- It is common to use the t-test and the t-statistic

$$t_{ij} = \frac{\bar{y}_{i\cdot} - \bar{y}_{j\cdot} - 0}{\hat{\sigma}\sqrt{1/n_i + 1/n_j}},$$

$t_{ij}^2 = F_{ij}$

$\mathrm{Var}(\bar{y}_{i\cdot} - \bar{y}_{j\cdot}) = \mathrm{Var}(\bar{y}_{i\cdot}) + \mathrm{Var}(\bar{y}_{j\cdot}) = \sigma^2/n_i + \sigma^2/n_j$

where n_i = number of observations for treatment i, $\hat{\sigma}^2 = \mathrm{RSS}_\Omega/\mathrm{df}_\Omega$ in ANOVA; declare "treatments i and j different at level α" if

$$|t_{ij}| > t_{N-k,\frac{\alpha}{2}}.$$

$F_{1,\mathrm{df}_\Omega} \to$ null dist.

df_Ω

null of ANOVA in LNp.3-6

R_{ij} = rejection region of H_0^{ij}

$P(R_{ij} \mid \mu_i = \mu_j \ (\tau_i = \tau_j)) = \alpha$

multiple testing

R_{12}, R_{13}, R_{23}

EER $= P(\bigcup R_{ij} \mid \mu_1 = \cdots = \mu_k) > \alpha$ ← why not good?

usually

- Suppose k' tests are performed to test $H_0: \tau_1 = \cdots = \tau_k$. *Experimentwise error rate* (EER) = Probability of declaring at least one pair of treatments significantly different under H_0. Need to use multiple comparisons to control EER.

e.g. $\binom{k}{2}$

$\bigcap_{i,j} H_0^{ij}$ ← union-intersection test

H_1: $\bigcup_{i,j} H_1^{ij}$

D C A B

Note. deductive logic does not hold (Why?)

$t_{16,\frac{0.05}{2}} = 2.12$

t_{ij}

A vs. B	A vs. C	A vs. D	B vs. C	B vs. D	C vs. D
−0.87	1.85	2.14	2.72	3.01	0.29

$|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > \hat{\sigma}\sqrt{\frac{1}{n_i} + \frac{1}{n_j}} \cdot t_{N-k,\frac{\alpha}{2}}$ ← a scale

RA

$\bar{y}_{(1)\cdot}$ $\bar{y}_{(2)\cdot}$ ⋯ $\bar{y}_{(k-1)\cdot}$ $\bar{y}_{(k)\cdot}$

$\mu_{(1)}$ $\mu_{(2)}$ ⋯ $\mu_{(k-1)}$ $\mu_{(k)}$

$\mu_{(1)}, \cdots, \mu_{(k)}$: sorted $\mu_1, \cdots, \mu_k$ $(\mu_{(1)} \le \mu_{(2)} \le \cdots \le \mu_{(k)})$

$\bar{y}_{(1)\cdot}, \cdots, \bar{y}_{(k)\cdot}$: order statistics of $\bar{y}_{1\cdot}, \cdots, \bar{y}_{k\cdot}$.

Bonferroni Method

- Declare "τ_i different from τ_j at level α" if $|t_{ij}| > t_{N-k,\frac{\alpha}{2k'}}$, where k' = number of tests.

cf. $\frac{\alpha}{2}$

EER ≤ α

very conservative when k' is large

- For one-way layout with k treatments, $k' = \binom{k}{2} = \frac{1}{2}k(k-1)$, as k increases, k' increases, and the critical value $t_{N-k,\frac{\alpha}{2k'}}$ gets bigger (i.e., method less powerful in detecting differences).

- Advantage: It works without requiring independence assumption.

some indep. assumptions btwn the test stat. of these tests, e.g., indep. $\bar{y}_{i\cdot}$'s

can be widely applied.

- For pulp experiment, take $\alpha = 0.05$, $k = 4$, $k' = 6$, $t_{16,0.05/12} = 3.008$. Among the 6 t_{ij}-values (see LNp.3-10), only the t-value for B-vs-D, 3.01, is larger. Declare "B and D different at level 0.05".

