Lecture 6
More on Complete Randomized Block Design (RBD)

Multiple test

Multiple test

- The multiple comparisons or multiple testing problem occurs when one considers a set of statistical inferences simultaneously.
- For *a* levels and their means $\mu_1, \mu_2, \cdots, \mu_a$, testing the following $\binom{a}{2}$ hypotheses:

$$H_0^{ij}: \mu_i = \mu_j, i < j,\ i, j = 1, 2, \cdots, a$$

Fisher Least Significant Different (LSD) Method

- This method builds on the equal variances t-test of the difference between two means.
- The test statistic is improved by using MS_ε rather than s_p^2.
- It is concluded that $\bar{y}_i$ and $\bar{y}_j$ differ at α significance level if $\left|\bar{y}_i - \bar{y}_j\right|$ > LSD, where

$$LSD = t_{\alpha/2, df_\varepsilon}\sqrt{MS_\varepsilon\left(\frac{1}{r_i} + \frac{1}{r_j}\right)}$$

LSD

- Where r_i and r_j are number of observations under level *i* and *j*.
- And $$df_\varepsilon = \sum_{i=1}^{a} r_i - a$$

Experiment-wise Type I error rate (α_E) (the effective Type I error)

- The Fisher's method may result in an increased probability of committing a type I error.
- The experiment-wise Type I error rate is the probability of committing at least one Type I error at significance level of α. It is calculated by

$$\alpha_E = 1-(1 - \alpha)^C$$

  where C is the number of pairwise comparisons (all: $C = a(a-1)/2$)
- The Bonferroni adjustment determines the required Type I error probability per pairwise comparison (α) , to secure a pre-determined overall α_E.

Bonferroni adjustment

- The procedure:
  - Compute the number of pairwise comparisons (C)
    [all: C=*a*(*a*-1)/2], where *a* is the number of populations.
  - Set $\alpha = \alpha_E/C$, where α_E is the true probability of making at least one Type I error (called *experimentwise Type I error).*
  - It is concluded that $\bar{y}_i$ and $\bar{y}_j$ differ at α/C significance level if

$$\left|\bar{y}_i - \bar{y}_j\right| > t_{\alpha/(2C),df_\varepsilon}\sqrt{MS_\varepsilon\left(\frac{1}{r_i}+\frac{1}{r_j}\right)}$$

Duncan's multiple range test

- The Duncan Multiple Range test uses different Significant Difference values for means next to each other along the real number line, and those with 1, 2, … , *a* means in between the two means being compared.
- The Significant Difference or the range value:

$$R_p = r_{\alpha,p,\nu}\sqrt{MS_\varepsilon / n}$$

- where $r_{\alpha,p,v}$ is the Duncan's Significant Range Value with parameters *p* (= range-value) and *v* (= MS_ε degree-of-freedom), and experiment-wise alpha level α (= α_{joint}).

Duncan's Multiple Range Test

- MS_ε is the mean square error from the ANOVA table and n is the number of observations used to calculate the means being compared.
- The range-value is:
  - 2 if the two means being compared are adjacent
  - 3 if one mean separates the two means being compared
  - 4 if two means separate the two means being compared
  - …

Significant ranges for Duncan's Multiple Range Test

Critical Points for Duncan's Multiple Range Statistic -- ALPHA = 0.05												
Degrees of freedom ν	p											
	2	3	4	5	6	7	8	9	10	20	50	100
1	18.00	18.00	18.00	18.00	18.00	18.00	18.00	18.00	18.00	18.00	18.00	18.00
2	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
3	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50
4	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
5	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83
6	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68
7	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61	3.61
8	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56	3.56
9	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52	3.52
10	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.48	3.48	3.48
11	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.48	3.48	3.48
12	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.46	3.48	3.48	3.48
13	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.47	3.47	3.47
14	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.47	3.47	3.47
15	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.47	3.47	3.47
16	3.00	3.15	3.23	3.30	3.34	3.37	3.39	3.41	3.43	3.47	3.47	3.47
17	2.98	3.13	3.22	3.28	3.33	3.36	3.38	3.40	3.42	3.47	3.47	3.47
18	2.97	3.12	3.21	3.27	3.32	3.35	3.37	3.39	3.41	3.47	3.47	3.47
19	2.98	3.11	3.19	3.26	3.31	3.35	3.37	3.39	3.41	3.47	3.47	3.47
20	2.95	3.10	3.18	3.25	3.30	3.34	3.36	3.38	3.40	3.47	3.47	3.47
30	2.89	3.04	3.12	3.20	3.25	3.29	3.32	3.35	3.37	3.47	3.47	3.47
40	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.47	3.47	3.47
60	2.83	2.98	3.08	3.14	3.20	3.24	3.28	3.31	3.33	3.47	3.48	3.48
100	2.80	2.95	3.05	3.12	3.18	3.22	3.26	3.29	3.32	3.47	3.53	3.53
inf	2.77	2.92	3.02	3.09	3.15	3.19	3.23	3.26	3.29	3.47	3.61	3.67

