# Normal Plot of Factorial Effects ← $\hat\theta_1, \cdots, \hat\theta_I \overset{i.i.d.}{\sim}$ Normal ?

$2\hat\beta_i$ — # of effects

- 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):

    quantile of $N(0,1)$

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

Recall.
① In unreplicated $2^k$ design,
⇒ no df left for $\sigma^2$
⇒ cannot do t-tests
⇒ cannot detect effect significance
② Normal (probability) plot for residual
⇒ check normality
⇒ detect outlier

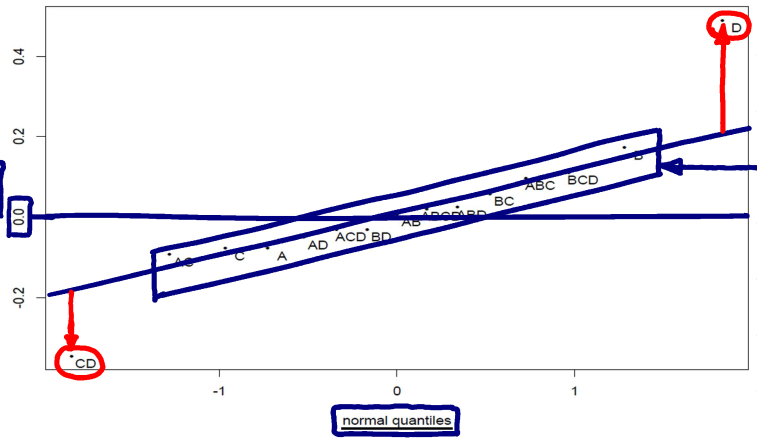
check linear trend

residuals

cf.

$\hat\theta_i$ $\hat\beta_i$

detect extreme values




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

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



☆ Under the conceptual model and $2^k$ full factorial design

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

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_A \\ \beta_B \\ \vdots \\ \beta_{ABCD} \end{bmatrix} \Rightarrow \hat\beta = \begin{bmatrix} \hat\beta_0 \\ \hat\beta_A \\ \vdots \\ \hat\beta_{ABCD} \end{bmatrix} = (X^TX)^{-1}X^TZ \text{ and}$$

$$\tfrac{1}{2}\hat\theta = \hat\beta \sim N(\beta, (X^TX)^{-1}\sigma^2), \quad \text{where } (X^TX)^{-1} = \tfrac{1}{N} I \leftarrow \boxed{\because \text{orthogonality}}$$

$N \leftarrow 2^k$

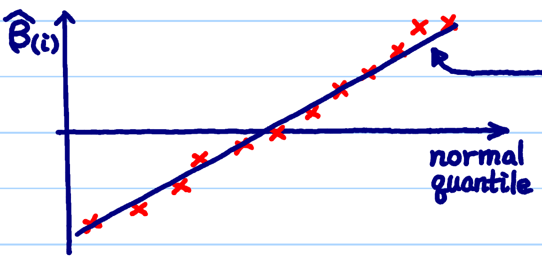
⇒ $\hat\beta_0, \hat\beta_A, \hat\beta_B, \cdots, \hat\beta_{ABCD}$ are independent and have same variance

☆ Under $H_0$: $\beta_{effect} = 0 \Rightarrow \hat\beta_{effect} \sim N(0, (\sigma^2/N) I)$ — run size $2^k$

⇒ $\hat\beta_A, \hat\beta_B, \cdots, \hat\beta_{ABCD} \overset{i.i.d.}{\sim} N(0, \sigma^2/N) \Leftrightarrow \hat\theta_{effect}$'s $\overset{i.i.d.}{\sim} N(0, 4\sigma^2/N)$, $4\sigma^2 = \sigma_*^2$

⇒ 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 $\sigma$.