2×6

$$P\left(\bigcup R_{ij} \,\middle|\, \bigcap H_0^{ij}\right) \le \sum_{i,j} P(R_{ij} \mid H_0^{ij}) = k' \cdot \alpha' = \alpha$$

$$\Rightarrow \alpha' = P(R_{ij} \mid H_0^{ij}) = \alpha/k'$$

D C A B

For simplicity, assume $n_1=\cdots=n_k=n$

Tukey Method

Same procedure can be applied to unequal sample size case → Tukey-Kramer test

- Declare "τ_i different from τ_j at level α" if

$$|t_{ij}| > \frac{1}{\sqrt{2}}\, q_{k,N-k,\alpha},$$

$\because \sqrt{1/n+1/n} = \sqrt{2}/\sqrt{n}$ (LNp.3-10)

EER ≤ α

where $q_{k,N-k,\alpha}$ is the upper α-quantile of the **Studentized range** (**SR**) distribution with parameter k and $N-k$ degrees of freedom. (see distribution table on LNp.3-13)

- For pulp experiment,

$\sqrt{RSS_\Omega/df_\Omega}$ (LNp.3-9)

LNp.3-13

$$\frac{1}{\sqrt{2}}\, q_{k,N-k,0.05} = \frac{1}{\sqrt{2}}\, q_{4,16,0.05} = \frac{4.05}{\sqrt{2}} = 2.86.$$

Again only B-vs-D has larger t_{ij}-value than 2.86 (See LNp.3-10). Tukey method is more powerful than Bonferroni method because 2.86 is smaller than 3.01 (why?)

- compare $\bar{y}_{i\cdot}-\bar{y}_{j\cdot}$, $\forall (i,j)$ ← check LNp.3-10
- For $1 \le i < j \le k$, $|\bar{y}_{i\cdot}-\bar{y}_{j\cdot}| \le \bar{y}_{(k)\cdot}-\bar{y}_{(1)\cdot}$.
- $P(\cup R_{ij} \mid \cap H_0^{ij})$ ← EER
  $= P(\text{at least one } |\bar{y}_{i\cdot}-\bar{y}_{j\cdot}| \text{ larger than } c^* \mid \cap H_0^{ij})$
  $= P(\bar{y}_{(k)\cdot}-\bar{y}_{(1)\cdot} > c^* \mid \cap H_0^{ij})$
  $= \alpha$
- Under $\cap H_0^{ij}$, $\bar{y}_{1\cdot}, \ldots, \bar{y}_{k\cdot} \overset{\text{indep.}}{\sim} N(\mu, \sigma^2/n) \Rightarrow$
  $\dfrac{\sqrt{n}(\bar{y}_{(k)\cdot}-\bar{y}_{(1)\cdot})}{\hat{\sigma}} \sim SR_{k,\nu}$
  where $\nu\hat{\sigma}^2 \sim \sigma^2\chi^2_\nu$ and indep. of $\bar{y}_{1\cdot}, \ldots, \bar{y}_{k\cdot}$.
  $\Rightarrow c^* = \hat{\sigma}/\sqrt{n} \cdot SR_{k,\nu,\alpha}$
  $\Rightarrow \sqrt{n}|\bar{y}_{i\cdot}-\bar{y}_{j\cdot}|/\hat{\sigma} > SR_{k,\nu,\alpha}$

Q: For the case in LNp.3-10, after getting $\bar{y}_{i\cdot}$'s, want to test $H_0^{ij}: \mu_D=\mu_B$ (only one test). Which critical value should we use? $H_0: \mu_{(1)}=\mu_{(k)}$?

Q: Among the 3 critical values of $|t_{ij}|$, which one is the smallest? the largest?

Selected values of $q_{k,\nu,\alpha}$ for $\alpha = 0.05$

ν = N-k

increasing (why?)

ν	k = 2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59	51.96	53.20	54.33	55.36
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39	14.75	15.08	15.38	15.65
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72	9.95	10.15	10.35	10.52
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03	8.21	8.37	8.52	8.66
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32	7.47	7.60	7.72
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65	6.79	6.92	7.03	7.14
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43	6.55	6.66	6.76
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18	6.29	6.39	6.48
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98	6.09	6.19	6.28
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72	5.83	5.93	6.03	6.11
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71	5.81	5.90	5.98
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51	5.61	5.71	5.80	5.88
13	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43	5.53	5.63	5.71	5.79
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46	5.55	5.64	5.71
15	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31	5.40	5.49	5.57	5.65
16	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26	5.35	5.44	5.52	5.59

α=upper tail probability, ν=degrees of freedom, k=number of treatments

decreasing (why?)

