ω (H_0): $y=\beta_0+\epsilon$
Ω (H_A): $y=\beta_0+\beta_1 x+\beta_2 x^2+\epsilon$

overall F-test

One-Way ANOVA (Contd)

same ANOVA table when laser power is treated as a qualitative factor

Table 3: ANOVA Table, Composite Experiment

Source	Degrees of Freedom	Sum of Squares	Mean Squares	F
laser	2	224.184	112.092	11.32
residual	6	59.422	9.904	
total	8	283.606		

$\hat{\sigma}^2$

large $\Rightarrow$ small p-value

cf. "how different" problems for qualitative and quantitative factors

- Conclusion from ANOVA : Laser power has a significant effect on strength.

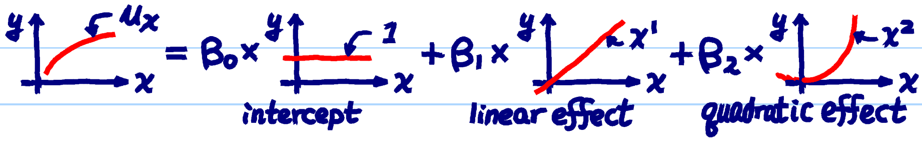
There exist difference between $\mu_{40}, \mu_{50}, \mu_{60}$

- To further understand the effect, use of multiple comparisons is not useful here. (Why?)

can only answer whether $\mu_i \overset{?}{=} \mu_j$ for $i,j\in\{40,50,60\}$

- The effects of a quantitative factor like laser power can be decomposed into linear, quadratic, etc.

base functions



coding/base functions for quantitative factors (cf. qualitative codings, LN p.3-8)

Linear and Quadratic Effects

orthogonal polynomial $\leftarrow$ (LM, LN p.8-4~5)

$y=\beta_0+\beta_\ell x_\ell+\beta_q x_q+\epsilon$

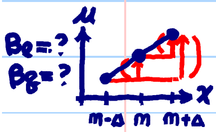
	x_ℓ	x_q	
$m-\Delta$	-1	1	$\to \mu_L$
m	0	-2	$\to \mu_M$
$m+\Delta$	1	1	$\to \mu_H$

equally spaced $\to$ $m-\Delta$ $\quad m \quad m+\Delta$

- Suppose there are three levels of x (low, medium, high) and the corresponding $E(y_x)$ values are $\mu=(\mu_L,\mu_M,\mu_H)^T$.

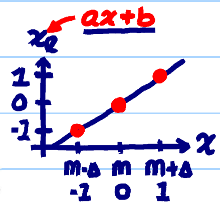
$$\underset{\sim}{\mu}=\begin{bmatrix}\mu_L\\ \mu_M\\ \mu_H\end{bmatrix}=\begin{bmatrix}\beta_0-\beta_\ell+\beta_q\\ \beta_0\quad -2\beta_q\\ \beta_0+\beta_\ell+\beta_q\end{bmatrix}=\underbrace{\begin{bmatrix}1&-1&1\\1&0&-2\\1&1&1\end{bmatrix}}_{A}\underset{\sim}{\beta}\Rightarrow\underset{\sim}{\beta}=\begin{bmatrix}\beta_0\\ \beta_\ell\\ \beta_q\end{bmatrix}=A^{-1}\underset{\sim}{\mu}=\underbrace{\begin{bmatrix}(\mu_L+\mu_M+\mu_H)/3\\ (-\mu_L\quad+\mu_H)/2\\ (\mu_L-2\mu_M+\mu_H)/6\end{bmatrix}}_{(\Delta)}$$

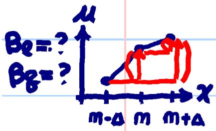
$A^TA=\begin{bmatrix}3&0&0\\0&2&0\\0&0&6\end{bmatrix}$, $A^{-1}=\begin{bmatrix}1/3&&0\\&1/2&\\0&&1/6\end{bmatrix}A^T$



Linear contrast : $\mu_H-\mu_L=(-1,0,1)\begin{pmatrix}\mu_L\\ \mu_M\\ \mu_H\end{pmatrix}$.