Tukey–Kramer method

- A procedure which controls the experimentwise error rate is "Tukey's honestly significant difference test ". It is used for levels with the same replications. Assume $r_1 = r_2 = \ldots = r_a = r$.
- Basic idea: if H_0^{ij} is true, the value $\left|\bar{y}_i - \bar{y}_j\right|$ should not be large. W: the rejection region of multiple tests (i.e. at least one H_0^{ij} is rejected)
- W= $\bigcup_{i<j} \left\{ \left|\bar{y}_i - \bar{y}_j\right| > c \right\}$

Tukey–Kramer method

- We need to determine *c* s.t. when all the H_0^{ij} are true, the probability of type I error is α, i.e. P(W)=α

$$P(\mathrm{W}) = \mathrm{P}\left(\bigcup_{i<j}\left\{\left|\bar{y}_i - \bar{y}_j\right| > c\right\}\right) = 1 - \mathrm{P}\left(\bigcap_{i<j}\left\{\left|\bar{y}_i - \bar{y}_j\right| \le c\right\}\right)$$

$$= 1 - P\left(\max_{i<j}\left|\bar{y}_i - \bar{y}_j\right| \le c\right) = P\left(\max_{i<j}\left|\bar{y}_i - \bar{y}_j\right| > c\right)$$

$$= P\left(\max_{i<j}\left|\frac{\bar{y}_i - \bar{y}_j}{\sqrt{MS_\varepsilon / r}}\right| > \frac{c}{\sqrt{MS_\varepsilon / r}}\right)$$

$$\overset{\text{all hypothesises are true}}{=} P\left(\max_{i<j}\left|\frac{(\bar{y}_i - \mu_\mathrm{i}) - (\bar{y}_j - \mu_\mathrm{j})}{\sqrt{MS_\varepsilon / r}}\right| > \frac{c}{\sqrt{MS_\varepsilon / r}}\right)$$

$$= P\left(\max_{i<j}\left(\frac{\bar{y}_i - \mu_\mathrm{i}}{\sqrt{MS_\varepsilon / r}}\right) - \min_{i<j}\left(\frac{\bar{y}_i - \mu_\mathrm{i}}{\sqrt{MS_\varepsilon / r}}\right) > \frac{c}{\sqrt{MS_\varepsilon / r}}\right)$$

Tukey–Kramer method

- MS_ε is mean square of errors in ANOVA, and is unbiased estimator of σ^2. It is independent with $\bar{y}_i$, so

$$\frac{\bar{y}_i - \mu_i}{\sqrt{MS_\varepsilon / r}} \sim t(f_\varepsilon) \qquad t_{(r)} = \max_i \left(\frac{\bar{y}_i - \mu_i}{\sqrt{MS_\varepsilon / r}} \right) \qquad t_{(1)} = \min_i \left(\frac{\bar{y}_i - \mu_i}{\sqrt{MS_\varepsilon / r}} \right)$$

- They are the largest and smallest order statistics from a sample (*a* observations, obey $t(f_\varepsilon)$). Denote $q(a, f_\varepsilon) = t_{(a)} - t_{(1)}$

Tukey–Kramer method

- Then $$P(\mathrm{W}) = \mathrm{P}\left(\mathrm{q}(a, \mathrm{f}_{\varepsilon}) > \frac{c}{\sqrt{MS_{\varepsilon}/r}} \right) = \alpha$$

$$i.e. \frac{c}{\sqrt{MS_{\varepsilon}/r}} = \mathrm{q}_{1\text{-}\alpha}(a, \mathrm{f}_{\varepsilon}), \mathrm{c} = \mathrm{q}_{1\text{-}\alpha}(a, \mathrm{f}_{\varepsilon})\sqrt{MS_{\varepsilon}/r}$$

