Normal Plot of Factorial Effects

$\hat{\theta}_1, \cdots, \hat{\theta}_I \overset{i.i.d.}{\sim}$ Normal ?

- Suppose $\hat{\theta}_i, i = 1, \cdots, I$, are the factorial effect estimates (example in Table 4, LNp.5-11).
  - Order them as $\hat{\theta}_{(1)} \le \cdots \le \hat{\theta}_{(I)}$.
  - **Normal probability plot** (LNp.3-28):

    $\hat{\theta}_{(i)}$ (vertical) vs. $\Phi^{-1}([i-0.5]/I)$ (horizontal)

$2\hat{\beta}_i$

\# of effects

quantile of $N(0,1)$

check linear trend

residuals

cf.

$\hat{\theta}_i$

$\hat{\beta}_i$

detect extreme values

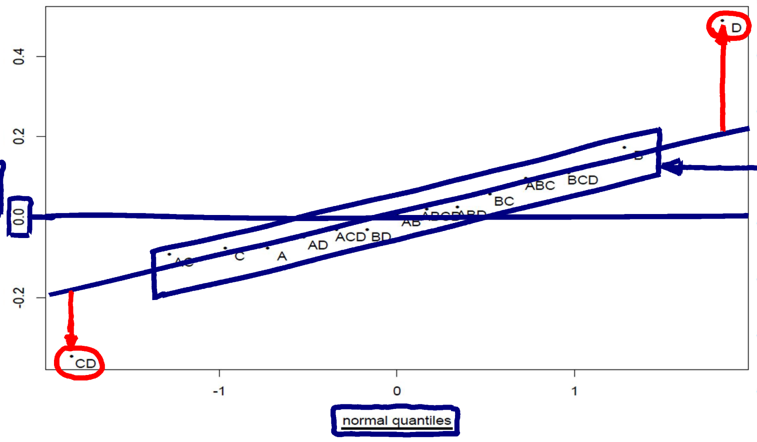


Figure 5: Normal Plot of Location Effects, Adapted Epitaxial Layer Growth Experiment

Recall.

① In unreplicated 2^k design,
⇒ no df left for σ^2
⇒ cannot do t-tests
⇒ cannot detect effect significance

② Normal (probability) plot for residual
⇒ check normality
⇒ detect outlier

★ Assumption: the effect parameters of these estimated factorial effects are zero.

★ Under the conceptual model and 2^k full factorial design

$$Z = X\beta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I), \text{ where}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_A \\ \beta_B \\ \vdots \\ \beta_{ABCD} \end{bmatrix} \Rightarrow \hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_A \\ \vdots \\ \hat{\beta}_{ABCD} \end{bmatrix} = (X^T X)^{-1} X^T Z \text{ and}$$

$$\frac{1}{2}\hat{\theta} = \hat{\beta} \sim N(\beta, (X^T X)^{-1}\sigma^2), \quad \text{where } (X^T X)^{-1} = \frac{1}{N} I$$

$\because$ orthogonality

2^k

⇒ $\hat{\beta}_0, \hat{\beta}_A, \hat{\beta}_B, \cdots, \hat{\beta}_{ABCD}$ are independent and have same variance

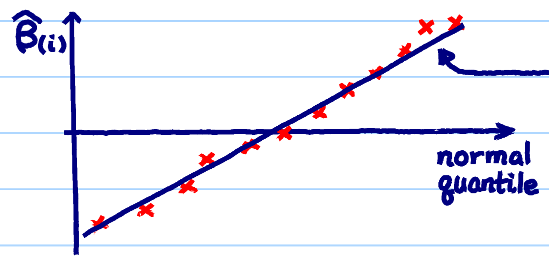
★ Under H_0: $\beta_{effect} = 0$ ⇒ $\hat{\beta}_{effect} \sim N(0, (\sigma^2/N) I)$

run size 2^k

⇒ $\hat{\beta}_A, \hat{\beta}_B, \cdots, \hat{\beta}_{ABCD} \overset{i.i.d.}{\sim} N(0, \sigma^2/N)$

⇒ The normal plot of $\hat{\beta}_A, \hat{\beta}_B, \cdots, \hat{\beta}_{ABCD}$ looks like

the required assumption for drawing normal plot



almost a straight line, and slope of the line $\approx \sigma/\sqrt{N}$

$\hat{\beta}$ contains information of σ.

