One-way layout and ANOVA: An Example

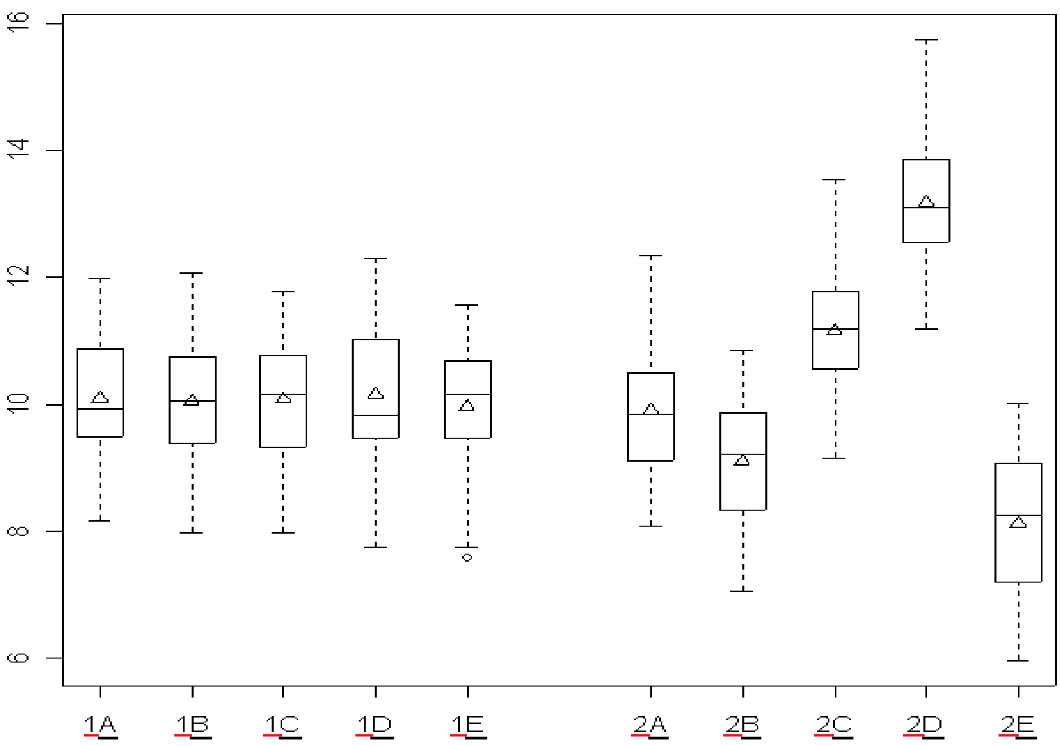
Reflectance data in pulp experiment: each of four operators made five pulp sheets; reflectance was read for each sheet using a brightness tester.

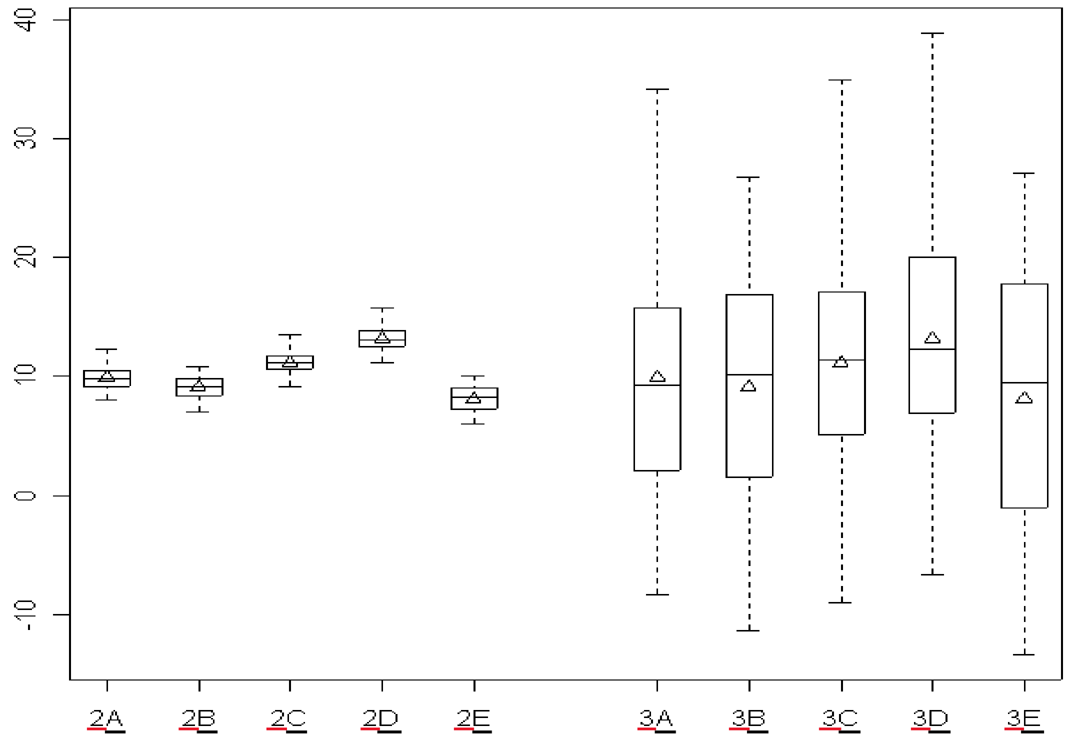
Randomization : assignment of 20 containers of pulp to operators and order of reading.