- So the rejection region of these multiple tests with significant level α is

$$\left| \overline{y}_i - \overline{y}_j \right| > \mathrm{q}_{1\text{-}\alpha}(a, \mathrm{f}_{\varepsilon})\sqrt{MS_{\varepsilon}/r}, i < j, i, j = 1, 2, \cdots, a$$

Scheffe method

- For different number of replications
- Under $H_0^{ij} : \mu_{\mathrm{i}} = \mu_j$

  $\overline{y}_i - \overline{y}_j$ ~ N(0, $\left(\frac{1}{r_i} + \frac{1}{r_j}\right)\sigma^2$)
- Replace σ^2 by MSε, and MSε is independent with $\overline{y}_i$, so

  $$F_{ij} = \frac{\left(\overline{y}_i - \overline{y}_j\right)^2}{\left(\frac{1}{r_i} + \frac{1}{r_j}\right)MS_\varepsilon}$$ ~ F(1, fε)
- If H_0^{ij} is true, F_{ij} should not be large.

Scheffe method

- When all the H_0^{ij} are true, the rejection region of multiple tests is

$$W = \bigcup_{i<j} \left\{ F_{\mathrm{ij}} > c \right\} \quad P(W) = P\left(\bigcup_{i<j} \left\{ F_{\mathrm{ij}} > c \right\} \right) = P\left(\max_{\mathrm{i<j}} F_{ij} > c \right)$$

- Scheffe proved that $\frac{\max_{\mathrm{i<j}} F_{ij}}{a-1}$ approximately obeys F(a-1, f_ε). Given significant level α, then $\mathrm{c} = (a-1)\mathrm{F}_{1-\alpha}(a-1, \mathrm{f}_\varepsilon)$
- The rejection region is

$$\left| \overline{y}_i - \overline{y}_j \right| > \sqrt{(a-1)\mathrm{F}_{1-\alpha}(a-1, \mathrm{f}_\varepsilon)(\frac{1}{r_\mathrm{i}} + \frac{1}{r_\mathrm{j}})\mathrm{MS}_\varepsilon}, i < j, i,j = 1,2,\cdots,a$$

Test of normality

Test of normality

- Many test procedures that we have developed rely on the assumption of Normality.
- There are several methods of assessing whether data are normally distributed or not (H_0: the data obeys Normal distribution; H_1: not obey Normal distribution).

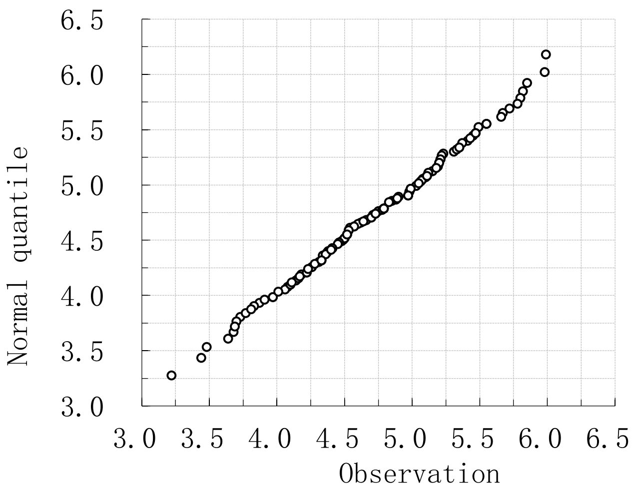
Test of normality

- They fall into two broad categories: graphical and statistical. The most common are:
- 1.Graphical
  - Q-Q probability plots
  - Cumulative frequency (P-P) plots
- 2. Statistical
  - Kolmogorov-Smirnov test
  - Shapiro-Wilk test

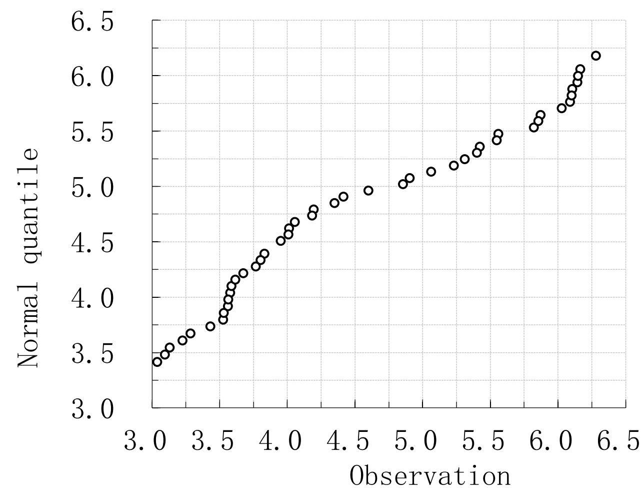
Q-Q probability plots

- Q-Q plots display the observed values against normally distributed data (represented by the line).