Use of Normal Plot to Detect Effect Significance

- **Deduction Step.** Null hypothesis H_0 : all factorial effects = 0 . Under H_0, $\hat{\theta}_i \sim N(0, \sigma_*^2)$ and the resulting normal plot should follow a *straight line*.

effect sparsity → add the assumption in LNp.5-17

→ slope → σ_*

→ intercept → mean of $\hat{\theta}_i$'s

- **Induction Step.** By fitting a straight line to the middle group of points (around 0) in the normal plot, *any effect whose corresponding point falls off the line is declared significant* (Daniel, 1959).

$Var(\hat{\theta}_i)$, not the variance of the response.

- Unlike t- or F-test, no estimate of σ^2 is required. Method is especially suitable for *unreplicated* experiments. In t-test, s^2 (i.e., $\hat{\sigma}^2$) is the *reference quantity*. For unreplicated experiments, Daniel's idea is to use the *normal curve as the reference distribution*. → compare to the empirical cdf of $\hat{\theta}_i$

- In Figure 5 (LNp.5-17), D, CD (and possibly B?) are significant. Method is informal and judgemental.

graphical method ⇒ subjective

Normal and Half Normal Plots

Recall. In t-tests, declare significant if $|\text{t-value}| > c$

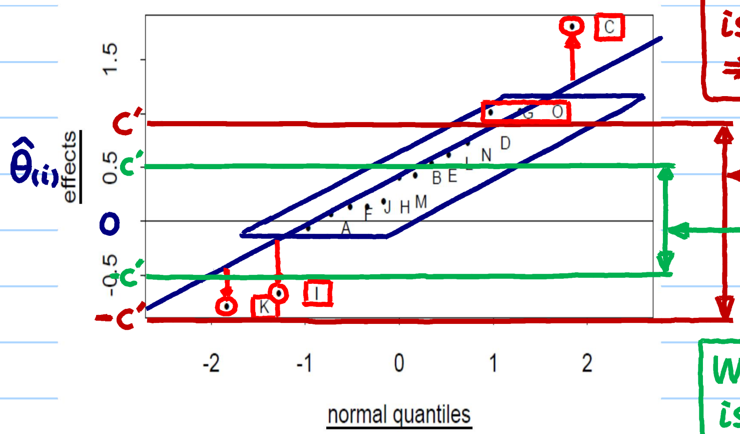
$\Leftrightarrow |\hat{\theta}| > c \cdot s.e.(\hat{\theta})$ $\;\llcorner \equiv c'$