$\propto \beta_\ell$



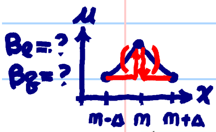


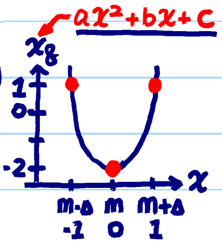
Quadratic contrast : $\mu_L-2\mu_M+\mu_H=(1,-2,1)\begin{pmatrix}\mu_L\\ \mu_M\\ \mu_H\end{pmatrix}$.

$\propto \beta_q$

$=2\times\left(\frac{\mu_L+\mu_H}{2}-\mu_M\right)$ $\leftarrow$ for qualitative (Helmert coding)

$=(\mu_H-\mu_M)-(\mu_M-\mu_L)$ $\leftarrow$ for quantitative





$(-1,0,1)$ and $(1,-2,1)$ are the linear and quadratic contrast vectors; they are **orthogonal** to each other. $\to$ & orthogonal to intercept $\mathbb{1}$.

$\underset{\sim}{a}$ s.t. $\underset{\sim}{a}^T\mathbb{1}=0$ (LN p.3-5) $\Rightarrow \underset{\sim}{a}$: a contrast

when # of replicates are identical for all treatments (balance), the columns in model matrix are orthogonal.

Linear and Quadratic Effects (Contd.)

$x_\ell \leftarrow$ coding $\rightarrow x_q$

- Using $(-1, 0, 1)$ and $(1, -2, 1)$, we can write a more detailed regression model $y = X\beta + \epsilon$, where the model matrix X is given below.

LM, LNp.8-2, location & scale change

- *Normalization* : Length of $(-1,0,1) = \sqrt{2}$, length of $(1,-2,1) = \sqrt{6}$, divide each vector by its length in the regression model. (Why ? It provides a *consistent* comparison of the regression coefficients. But the t-statistics in the next table are independent of such (and any) scaling.)

$\hat\beta_i \leftrightarrow t_i$ consistent

- Normalized contrast vectors:

linear: $(-1,0,1)/\sqrt{2} = (-1/\sqrt{2}, 0, 1/\sqrt{2})$,

quadratic: $(1,-2,1)/\sqrt{6} = (1/\sqrt{6}, -2/\sqrt{6}, 1/\sqrt{6})$.

vectors with different lengths

$$\begin{array}{c} x=40 \\ \\ \\ x=50 \\ \\ \\ x=60 \\ \\ \\ \end{array}\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{32} \\ y_{33} \end{bmatrix} = X\underbrace{\beta}_{\beta_0+\beta_\ell x_\ell+\beta_q x_q} + \epsilon \qquad X = \begin{array}{ccc} & x_\ell & x_q \\ \end{array}\begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 0 & -2 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} \hat\beta_0 \\ \hat\beta_\ell \\ \hat\beta_q \end{bmatrix} = (X^TX)^{-1}X^TY \leftarrow \text{check } (\triangle) \text{ in LNp.3-16}$$

$$X^TX = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 18 \end{bmatrix}$$

$$(X^TX)^{-1} = \begin{bmatrix} 1/9 & 0 & 0 \\ 0 & 1/6 & 0 \\ 0 & 0 & 1/18 \end{bmatrix} \Rightarrow \mathrm{cov}(\hat\beta) = \sigma^2(X^TX)^{-1}$$

Estimation of Linear and Quadratic Effects

- Let β_0^*, β_l^*, β_q^* denote respectively the intercept, the linear effect and the quadratic effect based on normalized contrasts and let $\beta = (\beta_0^*, \beta_l^*, \beta_q^*)'$. An estimator $\hat\beta$ of β is given by