Normal Q-Q plot: Normally distribution data



Normal Q-Q plot: Non-normally distribution data

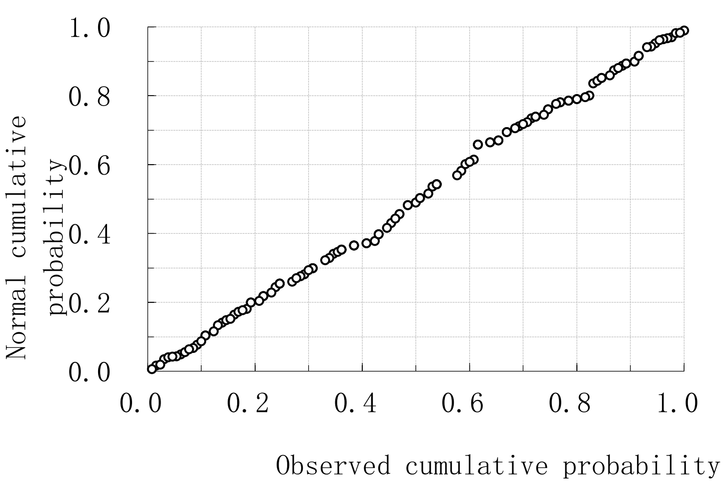


- Normally distributed data fall along the line.

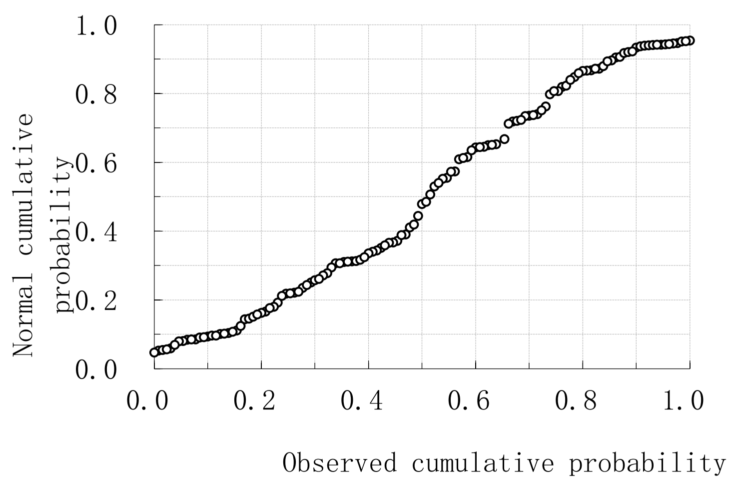
P-P plots

- P-P (cumulative) plot displays the cumulative probabilities against normally distributed data (represented by the line).

Normal P-P plot: Normally distribution data



Normal P-P plot: Non-normally distribution data



- Normally distributed data fall along the line.

Statistical tests

- Statistical tests for normality are more precise since actual
- probabilities are calculated.
- The Kolmogorov-Smirnov and Shapiro-Wilks tests for normality calculate the probability that the sample was drawn from a normal population.
- The hypotheses used are:
- H_0: The sample data are not significantly different than a normal population.
- H_a: The sample data are significantly different than a normal population.

Statistical tests

- Typically, we are interested in finding a difference between groups. When we are, we ‘look’ for small probabilities.
- If the probability of finding an event is rare (less than 5%) and we actually find it, that is of interest. When testing normality, we are not looking for a difference.
- In effect, we want our data set to be no different than normal. So when testing for normality:
- Probabilities > 0.05 mean the data are normal.
- Probabilities < 0.05 mean the data are NOT normal.