$\Leftrightarrow |\hat{\theta}| > c'$

When # of replicates is small, ⇒ s.e.$(\hat{\theta})$ is large

Note. Slope of the line ∝ $\sigma/\sqrt{N}$ if assumption* in LNp.5-17 holds

When # of replicates is large, ⇒ s.e.$(\hat{\theta})$ is small



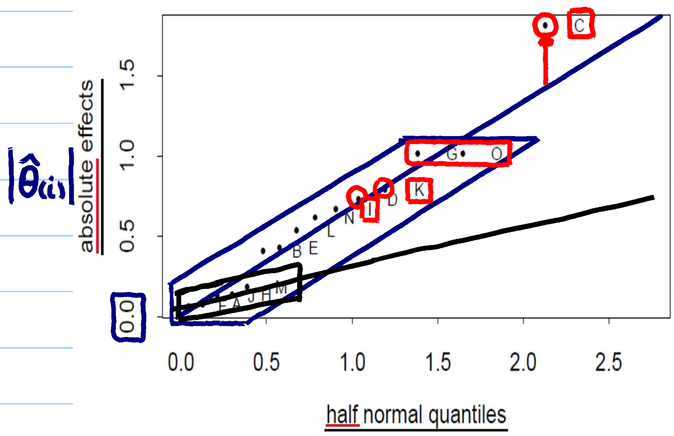


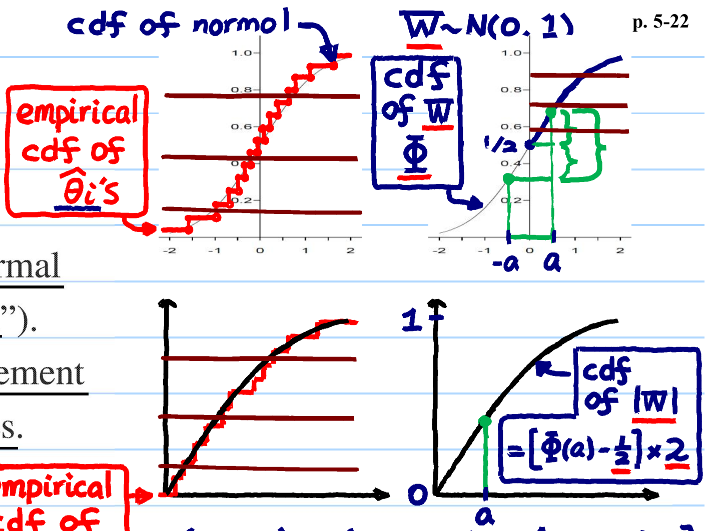
Figure 6: Comparison of Normal and Half-Normal Plots

Visual Misjudgement with Normal Plot

- **Potential misuse of normal plot :**
  - In Figure 6 (top, LNp.5-20), by following the procedure for detecting effect significance, one may declare C, K and I are significant, because they "deviate" from the middle straight line.
  - This is *wrong* because it ignores the obvious fact that K and I are smaller than G and O in magnitude. $|\cdot|$: absolute value
  - This points to a potential visual misjudgement and misuse with the normal plot.

such visual misjudgement appears more often in the normal plot of factorial effect estimates, but rarely happens in the normal plot of residuals (Why ? $\because \sum \hat{\varepsilon}_i = 0$)

Half-Normal Plot



$p(|W|<a) = p(-a<W<a) = 2[\Phi(a)-\Phi(0)]$

- **Idea:** Order the absolute $\hat{\theta}_i$ values as
$$|\hat{\theta}|_{(1)} \leq \cdots \leq |\hat{\theta}|_{(I)} .$$
Plot them on the positive axis of the normal distribution (thus the term "half-normal"). This would avoid the potential misjudgement between the positive and negative values.
- The **half-normal probability plot** consists of the points
$$(\Phi^{-1}(0.5+0.5[i-0.5]/I), |\hat{\theta}|_{(i)}), \quad (2)$$
for $i = 1, \ldots, 2^k - 1$.
- In Figure 6 (bottom, LNp.5-20), only C is declared significant. Notice that K and I no longer stand out in terms of the absolute values.

If we want to include more effects into the fitted model, we can remove the effects that have been identified as significant, and redraw a half-normal plot using the remaining effects

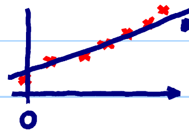
- For the rest of the book, **half-normal plots will be used for detecting effect significance.** → for exp'tal data with single replicate under saturate model

❖ **Reading**: textbook, 4.8

A Formal Test of Effect Significance: Lenth's Method

- Sometimes it is desirable to have a formal test that can assign p-values to the effects. The following method is also available in packages like SAS, JMP, or R.

Recall. ① normal plot / half-normal plot is subjective
② cannot do t-tests ($\because$ no df left to estimate σ)
③ If $W_1, \cdots, W_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, in the normal plot of $W_1, \cdots, W_n$



a straight line:
slope $\to \sigma$
intercept $\to \mu$

how?

Assume the t effects on/close to a straight line have parameters being zero, i.e., $\hat\theta_1, \cdots, \hat\theta_t \overset{i.i.d.}{\sim} N(0, \sigma_*^2)$ $\boxed{\text{not Var(z) or Var}(\epsilon)}$

$\text{Var}(\hat\theta)$

$\Rightarrow$ can use the concept in ③ to obtain information about σ_*^2.