# 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 → $\sigma_*$
→ 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 $\sigma^2$ 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 $\hat{\theta}_i$

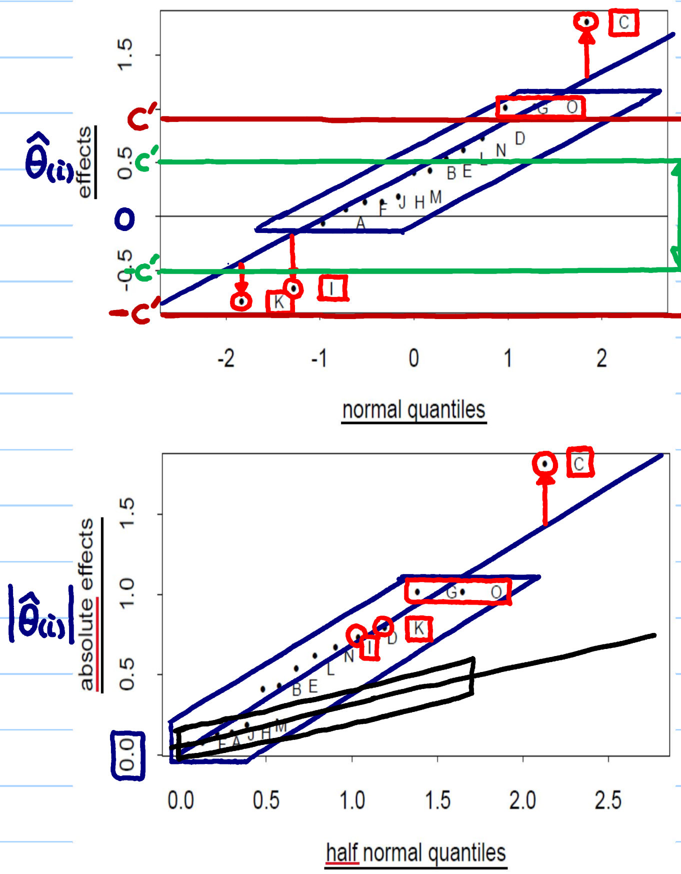
- In Figure 5 (LNp.5-17), $D$, $CD$ (and possibly $B$?) are significant. Method is informal and judgemental.

graphical method
⇒ subjective



# Normal and Half Normal Plots

Recall. In t-tests, declare significant if $|\text{t-value}| > c$.
$\Leftrightarrow |\hat{\theta}| > \underline{c \cdot s.e.(\hat{\theta})}$
$\quad \equiv c'$
$\Leftrightarrow |\hat{\theta}| > c'$




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

Note. Slope of the line ∝ $2\sigma/\sqrt{N}$ 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

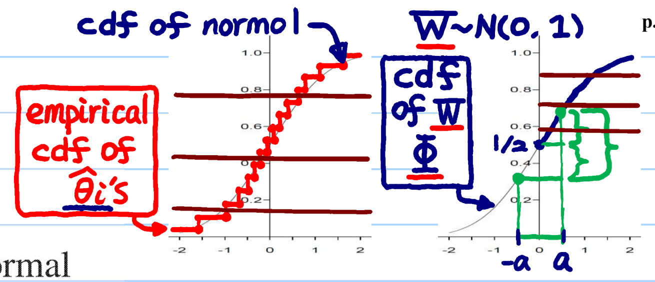
# Visual Misjudgement with Normal Plot

- **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.

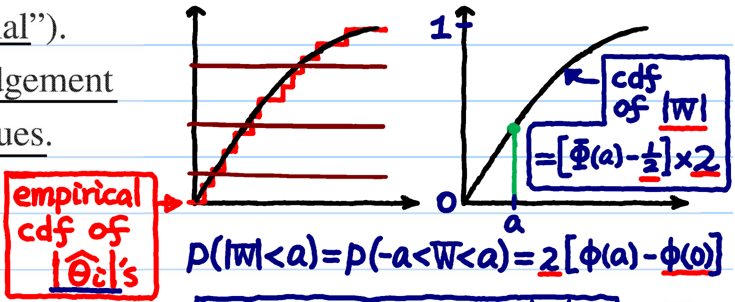
such visual misjudgement appears more often in the normal plot of factorial effect estimates, but rarely happens in the normal plot of residuals (Why? $\because \sum \hat{\varepsilon}_i = 0$)



# Half-Normal Plot




- **Idea:** ② Order the ① absolute $\hat{\theta}_i$ values as
$$|\hat{\theta}|_{(1)} \le \cdots \le |\hat{\theta}|_{(I)} .$$
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.