Table 1: Reflectance Data, Pulp Experiment

Operator			
A	B	C	D
59.8	59.8	60.7	61.0
60.0	60.2	60.7	60.8
60.8	60.4	60.5	60.6
60.8	59.9	60.9	60.5
59.8	60.0	60.3	60.5

Objective : determine if there are differences among operators in making sheets and reading brightness.





Model and ANOVA

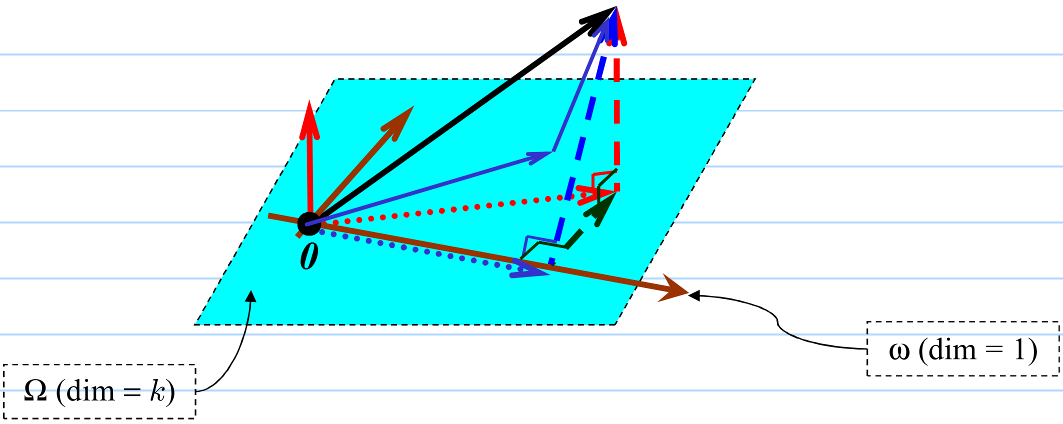
Model :

$$y_{ij} = \eta + \tau_i + \varepsilon_{ij}, \quad i = 1, \ldots, k; \; j = 1, \ldots, n_i,$$

where y_{ij} = jth observation with treatment i,

τ_i = ith treatment effect,

ε_{ij} = error, independent $N(0, \sigma^2)$.



Model fit:

$$y_{ij} = \hat{\eta} + \hat{\tau}_i + r_{ij}$$
$$= \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.}),$$

where "." means average over the particular subscript.

ANOVA Decomposition :

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{..})^2 = \sum_{i=1}^{k} n_i(\bar{y}_{i.}-\bar{y}_{..})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i.})^2.$$

F-Test

ANOVA Table

Source	Degrees of Freedom (df)	Sum of Squares	Mean Squares	Expected MS
treatment	$k-1$	$SSTr = \sum_{i=1}^{k} n_i(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$	$MSTr = SSTr/df$	$E_\Omega(MSTr)$
residual	$N-k$	$SSE = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i\cdot})^2$	$MSE = SSE/df$	$E_\Omega(MSE)$
total	$N-1$	$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{\cdot\cdot})^2$		

The F statistic for the null hypothesis that there is no difference between the treatments, i.e.,

$$H_0 : \tau_1 = \cdots = \tau_k,$$

is

$$F = \frac{\sum_{i=1}^{k} n_i(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2/(k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i\cdot})^2/(N-k)} = \frac{MSTr}{MSE},$$

which has an F distribution with parameters $k-1$ and $N-k$.

ANOVA for Pulp Experiment

Source	Degrees of Freedom (df)	Sum of Squares	Mean Squares	F
operator	3	1.34	0.447	4.20
residual	16	1.70	0.106	
total	19	3.04		

- $Prob(F_{3,16} > 4.20) = 0.02$ = p-value,
  thus declaring a significant operator-to-operator difference at level 0.02.