$\mu = A\beta$

$A^TA = I_3 \Rightarrow A^T = A^{-1}$

$\beta = A^{-1}\mu = A^T\mu$

$$\hat\beta = \begin{pmatrix} \hat\beta_0^* \\ \hat\beta_l^* \\ \hat\beta_q^* \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} \bar y_{1\cdot} \rightarrow \mu_1 \\ \bar y_{2\cdot} \rightarrow \mu_2 \\ \bar y_{3\cdot} \rightarrow \mu_3 \end{pmatrix}$$

cf. $(\triangle)$ in LNp.3-16

- We can write $\hat\beta = \mathbf{A}'\bar{\mathbf{y}}$, where

$$\mathbf{A} = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix} \quad \text{and} \quad \bar{\mathbf{y}} = \begin{pmatrix} \bar y_{1\cdot} \\ \bar y_{2\cdot} \\ \bar y_{3\cdot} \end{pmatrix}$$

$$E(y) = \beta_0 + \beta_\ell x_\ell + \beta_q x_q = \underbrace{(\sqrt{3}\beta_0)}_{\beta_0^*}\frac{1}{\sqrt{3}} + \underbrace{(\sqrt{2}\beta_\ell)}_{\beta_\ell^*}\frac{x_\ell}{\sqrt{2}} + \underbrace{(\sqrt{6}\beta_q)}_{\beta_q^*}\frac{x_q}{\sqrt{6}}$$

- Since the columns of $\mathbf{A}$ constitute a set of orthonormal vectors, i.e. $\mathbf{A}'\mathbf{A} = I_3$. Let $X = [\mathbf{A}' \cdots \mathbf{A}']'$. We have

repeat k times

\# of replicates

$$\hat\beta = \mathbf{A}'\bar{\mathbf{y}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\bar{\mathbf{y}} = (X'X)^{-1}X'Y, \quad \frac{1}{k}X^TY$$

$\mathrm{cov}(\hat\beta) = \sigma^2 \cdot \frac{1}{k} I_3$

$X'X = kI_3$ (k: \# of replicates) $\Rightarrow (X^TX)^{-1} = \frac{1}{k}(A^TA)^{-1}$

where X (9×3 matrix) is the model matrix and Y (9×1 vector) is the response vector. This shows that $\hat\beta$ is *identical to the least squares estimate of* β.

- Running a multiple linear regression with response y and predictors x_l and x_q, we get $\hat\beta_0^* = 31.0322$, $\hat\beta_l^* = 8.636$, $\hat\beta_q^* = -0.381$.

Tests for Linear and Quadratic Effects

t-test in LM

Q: What if we use linear coding : $(40, 50, 60) = x'$ quadratic coding : $(40^2, 50^2, 60^2) = x^2$

exercise

Q: Do we need to worry about "collinearity"?

Table 4: Tests for Polynomial Effects, Composite Experiment

consistent

Effect	Estimate	Standard Error	t	p-value
linear	$\hat\beta_\ell^* \to$ 8.636	$\boxed{1.817}$	4.75	0.003 V
quadratic	$\hat\beta_q^* \to$ −0.381	$\boxed{1.817}$	−0.21	0.841 ← not significant

$\widehat{cov}(\hat{\beta}) = \hat\sigma^2 (X^T X)^{-1} = \hat\sigma^2 \begin{bmatrix} 1/3 & & 0 \\ & 1/3 & \\ 0 & & 1/3 \end{bmatrix}$

$\hat\sigma^2$ = RSS/n-p (check Table3, LNp.15)

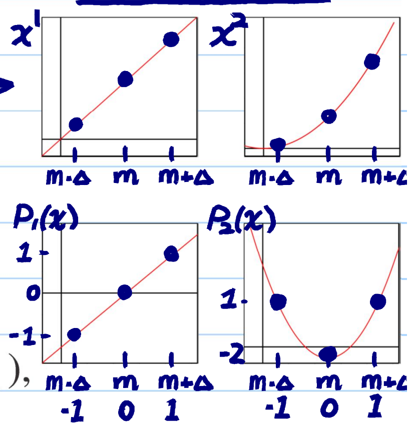
identical s.e. (why?)