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

- The **half-normal probability plot** consists of the points
$$\left(\Phi^{-1}(0.5+0.5[i-0.5]/I), |\hat{\theta}|_{(i)}\right), \quad (2)$$
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

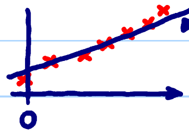
# A Formal Test of Effect Significance: Lenth's Method

- 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 $\sigma$)

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 $\to \mu$

Assume the $r$ effects on/close to a straight line have parameters being zero, i.e., $\hat{\theta}_1, \cdots, \hat{\theta}_r \overset{i.i.d.}{\sim} N(0, \sigma_*^2)$ [not Var(z) or Var($\bar{\epsilon}$)] how?

$\sigma_*^2$: $Var(\hat{\theta})$; $\hat{\sigma}_* = s.e.(\hat{\theta})$

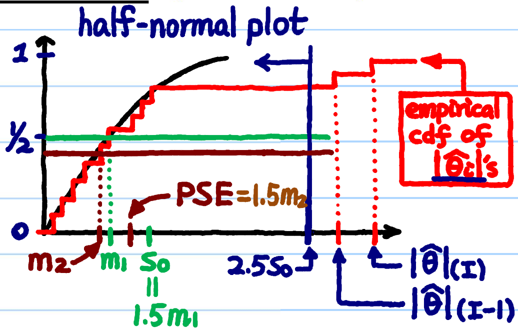
$\Rightarrow$ can use the concept in ③ to obtain information about $\sigma_*^2$. how?



# A Formal Test of Effect Significance (Contd.)

- Lenth's Method

1. Compute the pseudo standard error ($PSE$) — why? must assume $\hat{\theta}_i$ with $|\hat{\theta}_i| < 2.5 s_0$ has mean zero.




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

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

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

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

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

★ If $\hat{\theta}_1, \cdots, \hat{\theta}_r \overset{i.i.d.}{\sim} N(0, \sigma_*^2)$, 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 → $t_{PSE,i} = \dfrac{\hat{\theta}_i - 0}{PSE}$, 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 $\theta_i$'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 $\text{IER}_\alpha$ at level $\alpha$ 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 $i$.

$\because$ under $H_0$, $\hat{\theta}_i$'s are i.i.d. $\sim N(0, \sigma_*^2)$

for all effects

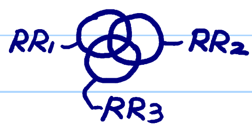
- **Experiment-wise Error Rate (EER)** $\xleftrightarrow{\text{cf.}}$ multiple testing

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

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

Recall. Tukey's method ← cf.

# of all effects




- $\text{EER}_\alpha > \text{IER}_\alpha$.
- 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 $\theta_i$'s are negligible (recall the effect sparsity principle). The EER critical values can be inflated by considering many $\theta_i$ values. (Why?) ← many $\theta_i$'s are actually very small & need not be tested

# Illustration with Adapted Epi-Layer Growth Experiment

1. In Table 4 (LNp.5-11),

closer to zero

- median$|\hat{\theta}_i| = 0.078$, ← median of 15 $|\hat{\theta}_i|$'s
- $s_0 = 1.5 \times 0.078 = 0.117$.
- trimming constant $2.5 s_0 = 2.5 \times 0.117 = 0.292$, which eliminates 0.490 ($D$) and 0.345 ($CD$). → 2 out of 15 effects
- $\text{median}_{\{|\hat{\theta}_i| < 2.5 s_0\}} |\hat{\theta}_i| = 0.058$ ← median of 13 smaller $|\hat{\theta}_i|$'s
- $PSE = 1.5 \times 0.058 = 0.087$ $\xleftrightarrow{\text{cf.}}$ s.e.$(\hat{\theta}_i)$

The corresponding $|t_{PSE}|$ values appear in Table 6 (LNp.5-28).