- Further question: among the $6 = \binom{4}{2}$ pairs of operators, what pairs show significant difference?
  Answer: Need to use multiple comparisons.

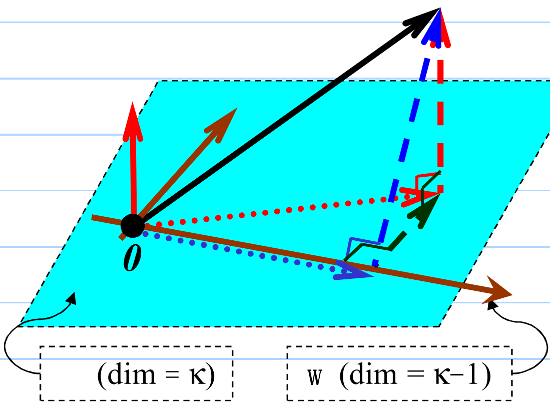
Constraint on the Parameters ι's

- The model in LNp.3-4 has κ distinct levels, but κ+1 regression parameters
  over-parameterized $\Xi^T\Xi$ singular unidentifiable
  cannot estimate parameters (**Q**: but why can do overall *F*-test?)
- Some common constraint on ι's
  - $\sum_{i=1}^{k} \tau_i = 0$ dummy variables: sum coding
  - $\tau_1 = 0$ dummy variables: treatment coding

❖ **Reading**: textbook, 2.1

Multiple Comparisons

- Consider the full model $y_{ij} = \eta + \tau_i + \varepsilon_{ij}$. For one pair, say (i, j), of treatments, test $H_0^{ij} : \tau_i = \tau_j$ against $H_A^{ij} : \tau_i \neq \tau_j$.



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

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

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

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

- 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.

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

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.

- 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.
- 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".

Tukey Method

- Declare "τ_i different from τ_j at level α" if
$$|t_{ij}| > \frac{1}{\sqrt{2}} q_{k,N-k,\alpha},$$
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,
$$\frac{1}{\sqrt{2}} q_{k,N-k,0.05} = \frac{1}{\sqrt{2}} q_{4,16,0.05} = \frac{4.05}{\sqrt{2}} = 2.86.$$
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?)

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

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

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

❖ **Reading**: textbook, 2.2

One-Way ANOVA with a Quantitative Factor

- **Data :**

  y = bonding strength of composite material,

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

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

One-Way ANOVA (Contd)

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		

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

Linear and Quadratic Effects

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

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

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

$(-1, 0, 1)$ and $(1, -2, 1)$ are the linear and quadratic contrast vectors; they are **orthogonal** to each other.

Linear and Quadratic Effects (Contd.)

- 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.

- *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.)

- Normalized contrast vectors:
  linear: $(-1,0,1)/\sqrt{2} = (-1/\sqrt{2}, 0, 1/\sqrt{2})$,
  quadratic: $(1,-2,1)/\sqrt{6} = (1/\sqrt{6}, -2/\sqrt{6}, 1/\sqrt{6})$.

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

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0^* \\ \hat{\beta}_l^* \\ \hat{\beta}_q^* \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} \bar{y}_{1\cdot} \\ \bar{y}_{2\cdot} \\ \bar{y}_{3\cdot} \end{pmatrix}$$

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

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

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

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

  where $\boldsymbol{X}$ is the model matrix and $\boldsymbol{Y}$ 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 $\hat{\beta}_0^* = 31.0322$, $\hat{\beta}_l^* = 8.636$, $\hat{\beta}_q^* = -0.381$.

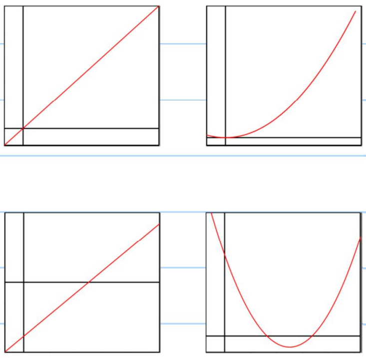
Tests for Linear and Quadratic Effects

Table 4: Tests for Polynomial Effects, Composite Experiment

Effect	Estimate	Standard Error	t	p-value
linear	8.636	1.817	4.75	0.003
quadratic	−0.381	1.817	−0.21	0.841

- Further conclusion : Laser power has a significant linear (but not quadratic) effect on strength.
- 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.

