 10g butter, 1/2 cup sugar, and 1/2 teaspoon of baking powder; the second run will use 15g butter, 1/2 cup sugar, and 1/2 teaspoon of baking powder; etc.

But, the experimenter decides to also study baking time on taste, but can't afford to do more than 8 runs since each run requires a different batch of cookie dough. He wants to test if a baking time of 12 minutes versus 16 minutes has an impact on taste

So, he decides to use the three factor interaction between butter, sugar, and powder to **assign if baking time will be 12 minutes (-1) or 16 minutes (+1) in each of the 8 runs**.

Run	butter	sugar	powder	baking time
1	-1	-1	-1	-1
2	1	-1	-1	1
3	-1	1	-1	1
4	1	1	-1	-1
5	-1	-1	1	1
6	1	-1	1	-1
7	-1	1	1	-1
8	1	1	1	1

In this factorial design with four factors in 8 runs the experimenter will bake the cookies with 10g butter, 1/2 cup sugar, 1/2 teaspoon of baking powder, and baking time 12 minutes in the first run; in the second run use 15g butter, 1/2 cup sugar, and 1/2 teaspoon of baking powder, and 16 minutes baking time; etc.

Let A=butter, B=sugar, C=powder, D=baking time. But,

$$D = ABC$$

since he used the three factor interaction to assign baking times to each run. In other words D is **aliased** with ABC.

This type of design is called a 2^{4-1} fractional factorial design. Instead of using a full factorial or $2^4 = 16$ runs to study 4 factors we are using $\frac{1}{2}2^4 = 8$ runs.

The aliasing relation

$$D = ABC \Rightarrow I = ABCD,$$

where I is the column of $+1$s.

The aliasing realtion also means that other factors in the design have aliases. We can find these aliases by multiplying $I = ABCD$ by all the possible main and interaction effects.

$$A = BCD, B = ACD, C = ABD, D = ABC, AB = CD, AC = BD, BC = AD$$

All the main effects are aliased with three factor interactions, and two factor interactions with two factor interactions.

Suppose that the experimental runs were conducted and the following results were obtained.

Run	butter	sugar	powder	baking time	taste (y)
1	-1	-1	-1	-1	3
2	1	-1	-1	1	4
3	-1	1	-1	1	5
4	1	1	-1	-1	7
5	-1	-1	1	1	2
6	1	-1	1	-1	8
7	-1	1	1	-1	9
8	1	1	1	1	6

The factorial effect of baking time (D) is really the effect of $D + ABC$. In other words the effects of D and ABC are confounded. They cannot be seperatley estimated which is why ABC is called an alias of D.

This means that effect of D

$$\frac{1}{4}(-y_1 + y_2 + y_3 - y_4 + y_5 - y_6 - y_7 + y_8) = -2.5$$

is really the effect of $D + ABC$ or baking time plus the interaction of butter, sugar, and baking powder. In order for this to be equal to the effect of baking time (D) we must assume the three-way interaction (ABC) is small enough to be ignored (i.e., the factorial estimate of ABC is close to 0).

If the experimenter were to use a full factorial then he would require $2^4 = 16$ different batches of cookies. In a full 2^4 design he would be estimating 4 main effects, 6 two-way interactions, 4 three-way interactions, and 1 four-way interaction. If we assume that we can ignore three-factor and higher order interactions then a 16 run design is being used to estimate then a 16 run design is being used to estimate 10 effects. Fractional factorials use these redundancies by arranging that lower order effects are confounded with higher order interactions that are assumed negligable.