2. Choose $\alpha = 0.01$.

larger

- $\text{IER}_{0.01} = 3.63$ for $I = 15$. By comparing with the $|t_{PSE}|$ values, $D$ and $CD$ are significant at 0.01 level.
- $\text{EER}_{0.01} = 6.45$ (for $I = 15$). No effect is detected as significant.

- Analysis of the $|t_{PSE}|$ values for $\ln s^2$ (Table 6, LNp.5-28) detects no significant effect (details on textbook, p.182), thus confirming the half-normal plot analysis in Figure 4.10 of section 4.8 (textbook, p.179).



# $|t_{PSE}|$ Values for Adapted Epi-Layer Growth Experiment

Table 6: $|t_{PSE}|$ Values, Adapted Epitaxial Layer Growth Experiment

| Effect | $\bar{y}$ | $\ln s^2$ |
|---|---|---|
| $A$ | 0.90 | 0.25 |
| $B$ | 1.99 | 1.87 |
| $C$ | 0.90 | 1.78 |
| $D$ | $\boxed{5.63}$ | 0.89 |
| $AB$ | 0.09 | 0.71 |
| $AC$ | 1.07 | 0.41 |
| $AD$ | 0.57 | 0.46 |
| $BC$ | 0.67 | 1.27 |
| $BD$ | 0.34 | 0.16 |
| $CD$ | $\boxed{3.97}$ | 1.35 |
| $ABC$ | 1.13 | 0.51 |
| $ABD$ | 0.29 | 0.67 |
| $ACD$ | 0.34 | 0.00 |
| $BCD$ | 1.26 | 0.05 |
| $ABCD$ | 0.23 | 1.63 |

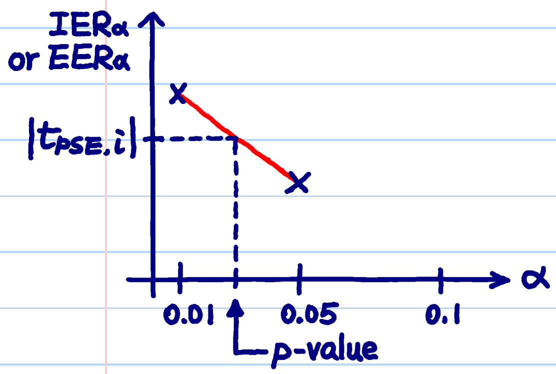
unreplicated

$y_x \sim N(\mu_x, \sigma_x^2)$

$\boxed{E(y_x)}$ → $\mu_x \approx \hat{\beta}_0 + \hat{\beta}_D x_D + \hat{\beta}_{CD} x_{CD}$

$\Rightarrow$ $\sigma_x^2 \approx$ a constant

$\boxed{Var(y_x)}$




- $p$-values of effects can be obtained from packages or by interpolating the critical values in the tables in appendix H (textbook, p.701). (See textbook, p.182 for illustration).

❖ **Reading**: textbook, 4.9

# Nominal-the-Best Problem ← Recall. objective in LNp.5-1

- There is a nominal or target value $t$ (=14.5 in the case) based on engineering design requirements.
- Define a quantitative loss due to deviation of $y_{\mathbf{x}}$ from $t$.

**Quadratic loss**: $L(y_{\mathbf{x}},t) = \cancel{c} \cdot (y_{\mathbf{x}} - t)^2$. $\boxed{-E(y_{\mathbf{x}})+E(y_{\mathbf{x}})}$

$\boxed{\underset{\mathbf{x}}{\text{argmin}}}$ → $E(L(y_{\mathbf{x}},t)) = \cancel{c} \cdot Var(y_{\mathbf{x}}) + \cancel{c} \cdot [E(y_{\mathbf{x}}) - t]^2$.
↑ variance ↑ bias²

$\boxed{\mu_{\mathbf{x}} = E(y_{\mathbf{x}}) \approx X_1\hat{\beta}_1}$

$\boxed{\ln\sigma^2_{\mathbf{x}} = \ln(Var(y_{\mathbf{x}})) \approx X_2\hat{\beta}_2}$