Kolmogorov-Smirnov Normality Test

- Based on comparing the observed frequencies and the expected frequencies
- Let $F(\mathrm{x}) = \mathrm{P}(\mathrm{X_i} \leq \mathrm{x})$ be the cdf for the distribution.
- In the uniform(0,1) case: $F(\mathrm{x}) = x, 0 \leq x \leq 1$
- Compare this to the "empirical distribution function":

$$\hat{F}_n(\mathrm{x}) = \frac{1}{n}(\text{number of } X_i \text{ in the sample} \leq x)$$

Kolmogorov-Smirnov Normality Test

- If X_1, X_2, …, X_n really come from the distribution with cdf F, the distance

$$D = D_n = \max_x \left| \hat{F}_n(x) - F(x) \right|$$

should be small.

- Example: Suppose we have 7 observations:
- 0.6 0.2 0.5 0.9 0.1 0.4 0.2
- Put them in order:
- 0.1 0.2 0.2 0.4 0.5 0.6 0.9

Kolmogorov-Smirnov Normality Test

- Now the empirical cdf is:

$\hat{F}_7(x) = 0 \quad$ for $x < 0.1$

$\hat{F}_7(x) = \frac{1}{7} \quad$ for $0.1 \le x < 0.2$

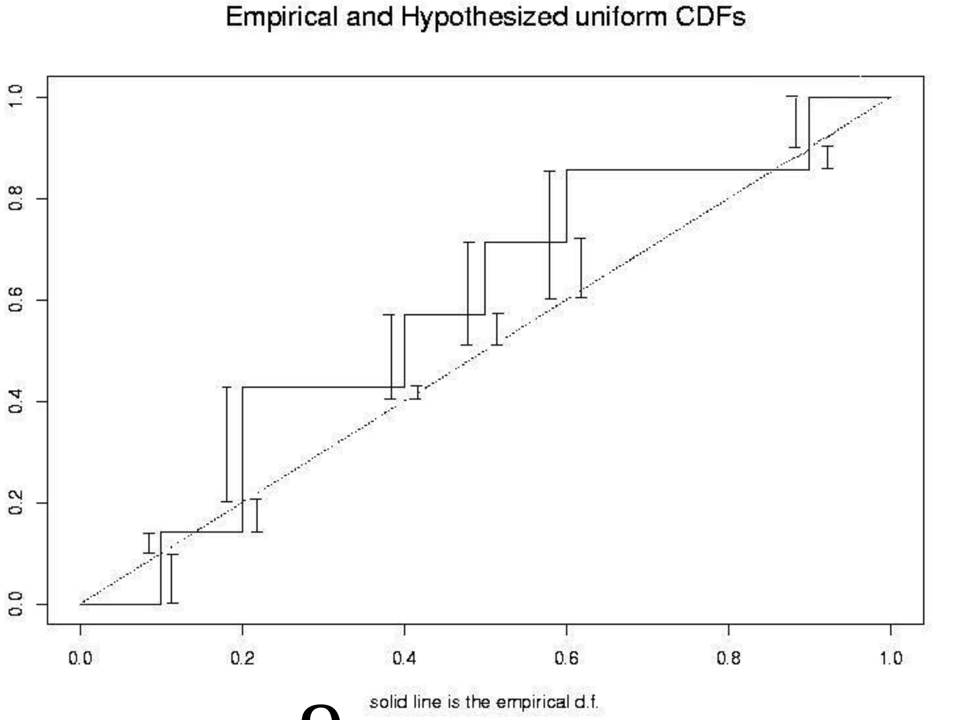
$\hat{F}_7(x) = \frac{3}{7} \quad$ for $0.2 \le x < 0.4$

$\hat{F}_7(x) = \frac{4}{7} \quad$ for $0.4 \le x < 0.5$

$\hat{F}_7(x) = \frac{5}{7} \quad$ for $0.5 \le x < 0.6$

$\hat{F}_7(x) = \frac{6}{7} \quad$ for $0.6 \le x < 0.9$

$\hat{F}_7(x) = 1 \quad$ for $x \ge 0.9$



$$D_7 = \frac{9}{35} \approx 0.2571429$$

Kolmogorov-Smirnov Normality Test

- Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ be the ordered sample.
- Then D_n can be estimated by

$$D_n = \max\left\{ D_n^+, D_n^- \right\}$$

- Where

$$D_n^+ = \max_{1 \le i \le n} \left\{ \frac{i}{n} - F(X_{(i)}) \right\}$$

$$D_n^- = \max_{1 \le i \le n} \left\{ F(X_{(i)}) - \frac{i-1}{n} \right\}$$

- (assuming non-repeating values)

Kolmogorov-Smirnov Normality Test

- We reject that this sample came from the proposed distribution if the empirical cdf is too far from the true cdf of the proposed distribution
- i.e.: We reject if D_n is too “large”.