Orthogonal Polynomials



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

$$P_1(x) = \frac{x-m}{\Delta},$$
$$(= -1, 0 \text{ and } 1, \text{ for } x = m-\Delta, m, m+\Delta),$$

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

Therefore, $P_1(x)$ and $P_2(x)$ are extensions of the linear and quadratic contrast vectors. (Why ?)

- **Polynomial regression model :**

$$y = \beta_0^* + \beta_1^* \times P_1(x)/\sqrt{2} + \beta_2^* \times P_2(x)/\sqrt{6} + \varepsilon,$$

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:

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

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

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

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

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

❖ **Reading**: textbook, 2.3

Residual Analysis: Theory

- Theory: define the **residual** for the i^{th} observation (x_i, y_i) as

$$r_i = y_i - \hat{y}_i, \qquad \hat{y}_i = \mathbf{x}_i^T \hat{\beta},$$

  $\hat{y}_i$ contains information given by the model; r_i is the "difference" between y_i (observed) and $\hat{y}_i$ (fitted) and contains information on possible *model inadequacy*.

- Vector of residuals $\mathbf{r} = (r_1, \ldots, r_N)^T = \mathbf{y} - \mathbf{X}\hat{\beta}$.

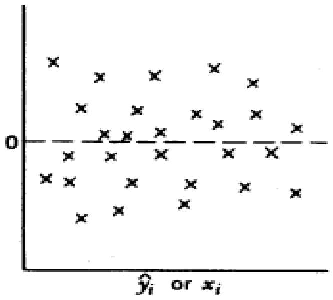
- Under the model assumption $E(\mathbf{y}) = \mathbf{X}\beta$, it can be shown that
  - (a) $E(\mathbf{r}) = 0$,
  - (b) $\mathbf{r}$ and $\hat{\mathbf{y}}$ are independent,
  - (c) variances of r_i are nearly constant for "nearly balanced" designs.

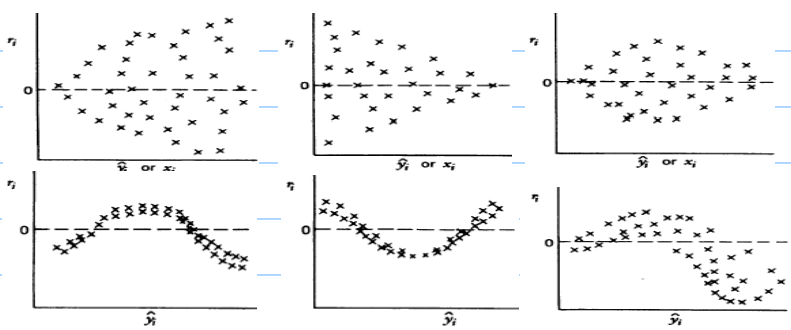
$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} = \hat{\boldsymbol{Y}} + \hat{\boldsymbol{\varepsilon}}$$

$$\boldsymbol{Y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon} = (\boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{H}_1\boldsymbol{X}_2\boldsymbol{\beta}_2) + ((\boldsymbol{I}-\boldsymbol{H}_1)\boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}) = \hat{\boldsymbol{Y}}_{X_1} + \hat{\boldsymbol{\varepsilon}}_{X_1})$$

Residual Plots

- Plot r_i vs. $\hat{y}_i$ (see Figure 1): It should appear as a parallel band around 0. Otherwise, it would suggest model violation. If spread of r_i increases as $\hat{y}_i$ increases, error variance of y increases with mean of y. May need a transformation of y. (Will be explained in future lecture.)
- Plot r_i from replicates per treatment (see Figure 2): to see if error variance depends on treatment.
- Plot r_i vs. x_i: If not a parallel band around 0, relationship between y_i and x_i not fully captured, revise the $\mathbf{X}\beta$ part of the model.
- Plot r_i vs. time sequence: to see if there is a time trend or autocorrelation over time.