- **Two-step procedure for nominal-the-best problem**:
  (i) Select levels of some factors to minimize $Var(y_{\mathbf{x}})$.
  (ii) Select the level of a factor not in (i) to move $E(y_{\mathbf{x}})$ closer to $t$.
  - A factor in step (ii) is an *adjustment* factor if it has a significant effect on $E(y_{\mathbf{x}})$ but not on $Var(y_{\mathbf{x}})$. ← $\boxed{\text{a factor whose effects appears in } \hat{\beta}_1\text{, but not } \hat{\beta}_2}$
  - Procedure is effective only if an adjustment factor can be found. This is often done on engineering ground.
    → If an adjustment factor does not exist, some trade-off between (i) and (ii) is required.
  - Examples of adjustment factors : deposition time in surface film deposition process, mold size in tile fabrication, location and spacing of markings on the dial of a weighing scale.

❖ **Reading**: textbook, 4.10



# Why Take $\ln s^2$ ? → as a response of linear model

- It maps $s^2$ over $(0,\infty)$ to $\ln s^2$ over $(-\infty,\infty)$.
  Regression and ANOVA assume the responses are nearly normal, i.e. over $(-\infty,\infty)$.
- Better for variance prediction.
  - Suppose $z_{\mathbf{x}} = \ln s^2_{\mathbf{x}}$.
  - $\hat{z}_{\mathbf{x}}$ = predicted value of $\ln\sigma^2_{\mathbf{x}}$
  - $e^{\hat{z}_{\mathbf{x}}}$ = predicted value of $\sigma^2_{\mathbf{x}}$, always nonnegative.

response $\xleftarrow{\text{approximate}}$ linear structure
$s^2 \in (0,\infty)$
$\ln(s^2) \in (-\infty,\infty)$ ← $X\beta + \varepsilon \in (-\infty,\infty)$

- Most physical laws have a multiplicative component.
  Log converts *multiplicity* into *additivity*.
- Variance stabilizing property: next slide.
  → produce "constant" error variance in the model: $\ln s^2_{\mathbf{x}} = X\beta + \varepsilon$

→ error part of $y$:
$\varepsilon = \varepsilon_1 \times \cdots \times \varepsilon_k$
Assume independence and zero means:
$Var(\varepsilon) = Var(\varepsilon_1) \times \cdots \times Var(\varepsilon_k)$ ← cf.
→ $\ln(Var(\varepsilon)) = \sum_{i=1}^{k} \ln(Var(\varepsilon_i))$

$\boxed{\text{may get a better approximation by linear structure } X\beta \text{ and follow normal.}}$

# $\ln s^2$ as a Variance Stabilizing Transformation

- Assume $y_{\mathbf{x},j} \overset{i.i.d.}{\sim} N(\mu_{\mathbf{x}}, \sigma^2_{\mathbf{x}})$, $j = 1, \ldots, n_{\mathbf{x}}$. Then,

$$(n_{\mathbf{x}} - 1)s^2_{\mathbf{x}} = \sum_{j=1}^{n_{\mathbf{x}}} (y_{\mathbf{x},j} - \bar{y}_{\mathbf{x}})^2 \sim \sigma^2_{\mathbf{x}} \chi^2_{n_{\mathbf{x}}-1},$$

and

$$\left.\begin{array}{l} E(\ln S^2_{\mathbf{x}}) = ? \\ Var(\ln S^2_{\mathbf{x}}) = ? \end{array}\right\} \leftarrow \overbrace{\ln s^2_{\mathbf{x}}}^{\text{a random variable}} = \ln \sigma^2_{\mathbf{x}} + \ln\Big(\underbrace{\overbrace{\chi^2_{n_{\mathbf{x}}-1}}^{\text{a random variable} \sim \chi^2_{n_{\mathbf{x}}-1}}/(n_{\mathbf{x}} - 1)}_{W_{\mathbf{x}}}\Big). \tag{3}$$