- Further conclusion : Laser power has a significant linear (but not quadratic) effect on strength. ← answer to "how different" problem
- Another question : How to predict y-value (strength) at a setting not in the experiment (i.e., other than 40, 50, 60) ? Need to extend the concept of linear and quadratic contrast vectors to cover a *whole interval* for x. This requires building a model using polynomials.

LM, LNp. 8-3~6

Orthogonal Polynomials

base functions



x_q: 1 −2 1
x_ℓ: −1 0 1

- For three evenly spaced levels $m-\Delta$, m, and $m+\Delta$, define the first and second degree polynomials :

x', x² might have strong collinearity

$$P_1(x) = \frac{x-m}{\Delta},$$

($= -1, 0$ and 1, for $x = m-\Delta, m, m+\Delta$),

$$P_2(x) = 3\left[\left(\frac{x-m}{\Delta}\right)^2 - \frac{2}{3}\right] \quad (= 1, -2 \text{ and } 1, \text{ for } x = m-\Delta, m, m+\Delta) .$$

$E(y_x) = \beta_0^* + \beta_1^* \frac{P_1(x)}{\sqrt{2}} + \beta_2^* \frac{P_2(x)}{\sqrt{3}}$ ↑ defined for any x

Therefore, $P_1(x)$ and $P_2(x)$ are extensions of the linear and quadratic contrast vectors. (Why ?) ← cf. functional form & matrix form of LM

- **Polynomial regression model :**

$$y = \beta_0^* + \beta_1^* \times P_1(x)/\sqrt{2} + \beta_2^* \times P_2(x)/\sqrt{6} + \varepsilon, \xleftrightarrow{\text{cf.}} \beta_0^* + \beta_1^* x_\ell/\sqrt{2} + \beta_2^* x_q/\sqrt{6}$$

obtain regression (i.e., least squares) estimates $\hat\beta_0^* = 31.03$, $\hat\beta_1^* = 8.636$, $\hat\beta_2^* = -0.381$. (Note : $\hat\beta_1^*$ and $\hat\beta_2^*$ values are same as in Table 4).

Prediction based on Polynomial Regression Model

- Fitted model:

$$\widehat{E(y_x)} = \hat{\mu}_x = 31.0322 + 8.636 \times P_1(x)/\sqrt{2} - 0.381 \times P_2(x)/\sqrt{6},$$

can be any x

functional form

cf.

- To predict $\hat{\mu}_x$ at any $x = x^*$, plug in the x^* on the right side of the regression equation. For $x = 55$, because $m = 50$, $\Delta = 10$,

$$P_1(55) = \frac{55-50}{10} = \frac{1}{2},$$

cf. interpolation / extrapolation

model uncertainty : model might not be close to the 2nd-order model ouside the exp'tal region

$$P_2(55) = 3\left[\left(\frac{55-50}{10}\right)^2 - \frac{2}{3}\right] = -\frac{5}{4},$$

matrix form

$x_0 = (1, P_1(x^*)/\sqrt{2}, P_2(x^*)/\sqrt{6})^T$

$\hat{\mu}_{x^*} = x_0^T \hat{\beta}$

Also, predition for future obs.

$$\begin{aligned}\hat{\mu}_{55} &= 31.0322 + 8.636(0.5/\sqrt{2}) - 0.381(-1.25/\sqrt{6}) \\ &= 34.2803.\end{aligned}$$

point estimation

can apply LM technique to construct confidence interval for mean response / future obs.

s.e.

$\widehat{Var}(x_0^T\hat{\beta}) = \hat{\sigma}^2 x_0^T (X^TX)^{-1} x_0$

s.e.

$\widehat{Var}(x_0^T\hat{\beta}) + \widehat{Var}(\varepsilon)$ future error

$= \hat{\sigma}^2[x_0^T (X^TX)^{-1} x_0 + 1]$