Plot of r_i vs. $\hat{y}_i$

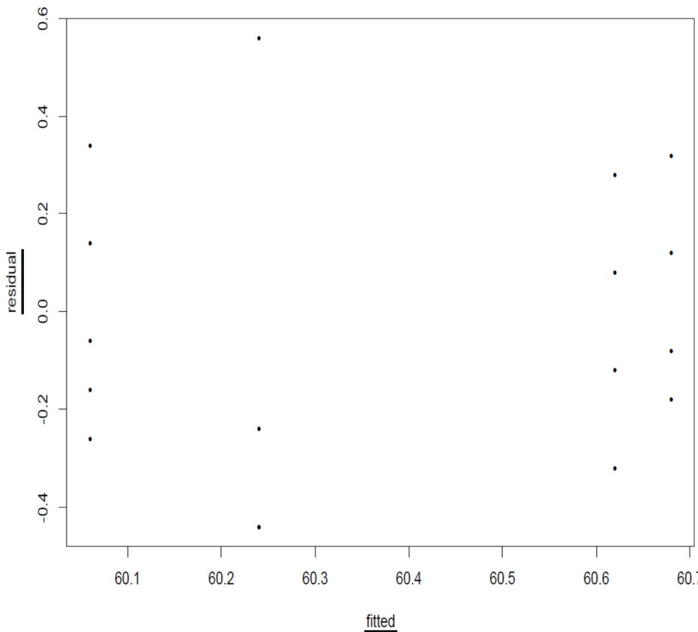


Figure 1: r_i vs. $\hat{y}_i$, Pulp Experiment

Plot of r_i vs. treatment

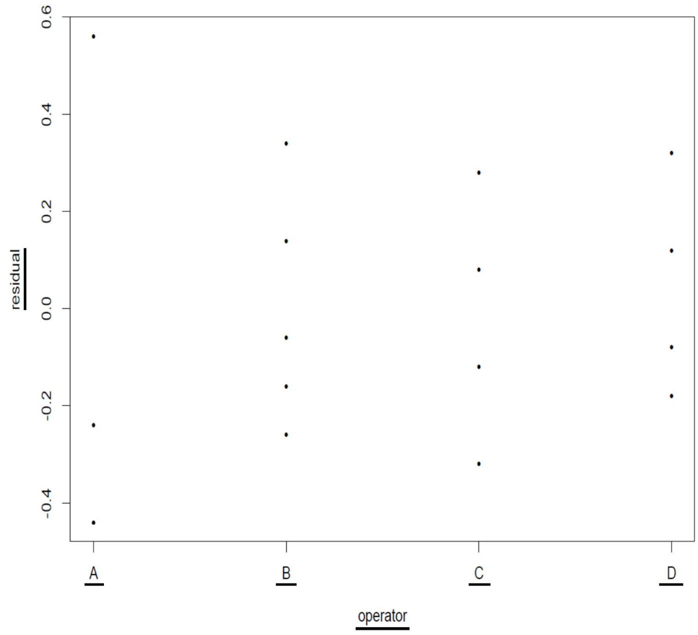


Figure 2: r_i vs. treatment, Pulp Experiment

Box-(Whisker) Plot

- A powerful graphical display (due to Tukey) to capture the location, dispersion, skewness and extremity of a distribution. See Figure 3.
- Q_1 = lower quartile (25% quantile), Q_3 = upper quartile (75% quantile), Q_2 = median (50% quantile, estimate of *location* parameter) is the white line in the box. Q_1 and Q_3 are boundaries of the *black box*.
- IQR = interquartile range (length of box) = $Q_3 - Q_1$: measure of *dispersion*.
- Minimum and maximum of **observed** values within

$$[Q_1 - 1.5 \times IQR, \;\; Q_3 + 1.5 \times IQR]$$

  are denoted by two *whiskers*. Any values outside the whiskers are regarded as extreme values and displayed (possible *outliers*).
- If Q_1 and Q_3 are not symmetric around the median, it indicates *skewness*.
- Side-by-side box plots (LNp. 3-2∼3) are useful to compare the difference between the distributions of several groups of data.