- $W_{\mathbf{x}}$: a random variable, $h$: a smooth function, by $\delta$-method,

$$E(h(W_{\mathbf{x}})) \approx h(E(W_{\mathbf{x}})) \quad \text{and} \quad Var(h(W_{\mathbf{x}})) \approx [h'(E(W_{\mathbf{x}}))]^2 Var(W_{\mathbf{x}}) \tag{4}$$

- Suppose $W_{\mathbf{x}} \sim \chi^2_{\nu_{\mathbf{x}}}/\nu_{\mathbf{x}}$. Then, $E(W_{\mathbf{x}}) = 1$ and $Var(W_{\mathbf{x}}) = 2/\nu_{\mathbf{x}}$.
- Take $h = \ln$. Applying (4) to $W_{\mathbf{x}}$ $(\sim \chi^2_{\nu_{\mathbf{x}}}/\nu_{\mathbf{x}})$ leads to

$$E(\ln(W_{\mathbf{x}})) \approx \ln(E(W_{\mathbf{x}})) = \ln(1) = 0, \rightarrow E(\ln S^2_{\mathbf{x}}) \approx \ln \sigma^2_{\mathbf{x}}$$

$$Var(\ln(W_{\mathbf{x}})) \approx [h'(1)]^2 (2/\nu_{\mathbf{x}}) = 2/\nu_{\mathbf{x}}. \rightarrow Var(\ln S^2_{\mathbf{x}}) \approx 2/\nu_{\mathbf{x}}$$

In LM, $Z_{\mathbf{x}} \sim N(\mu_{\mathbf{x}}, \sigma^2)$

cf.

- In (3), $\nu_{\mathbf{x}} = n_{\mathbf{x}} - 1$, we have $\ln s^2_{\mathbf{x}} \sim N(\ln \sigma^2_{\mathbf{x}}, 2(n_{\mathbf{x}} - 1)^{-1})$. The variance of $\ln s^2_{\mathbf{x}}$, i.e., $2(n_{\mathbf{x}} - 1)^{-1}$, is nearly constant for $n_{\mathbf{x}} - 1 \geq 9$.

a constant if we have same # of replicates for each run.

$\sigma^2_{\mathbf{x}}$ only appears in the mean structure of $\ln(S^2_{\mathbf{x}})$

❖ **Reading**: textbook, 4.11 — (exercise) Compare the result with $E(S^2_{\mathbf{x}})$ and $Var(S^2_{\mathbf{x}})$



# Epi-layer Growth Experiment Revisited

- Original data from Shoemaker, Tsui and Wu (1991) ★Conceptual model:

$$Z_{\mathbf{x}} \sim \beta_0 + \beta_A x_A + \cdots$$
$$+ \beta_{AB} x_A x_B + \cdots$$
$$+ \beta_{ABC} x_A x_B x_C + \cdots$$
$$+ \beta_{ABCD} x_A x_B x_C x_D$$
$$+ \varepsilon \leftarrow \text{constant variance}$$

Table 2 (LNp.5-2) cf. ↔ Table 7: Design Matrix and Thickness Data, Original Epitaxial Layer Growth Experiment