For complete tables corresponding to various values of α refer to Appendix E.

❖ **Reading**: textbook, 2.2

One-Way ANOVA with a Quantitative Factor

cf. Qualitative

★ response: bonding strength
★ (treament) factor: power (quantitative)
3 levels — 40, 50, 60
equally spaced
★ Exp'tal units: one composite
9 EUs
★ Each treatment repeats 3 times (3 replicates)

- **Data** :

Design matrix

power	strength
40	25.66
40	28.00
40	20.65
50	⋮
50	
50	
60	
60	
60	35.66

y = bonding strength of composite material,

x = laser power at 40, 50, 60 watt.

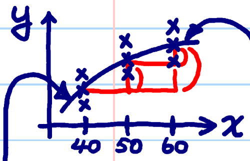
homogeneous → EU1 ··· EU9

$\frac{9!}{3!3!3!}$ ← completely randomized → r → 40, 50, 60 ← r

Table 2: Strength Data, Composite Experiment

Laser Power (watts)		
40	50	60
25.66	29.15	35.73
28.00	35.09	39.56
20.65	29.79	35.66

Initial Data Analysis: scatter plot



the line is meaningful only when the factor is quantitative

major difference between quantitative & qualitative factors

conceptual model

$\mu_x = E(y_x)$

$y_x = \mu_x + \varepsilon$

$\mu_x = \beta_0 + \beta_1 x$

★ $\mu_x = \beta_0 + \beta_1 x + \beta_2 x^2$

$\mu_x = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$

$\mu_x = \beta_0 + \beta_1 x \log x + \beta_2 e^x$

⋮

One-Way ANOVA (Contd)

ω (H₀): $y = \beta_0 + \varepsilon$

Ω (H$_A$): $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$

overall F-test

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 ⇒ small p-value

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

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

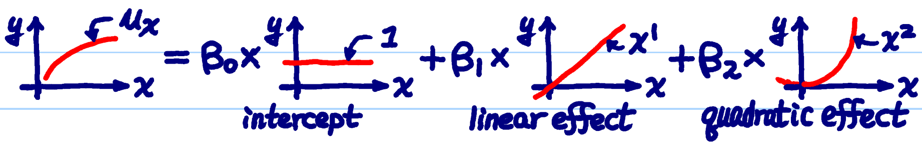
- Conclusion from ANOVA : Laser power has a significant effect on strength.

- 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, LNp.3-8)

Linear and Quadratic Effects

orthogonal polynomial ← (LM, LNp.8-4~5)

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

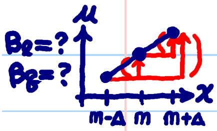
	x_ℓ	x_q	
m−Δ	−1	1	→ μ_L
m	0	−2	→ μ_M
m+Δ	1	1	→ μ_H

equally spaced → m−Δ m m+Δ

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

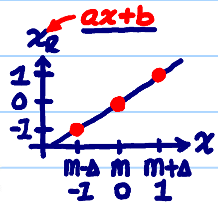
$$\underset{\sim}{\mu} = \begin{bmatrix}\mu_L\\ \mu_M\\ \mu_H\end{bmatrix} = \begin{bmatrix}\beta_0 - \beta_\ell + \beta_q\\ \beta_0 \quad -2\beta_q\\ \beta_0 + \beta_\ell + \beta_q\end{bmatrix} = \underbrace{\begin{bmatrix}1 & -1 & 1\\ 1 & 0 & -2\\ 1 & 1 & 1\end{bmatrix}}_{=A} \underset{\sim}{\beta}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

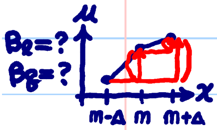
$$A^TA = \begin{bmatrix}3 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 6\end{bmatrix} \qquad A^{-1} = \begin{bmatrix}1/3 & & 0\\ & 1/2 & \\ 0 & & 1/6\end{bmatrix} A^T$$