Box-(Whisker) Plot

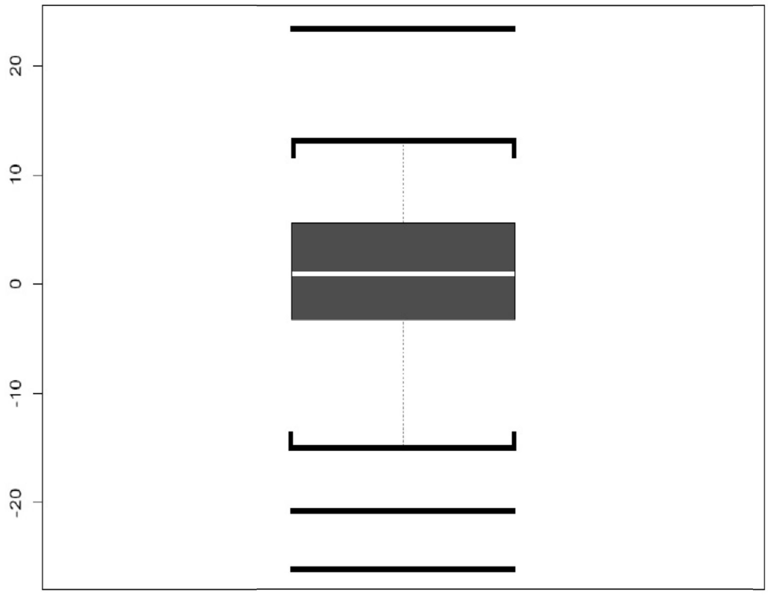


Figure 3: Box-Whisker Plot

Normal Probability Plot

- Original purpose : To test if a distribution is normal, e.g., if the residuals follow a normal distribution (see Figure 5).
- More powerful use in factorial experiments (discussed in Units 5 and 6).
- Let $r_{(1)} \leq \ldots \leq r_{(N)}$ be the ordered residuals. The cumulative probability for $r_{(i)}$ is $p_i = (i-0.5)/N$. Thus the plot of p_i vs. $r_{(i)}$ should be S-shaped as in Figure 4(a) if the errors are normal. By transforming the scale of the horizontal axis, the S-shaped curve is straightened to be a line (see Figure 4(b)).
- **Normal probability plot** of residuals :

$$\left(\Phi^{-1}\left(\frac{i-0.5}{N}\right), r_{(i)}\right), \qquad i=1,\ldots,N, \qquad \Phi = \text{normal cdf}.$$

If the errors are normal, it should plot roughly as a straight line. See Figure 5.