Kolmogorov-Smirnov Normality Test

- In the 1930's, Kolmogorov and Smirnov showed that $\lim_{n \to \infty} P\left(n^{1/2} D_n \le t\right) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 t^2}$
- So, for large sample sizes, you could assume
$$P(n^{1/2} D_n \le t) \approx 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 t^2}$$
- and find the value of t that makes the right hand side 1-α for an α level test.

Kolmogorov-Smirnov Normality Test

- For small samples, people have worked out and tabulated critical values, but there is no nice closed form solution.
  - J. Pomeranz (1973)
  - J . Durbin (1968)
- Good approximations for n>40:

α	0.20	0.10	0.05	0.02	0.01
CV	$\frac{1.0730}{\sqrt{n}}$	$\frac{1.2239}{\sqrt{n}}$	$\frac{1.3581}{\sqrt{n}}$	$\frac{1.5174}{\sqrt{n}}$	$\frac{1.6276}{\sqrt{n}}$

Kolmogorov-Smirnov Normality Test

- From a table, the critical value for a 0.05 level test for n=7 is 0.483.

$$D_7 = \frac{9}{35} \approx 0.2571429 < 0.483$$

- So we cannot reject H_0, i.e. the data obeys Normal distribution.

Shapiro-Wilk test

- The test statistic is: $$W = \frac{\left(\sum_{i=1}^{[n/2]} a_i \left(x_{(n+1-i)} - x_{(i)}\right)\right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$
- Where $x_{(i)}$ is the ith order statistic
- The constants a_i are given by

$$(a_1, \cdots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m^T)^{1/2}}, \textit{where } m = (m_1, \cdots, m_n)^T$$

- m_1, ..., m_n are the expected values of the order statistics of independent and identically distributed random variables sampled from the standard normal distribution, and V is the covariance matrix of those order statistics.

Shapiro-Wilk test

- The user may reject the null hypothesis if is too small
- W should be closed to 1 under H_0. The rejection region is $\{W \leq c\}$
- Where $P(W \leq c) = \alpha$

Example

- The data from FDP activities of mice.

Obs	1	2	3	4	5	6	7	8
x	3.83	3.16	4.70	3.97	2.03	2.87	3.65	5.09

- Put the data in order and divide them into two.

Obs in order	1	2	3	4	5	6	7
x	2.03	2.87	3.16	3.65	3.83	3.97	4.70

Example

Order	a_i	$x_{(n+1-i)}$	$x_{(i)}$	$d_i=x_{(n+1-i)}-x_{(i)}$	$a_i d_i$
1	0.6052	5.09	2.03	3.06	1.851912
2	0.3164	4.70	2.87	1.83	0.579012
3	0.1743	3.97	3.16	0.81	0.141183
4	0.0561	3.83	3.65	0.18	0.010098
Total					2.582205

- So

$$W = \frac{\left(\sum_{i=1}^{[n/2]} a_i \left(x_{(n+1-i)} - x_{(i)}\right)\right)^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2} = \frac{(2.582205)^2}{6.782549963} = 0.98307903$$

- From the reference table of W, $W_{(8,0.05)}=0.818<W$, so we cannot reject H_0.

Choosing the methods

- Which normality test should I use?
- Kolmogorov-Smirnov:
  - More suitable for large samples.
- Shapiro-Wilk:
  - Works best for small data sets

Test of homogeneity

Test of homogeneity

- If we have various groups or levels of a variable, we want to make sure that the variance within these groups or levels is the same. It is the basic assumption for ANOVA.
- H_0: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_a^2$
- H_1: not H_0

Test of homogeneity

- Some common methods:
- 1. Hartley test: only used for the same sample size for each level.
- 2. Bartlett test: sample size may be different, but each size>=5
- 3. Adjusted Bartlett test: sample size may be different, and no restriction on size

Hartley test

- The numbers of replications are the same for each level, i.e. $r_1 = r_2 = \cdots = r_a = r$
- Hartley proposed the statistic:

$$H = \frac{\max\{S_1^2, S_2^2, \cdots, S_a^2\}}{\min\{S_1^2, S_2^2, \cdots, S_a^2\}}$$

- The values of H under H_0 can be simulated, and denote the distribution as H(a, f), f=r-1

Hartley test

- Under H_0, the value of H should be close to 1.
- Given significance level α, the rejection region should be

$$W_1 = \{\mathrm{H} > \mathrm{H}_{1\text{-}\alpha}(a, \mathrm{f})\}$$

- Where $\mathrm{H}_{1\text{-}\alpha}(a, f)$ is the 1- α quantile of H distribution.