$2^4$ design

original response (6 replicates) — unreplicated

| Design A | B | C | D | Thickness ($y$) | | | | | | $\bar{y}$ | $s^2$ | $\ln s^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| − | − | − | + | 14.812 | 14.774 | 14.772 | 14.794 | 14.860 | 14.914 | 14.821 | 0.003 | −5.771 |
| − | − | − | − | 13.768 | 13.778 | 13.870 | 13.896 | 13.932 | 13.914 | 13.860 | 0.005 | −5.311 |
| − | − | + | + | 14.722 | 14.736 | 14.774 | 14.778 | 14.682 | 14.850 | 14.757 | 0.003 | −5.704 |
| − | − | + | − | 13.860 | 13.876 | 13.932 | 13.846 | 13.896 | 13.870 | 13.880 | 0.001 | −6.984 |
| − | + | − | + | 14.886 | 14.810 | 14.868 | 14.876 | 14.958 | 14.932 | 14.888 | 0.003 | −5.917 |
| − | + | − | − | 14.182 | 14.172 | 14.126 | 14.274 | 14.154 | 14.082 | 14.165 | 0.004 | −5.485 |
| − | + | + | + | 14.758 | 14.784 | 15.054 | 15.058 | 14.938 | 14.936 | 14.921 | 0.016 | −4.107 |
| − | + | + | − | 13.996 | 13.988 | 14.044 | 14.028 | 14.108 | 14.060 | 14.037 | 0.002 | −6.237 |
| + | − | − | + | 15.272 | 14.656 | 14.258 | 14.718 | 15.198 | 15.490 | 14.932 | 0.215 | −1.538 |
| + | − | − | − | 14.324 | 14.092 | 13.536 | 13.588 | 13.964 | 14.328 | 13.972 | 0.121 | −2.116 |
| + | − | + | + | 13.918 | 14.044 | 14.926 | 14.962 | 14.504 | 14.136 | 14.415 | 0.206 | −1.579 |
| + | − | + | − | 13.614 | 13.202 | 13.704 | 14.264 | 14.432 | 14.228 | 13.907 | 0.226 | −1.487 |
| + | + | − | + | 14.648 | 14.350 | 14.682 | 15.034 | 15.384 | 15.170 | 14.878 | 0.147 | −1.916 |
| + | + | − | − | 13.970 | 14.448 | 14.326 | 13.970 | 13.738 | 13.738 | 14.032 | 0.088 | −2.430 |
| + | + | + | + | 14.184 | 14.402 | 15.544 | 15.424 | 15.036 | 14.470 | 14.843 | 0.327 | −1.118 |
| + | + | + | − | 13.866 | 14.130 | 14.256 | 14.000 | 13.640 | 13.592 | 13.914 | 0.070 | −2.653 |

① Use $Z_{\mathbf{x}} = \bar{y}_{\mathbf{x}}$ to build a mean model

(Q: Why not use $Z_{\mathbf{x}} = y_{\mathbf{x}}$ to build a mean model?

Ans. $Var(y_{\mathbf{x}})$ not regarded constant over $\mathbf{x}$)

(Q: Why $\bar{y}_{\mathbf{x}}$ better than $y_{\mathbf{x}}$?

Ans. $0 < Var(y_{\mathbf{x}}) = \sigma^2_{\mathbf{x}}$

$0 < Var(\bar{y}_{\mathbf{x}}) = \sigma^2_{\mathbf{x}}/6$)

② Use $Z_{\mathbf{x}} = \ln S^2_{\mathbf{x}}$ to build a variance model

# Epi-layer Growth Experiment: Effect Estimates

Table 8: Factorial Effects, Original Epitaxial Layer Growth Experiment

$2\hat{\beta}_i = \hat{\theta}_i$

$(X^TX)^{-1}X^TZ$

$16I$

| Effect | $\bar{y}$ | $\ln s^2$ |
|---|---|---|
| $A$ | −0.055 | 3.834 |
| $B$ | 0.142 | 0.078 |
| $C$ | −0.109 | 0.077 |
| $D$ | 0.836 | 0.632 |
| $AB$ | −0.032 | −0.428 |
| $AC$ | −0.074 | 0.214 |
| $AD$ | −0.025 | 0.002 |
| $BC$ | 0.047 | 0.331 |
| $BD$ | 0.010 | 0.305 |
| $CD$ | −0.037 | 0.582 |
| $ABC$ | 0.060 | −0.335 |
| $ABD$ | 0.067 | 0.086 |
| $ACD$ | −0.056 | −0.494 |
| $BCD$ | 0.098 | 0.314 |
| $ABCD$ | 0.036 | 0.109 |