Regular and Normal Probability Plots of Normal CDF

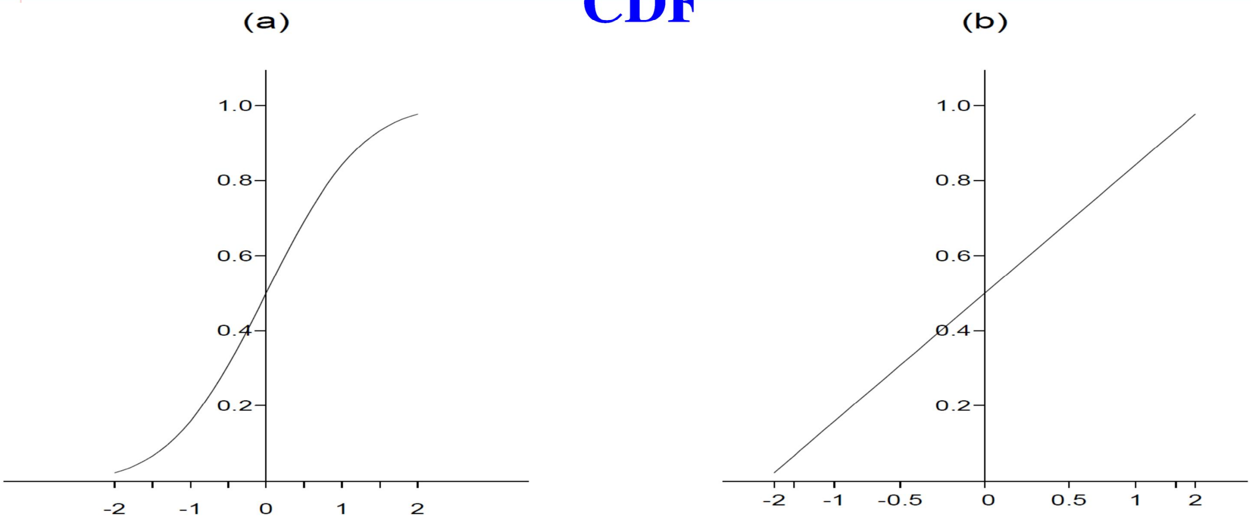


Figure 4: Normal Plot of r_i, Pulp Experiment

Normal Probability Plot : Pulp Experiment

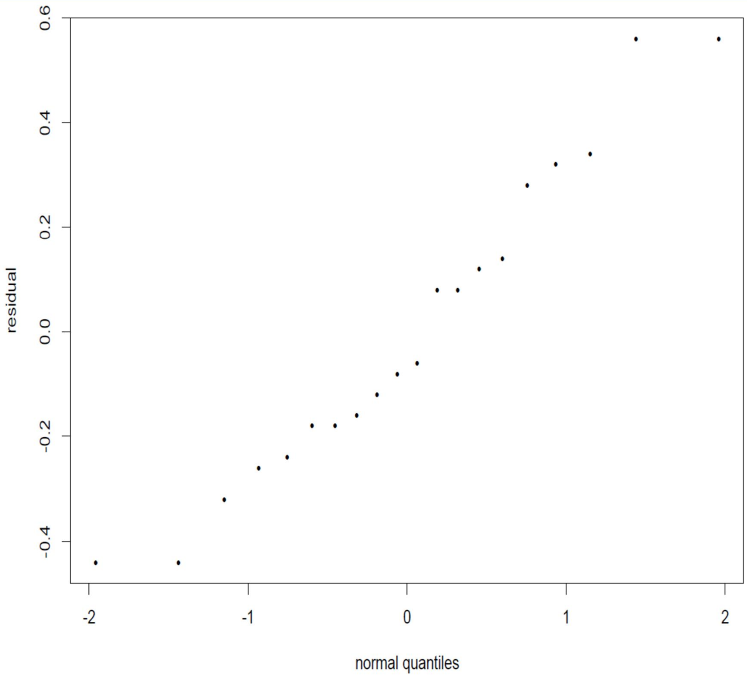


Figure 5: Normal Plot of r_i, Pulp Experiment

❖ **Reading**: textbook, 2.6

Pulp Experiment Revisited

- Compare the 2 scenarios

  (S1) plant has only 4 operators (or only interested in these 4 operators)
  - τ_i's: parameters (unknown fixed values)
  - interest: difference btwn the 4 specific τ_i's

  (S2) 4 operators randomly sampled from a large population of operators
  - τ_i's: random variables
  - interest: difference btwn all operators in this population
- In the pulp experiment the effects τ_i are called *fixed* effects because the interest was in comparing the four *specific* operators in the study. If these four operators were chosen randomly from the population of operators in the plant, the interest would usually be in the variation among all operators in the population. Because the observed data are from operators randomly selected from the population, the variation among operators in the *population* is referred to as *random* effects.

- **One-way random effects model (REM $\longleftrightarrow$ FEM) :**

$$y_{ij} = \eta + \tau_i + \epsilon_{ij},$$

where ϵ_{ij}'s: independent error terms with $N(0,\sigma^2)$,

τ_i's: independent $N(0,\sigma_\tau^2)$,

and τ_i and ϵ_{ij} are independent (Why? Give an example.);

σ^2 and σ_τ^2 are the two *variance components* of the model.

The variance among operators in the population is measured by σ_τ^2.

One-way Random Effects Model: ANOVA

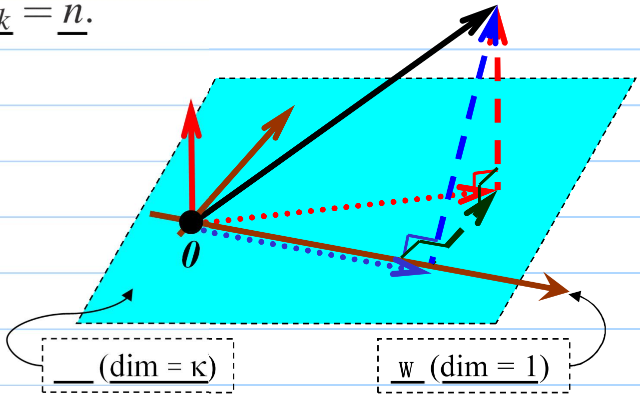
- In the following, assume $n_1 = \cdots = n_k = n$.
- The null hypothesis for the FEM:

  $$H_0 : \tau_1 = \cdots = \tau_k$$

  should be replaced by

  $$H_0^* : \sigma_\tau^2 = 0.$$

  Under H_0^*, the F-test and the ANOVA table in LNp. 3-6 still holds.



- Reason: under H_0^*,

  $$SSTr \sim \sigma^2 \chi^2_{k-1},$$

  and

  $$SSE \sim \sigma^2 \chi^2_{N-k},$$

  and they are independent.

  Therefore the F-test has the distribution $F_{k-1,N-k}$ under H_0^*.

ANOVA Tables ($n_i = n$)

- We can apply the *same* ANOVA and F-test in the fixed effects case for analyzing data.

Source	d.f.	SS	MS	E(MS)
treatment	$k-1$	$SSTr$	$MSTr = \frac{SSTr}{k-1}$	$\sigma^2 + n\sigma_\tau^2$
residual	$N-k$	SSE	$MSE = \frac{SSE}{N-k}$	σ^2
total	$N-1$			

Pulp Experiment

Source	d.f.	SS	MS	E(MS)
treatment	3	1.34	0.447	$\sigma^2 + 5\sigma_\tau^2$
residual	16	1.70	0.106	σ^2
total	19	3.04		

- However, we need to compute the expected mean squares under the alternative of $\sigma_\tau^2 > 0$,

  (i) for sample size determination, and
  (ii) to estimate the variance components.