Bartlett test

- The ith sample variance is

$$S_i^2 = \frac{1}{r_i - 1}\sum_{j=1}^{r_i}(\mathrm{y}_{\mathrm{ij}} - \bar{\mathrm{y}}_{\mathrm{i}})^2 = \frac{Q_i}{f_i}, i = 1, 2, \cdots, a$$

- Where $Q_i = \sum_{j=1}^{r_i}(\mathrm{y}_{\mathrm{ij}} - \bar{\mathrm{y}}_{\mathrm{i}})^2, \mathrm{f}_{\mathrm{i}} = r_i - 1 \text{(degree of freedom)}$

- We know that $MS_\varepsilon = \frac{1}{f_\varepsilon}\sum_{i=1}^{a} Q_i = \sum_{i=1}^{a}\frac{f_i}{f_\varepsilon}S_i^2$

- It is the (average) arithmetic mean of

$$S_1^2, S_2^2, \cdots, S_a^2$$

Bartlett test

- Denote geometric mean

$$GMS_\varepsilon = \left[(S_1^2)^{f_1} (S_2^2)^{f_2} \cdots (S_a^2)^{f_a} \right]^{\frac{1}{f_\varepsilon}}$$

- Where $f_\varepsilon = \sum_{i=1}^{a} \mathrm{f_i} = \sum_{i=1}^{a} (r_\mathrm{i} - 1) = n - a$
- It is true that $GMS_\varepsilon \le MS_\varepsilon$

$$GMS_\varepsilon = MS_\varepsilon \Leftrightarrow S_1^2 = S_2^2 = \cdots = S_a^2$$

- So under H_0, $\frac{MS_\varepsilon}{GMS_\varepsilon}$ should be close to 1. If it is too large, reject H_0. Rejection region is

$$W_1 = \{\ln \frac{MS_\varepsilon}{GMS_\varepsilon} > d\}$$

Bartlett test

- Bartlett proved that: for large sample, one function of $\ln \frac{MS_\varepsilon}{GMS_\varepsilon}$ approximately obeys $\chi^2(a-1)$
- i.e. $$B = \frac{f_\varepsilon}{C}(\ln \mathrm{MS}_\varepsilon - \ln \mathrm{GMS}_\varepsilon) \overset{\bullet}{\sim} \chi^2(a-1)$$
- Where $$C = 1 + \frac{1}{3(a-1)}\left(\sum_{i=1}^{a}\frac{1}{f_i} - \frac{1}{f_\varepsilon}\right)$$

is always larger than 1.

Bartlett test

- Taking the statistic

$$B = \frac{1}{C}(f_\varepsilon \ln MS_\varepsilon - \sum_{i=1}^{a} f_i \ln S_i^2) \overset{\bullet}{\sim} \chi^2(a-1)$$

- The rejection region is

$$W_1 = \{B > \chi^2_{1-\alpha}(a-1)\}$$

- Here B approximately obeys χ^2 . So the method is more suitable for data with more than 5 replications in each level.

Adjusted Bartlett test

- Box proposed the adjusted Bartlett statistic

$$B' = \frac{f_2 BC}{f_1(\mathrm{A} - \mathrm{BC})}$$

- B and C are given above. And

$$f_1 = a - 1,\ f_2 = \frac{a+1}{(\mathrm{C}-1)^2},\ A = \frac{f_1}{2 - C + 2/f_2}$$

- Under H_0, B' approximately obeys $F(f_1, f_2)$

Adjusted Bartlett test

- The rejection region is

$$W_1 = \{B' > \mathrm{F}_{1\text{-}\alpha}(\mathrm{f}_1, \mathrm{f}_2)\}$$

- Sometimes, f_2 is not an integer. We can use Interpolation method of the quantiles for F distribution.

Example

- Testing folic acid content for teas from 4 locations.