Understand the meaning of these effect estimates, e.g.,

$ME(A) = -0.055 = \bar{Z}(A+) - \bar{Z}(A-)$

$INT(A,B) = -0.032 = ME(A|B+) - ME(A|B-)$

⋮



# Epi-layer Growth Experiment: Half-Normal Plots

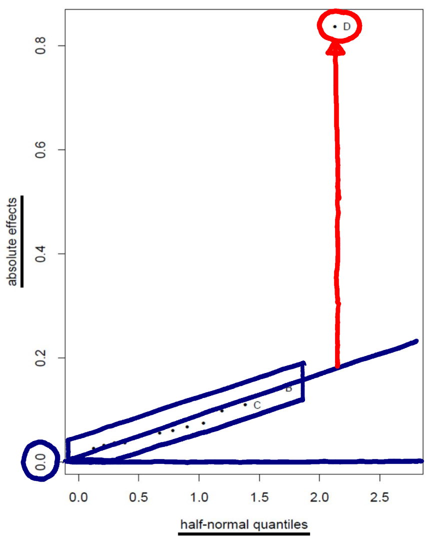



Figure 7 : Location effects ($\bar{y}_{\mathbf{x}}$)

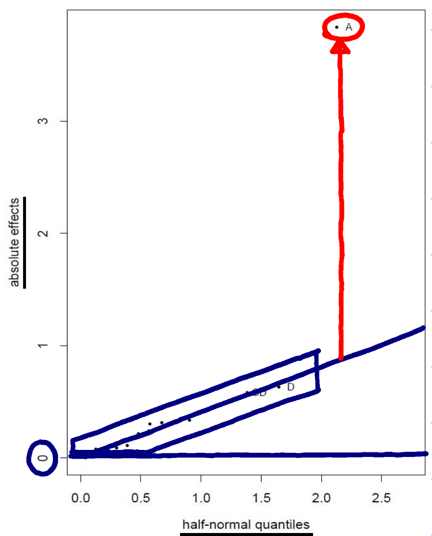



Figure 8 : Dispersion effects ($\ln s^2_{\mathbf{x}}$)

(exercise) Perform Lenth's method to identify significant effects

# Epi-layer Growth Experiment: Analysis and Optimization

positive sign ⇒ increasing effect

- From the two plots, ME $D$ is significant for $z = \bar{y}$ and ME $A$ is significant for $z = \ln s^2$. Fitted models :

$\mathrm{ME}(D)/2$

used to reach the target value, i.e., $E(y_x) \approx 14.5$

location model → $\hat{y}_{\mathbf{x}} = \hat{\beta}_0 + \hat{\beta}_D x_D = 14.389 + 0.418 x_D,$

dispersion model → $\ln \hat{s}^2_{\mathbf{x}} = \hat{\gamma}_0 + \hat{\gamma}_A x_A = -3.772 + 1.917 x_A.$ → used to minimize variance

$\mathrm{ME}(A)/2$

  – Factor $D$ is an adjustment factor.
  – All the effects related to factors $B$ and $C$ do not appear in the fitted models.

B,C could be important variables with small exp'tal range

useful information?

- Two-step procedure: qualitative factor
  (i) Choose $A$ at $-$ level (continuous rotation). ← Why not set A=−2,−3,...

  quantitative factor

  (ii) Choose $x_D = 0.266$ to satisfy $14.5 = 14.389 + 0.418 x_D$. (If $D$ = 30 and 40 sec for $x_D = -1$ and $+1$, $x_D = 0.266$ corresponds to $35 + 0.266(5) = 36.33$ sec.) ← cf.

- Predicted variance (Q: What if D is a qualitative factor?)

$\ln(\hat{\sigma}^2) = -5.695$ ← $\hat{\sigma}^2 = \exp[-3.772 + 1.917(-1)] = (0.058)^2 \approx 0.0034$ ← check LNp.5-32

This is too optimistic! Predicted values should be validated with a *confirmation experiment*.

(Note. $s^2_{\mathbf{x}}$ estimates subplot error variance) ← check LNp.5-3

| optimal setting | | | |
|---|---|---|---|
| $x_A$ | $x_B$ | $x_C$ | $x_D$ |
| $-1$ | ? | ? | 0.266 |

Q: Do we have information about the whole-plot error variance?
Ans. Check Fig.7 in LNp.5-34

❖ **Reading**: textbook, 4.12