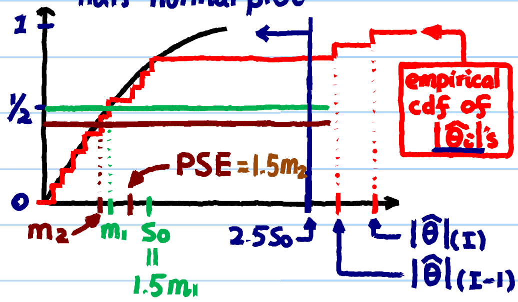
$\hat\sigma_* = \text{s.e.}(\hat\theta)$

how?

A Formal Test of Effect Significance (Contd.)

- Lenth's Method

why? must assume $\hat\theta_i$ with $|\hat\theta_i| < 2.5 s_0$ has mean zero



1. Compute the pseudo standard error (PSE)

$$\hat\sigma_*^{(2)} = PSE = 1.5 \times \underbrace{\text{median}_{\{|\hat\theta_i| < 2.5 s_0\}} |\hat\theta_i|}_{\equiv m_2},$$

where the median is computed among the $|\hat\theta_i|$ with $|\hat\theta_i| < 2.5 s_0$ and

$$\hat\sigma_*^{(1)} = s_0 = 1.5 \times \underbrace{\text{median} |\hat\theta_i|}_{\equiv m_1}.$$

$\bigstar$ $P(|W| > 2.5) \approx 0.01$, $W \sim N(0,1)$

⊖ Justification : If $\theta_i = 0$ and error is normal, s_0 is a *consistent* estimate of the standard deviation of $\hat\theta_i$. Use of median gives "robustness" to outlying values.

Why?

extremely large $|\hat\theta_i|$'s

$\bigstar$ If $\hat\theta_1, \cdots, \hat\theta_t \overset{i.i.d.}{\sim} N(0, \sigma_*^2)$, then median $\dfrac{|\hat\theta_i|\text{'s}}{\sigma_*} \xrightarrow{\text{estimate}}$ median$(|W|) = \Phi^{-1}\left(\frac{3}{4}\right) \approx \dfrac{1}{1.5}$, $W \sim N(0,1)$

check LN p.5-22

A Formal Test of Effect Significance (Contd.)

2. Compute $t\text{-statistic} = \frac{\hat{\theta} - 0}{s.e.(\hat{\theta})}$ cf. cf.

① test statistic → $t_{PSE,i} = \frac{\hat{\theta}_i - 0}{PSE}$, for each i.

If $|t_{PSE,i}|$ exceeds the ② critical value given in Appendix H (textbook, p.701) or from software, $\hat{\theta}_i$ is declared significant. obtained from simulation

- Two versions of the critical values are considered next.

Two Versions of Lenth's Method

- Null hypothesis. H_0 : all θ_i's $= 0$, normal error.
- **Individual Error Rate (IER)** $\xleftrightarrow{\text{cf.}}$ individual t-test

for some fixed i, a specific effect

For $i = 1, \cdots, I$, the IER_α at level α is determined by

$$Prob(|t_{PSE,i}| > \text{IER}_\alpha | H_0) \quad = \quad \alpha.$$

– Note : Because $\theta_i = 0$, $t_{PSE,i}$ has the *same* distribution under H_0 for all i. ← ∵ under H_0, $\hat{\theta}_i$'s are i.i.d. $\sim N(0, \sigma_*^2)$

for all effects

- **Experiment-wise Error Rate (EER)** $\xleftrightarrow{\text{cf.}}$ multiple testing

$$Prob(|t_{PSE,i}| > \text{EER}_\alpha \text{ for at least one } i,\ i = 1, \ldots, I | H_0)$$

← # of all effects

$$= Prob\left(\max_{1 \le i \le I} |t_{PSE,i}| > \text{EER}_\alpha \middle| H_0\right) = \alpha.$$

Recall. Tukey's method ← cf.

RR₁ RR₂ RR₃

- $\text{EER}_\alpha > \text{IER}_\alpha$.
- EER accounts for the number of tests done in the experiment but often gives conservative results (less powerful). In screening experiments, IER is more powerful and preferable because many of the θ_i's are negligible (recall the effect sparsity principle). The EER critical values can be inflated by considering many θ_i values. (Why?) ← many θ_i's are actually very small & need not be tested