Expected Mean Squares for Treatments

- Equation (1) holds independent of σ_τ^2,

$$E(MSE) = E\left(\frac{SSE}{N-k}\right) = \sigma^2. \qquad (1)$$

- Under the alternative: $\sigma_\tau^2 > 0$, and for $n_i = n$,

$$E(MSTr) = E\left(\frac{SSTr}{k-1}\right) = \sigma^2 + n\sigma_\tau^2. \qquad (2)$$

- For unequal n_i's, n in (2) is replaced by

$$n' = \frac{1}{k-1}\left[\sum_{i=1}^{k} n_i - \frac{\sum_{i=1}^{k} n_i^2}{\sum_{i=1}^{k} n_i}\right].$$

Proof of (2)

$$\bar{y}_{i.} - \bar{y}_{..} = \left(\tau_i - \bar{\tau}_.\right) + \left(\bar{\varepsilon}_{i.} - \bar{\varepsilon}_{..}\right)$$

$$SSTr = \sum_{i=1}^{k} n\left(\bar{y}_{i.} - \bar{y}_{..}\right)^2$$

$$= n\left\{\sum (\tau_i - \bar{\tau}_.)^2 + \sum (\bar{\varepsilon}_{i.} - \bar{\varepsilon}_{..})^2 + 2\sum (\bar{\varepsilon}_{i.} - \bar{\varepsilon}_{..})(\tau_i - \bar{\tau}_.)\right\}.$$

The cross product term has mean 0 (because τ and ε are independent). It can be shown that

$$E\left(\sum_{i=1}^{k} (\tau_i - \bar{\tau}_.)^2\right) = (k-1)\sigma_\tau^2 \quad \text{and} \quad E\left(\sum_{i=1}^{k} (\bar{\varepsilon}_{i.} - \bar{\varepsilon}_{..})^2\right) = \frac{(k-1)\sigma^2}{n}.$$

Therefore

$$E(SSTr) = n(k-1)\sigma_\tau^2 + (k-1)\sigma^2,$$

$$E(MSTr) = E\left(\frac{SSTr}{k-1}\right) = \sigma^2 + n\sigma_\tau^2.$$

Variance components: estimation of σ^2 and σ_τ^2

- From equations (1) and (2) in LNp. 3-35, we obtain the following unbiased estimates of the variance components:

$$\hat{\sigma}^2 = MSE \quad \text{and} \quad \hat{\sigma}_\tau^2 = \frac{MSTr - MSE}{n}.$$

Note that $\hat{\sigma}_\tau^2 \geq 0$ if and only if $MSTr \geq MSE$, which is equivalent to $F \geq 1$. Therefore a *negative* variance estimate $\hat{\sigma}_\tau^2$ occurs only if the value of the F statistic is less than 1. Obviously the null hypothesis H_0 is not rejected when $F \leq 1$. Since variance cannot be negative, a negative variance estimate is replaced by 0. This does not mean that σ_τ^2 is zero. It simply means that there is not enough information in the data to get a good estimate of σ_τ^2.

- For the pulp experiment, $n = 5$, $\hat{\sigma}^2 = 0.106$, $\hat{\sigma}_\tau^2 = (0.447 - 0.106)/5 = 0.068$, i.e., sheet-to-sheet variance (within same operator) is 0.106, which is about 50% higher than operator-to-operator variance 0.068.

Implications on process improvement: try to reduce the two sources of variation, also considering costs.

Estimation of Overall Mean η

- In REM, η, the population mean, is often of interest. From $E(y_{ij}) = \eta$, we use the estimate

$$\hat{\eta} = \bar{y}_{..}\ .$$

- $Var(\hat{\eta}) = Var(\bar{\tau}_. + \bar{\varepsilon}_{..}) = \frac{\sigma_\tau^2}{k} + \frac{\sigma^2}{N}$, where $N = \sum_{i=1}^k n_i$.

For $n_i = n$, $Var(\hat{\eta}) = \frac{\sigma_\tau^2}{k} + \frac{\sigma^2}{nk} = \frac{1}{nk}\left(\sigma^2 + n\sigma_\tau^2\right)$.
Using (2) in LNp.3-35, $\frac{MSTr}{nk}$ is an unbiased estimate of $Var(\hat{\eta})$.
Confidence interval for η:

$$\hat{\eta} \pm t_{k-1,\frac{\alpha}{2}} \sqrt{\frac{MSTr}{nk}}$$

- In the pulp experiment, $\hat{\eta} = 60.40$, $MSTr = 0.447$, and the 95% confidence interval for η is

$$60.40 \pm 3.182\sqrt{\frac{0.447}{5 \times 4}} = [59.92, 60.88].$$

❖ **Reading**: textbook, 2.5