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

$\propto \beta_\ell$

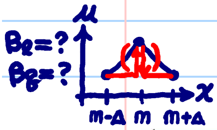


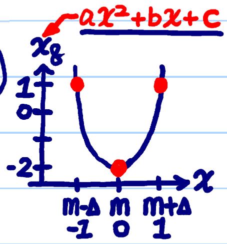


Quadratic contrast : $\mu_L - 2\mu_M + \mu_H = (1, -2, 1)\begin{pmatrix}\mu_L\\ \mu_M\\ \mu_H\end{pmatrix}$.

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

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





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

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

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

Linear and Quadratic Effects (Contd.)

x_ℓ ← coding → x_q

- Using $(-1, 0, 1)$ and $(1, -2, 1)$, we can write a more detailed regression model $y = X\beta + \varepsilon$, 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.)

$\hat{\beta}_i \longleftrightarrow t_i$ 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

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

$$\hat{\underset{\sim}{\beta}} = \begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_\ell\\ \hat{\beta}_q\end{bmatrix} = (X^TX)^{-1}X^TY \leftarrow \text{check } (\Delta) \text{ in LNp.3-16}$$

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

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

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. (Δ) in LNp.3-16

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

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

$$E(y) = \beta_0 + \beta_l x_l + \beta_q x_q = (\sqrt{3}\beta_0)\frac{1}{\sqrt{3}} + (\sqrt{2}\beta_l)\cdot\frac{x_l}{\sqrt{2}} + (\sqrt{6}\beta_q)\frac{x_q}{\sqrt{6}}$$
$$\sqrt{3}\beta_0 = \beta_0^*,\quad \sqrt{2}\beta_l = \beta_l^*,\quad \sqrt{6}\beta_q = \beta_q^*$$

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

# of replicates

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

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

$X'X = kI_3$, 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 : $(40, 50, 60) = x^l$, 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	$\hat{\beta}_l^* \rightarrow$ 8.636	1.817	4.75	0.003 V
quadratic	$\hat{\beta}_q^* \rightarrow$ −0.381	1.817	−0.21	0.841 ← not significant

identical s.e. (why?)

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

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

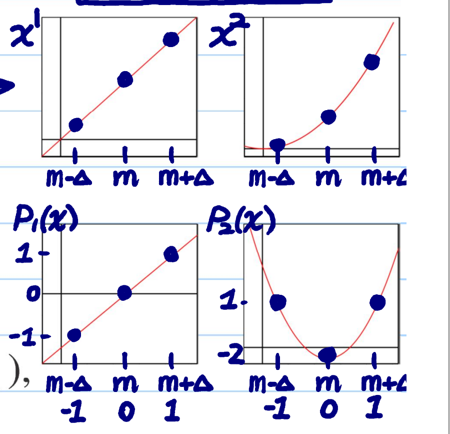
- Further conclusion : Laser power has a significant linear (but not quadratic) effect on strength. ← answer to "how different" problem

- 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_ℓ: -1 0 1
x_q: 1 -2 1

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

x^1, x^2 might have strong collinearity

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

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

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

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.

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

- **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:

functional form

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

cf.

↑ can be any x

- 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 $m = 50$, $\Delta = 10$,

interpolation

cf. extrapolation

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

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

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

matrix form

$\underset{\sim}{x}_0 = (1, P_1(x^*)/\sqrt{2}, P_2(x^*)/\sqrt{6})^T$

$\hat{\mu}_{x^*} = \underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}}$

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.

$\widehat{Var}(\underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}}) + \widehat{Var}(\varepsilon)$ ← future error

$= \hat{\sigma}^2 [\underset{\sim}{x}_0^T (X^T X)^{-1} \underset{\sim}{x}_0 + 1]$ →s.e.→ future obs.

s.e. $\widehat{Var}(\underset{\sim}{x}_0^T \hat{\underset{\sim}{\beta}}) = \hat{\sigma}^2 \underset{\sim}{x}_0^T (X^T X)^{-1} \underset{\sim}{x}_0$ → mean response