Level	Data							Rep	Sum	Mean	SS within groups
A_1	7.9	6.2	6.6	8.6	8.9	10.1	9.6	$r_1 = 7$	$T_1 = 57.9$	8.27	$Q_1 = 12.83$
A_2	5.7	7.5	9.8	6.1	8.4			$r_2 = 5$	$T_2 = 37.5$	7.50	$Q_2 = 11.30$
A_3	6.4	7.1	7.9	4.5	5.0	4.0		$r_3 = 6$	$T_3 = 34.9$	5.82	$Q_3 = 12.03$
A_4	6.8	7.5	5.0	5.3	6.1	7.4		$r_4 = 6$	$T_4 = 38.1$	6.35	$Q_4 = 5.61$
								$r = 24$	$T = 168.4$		$S_\varepsilon = 41.77$

Example

- For Bartlett test,

$$S_1^2 = 2.14, S_2^2 = 2.83, S_3^2 = 2.41, S_4^2 = 1.12$$

- And MS_ε=2.09. Then

$$C = 1 + \frac{1}{3(a-1)}\left(\sum_{i=1}^{a}\frac{1}{f_i} - \frac{1}{f_\varepsilon}\right)$$

$$= 1 + \frac{1}{3\times(4-1)}\left(\frac{1}{6} + \frac{1}{4} + \frac{1}{5} + \frac{1}{5} - \frac{1}{20}\right) = 1.0856$$

$$B = \frac{1}{C}\left(f_\varepsilon \ln MS_\varepsilon - \sum_{i=1}^{a} f_i \ln S_i^2\right)$$

$$= \frac{1}{1.0856}\cdot\left(20\times\ln 2.09 - \left(6\times\ln 2.14 + 4\times\ln 2.83 + 5\times\ln 2.41 + 5\times\ln 1.12\right)\right) = 0.970$$

Example

- Given α=0.05,

$$\chi^2_{1-\alpha}(a-1) = \chi^2_{0.95}(4-1) = 7.815 > B$$

- So we cannot reject H_0, i.e. we agree with that

$$\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_{\mathrm{a}}^2$$

Example

- For Adjusted Bartlett test,

$$f_1 = 4 - 1 = 3 \qquad f_2 = \frac{4+1}{(1.0856C - 1)^2} = 682.4$$

$$A = \frac{682.4}{2 - 1.0856 + 2/682.4} = 743.9 \quad B' = \frac{682.4 \times 0.970 \times 1.0856}{3 \times (743.9 - 0.970 \times 1.0856)} = 0.322$$

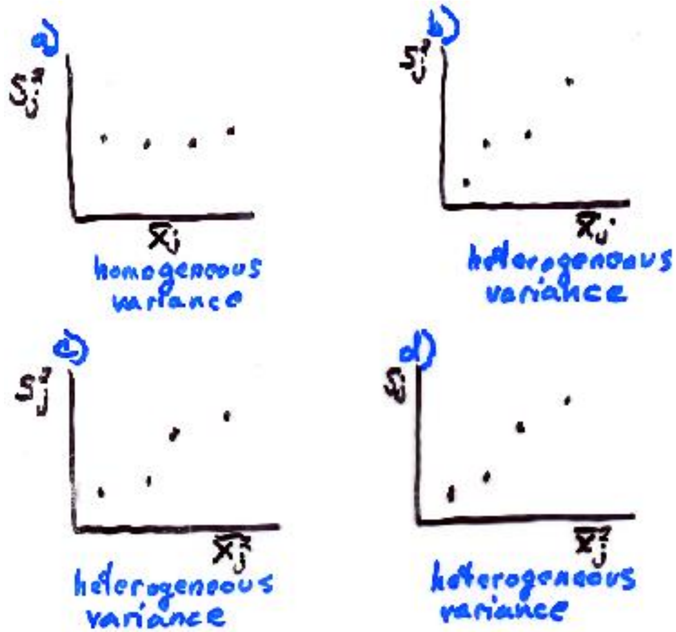
- Given α=0.05,

$$F_{0.95}(3,\quad 682.4) = F_{0.95}(3,\quad +\infty) = 2.60 > B'$$

- We cannot reject H_0

Data transformation

- Data transformation is used to make your data obey Normal distribution.
- Three normal transformation methods：



a) No need for transformation
b) Use square root transformation
c) Use logarithmic transformation
d) Use reciprocal transformation

Let’s work on the previous example together

Mutants	Rep I	Rep II	Rep III
A	10.9	9.1	12.2
B	10.8	12.3	14.0
C	11.1	12.5	10.5
D	9.1	10.7	10.1
E	11.8	13.9	16.8
F	10.1	10.6	11.8
G	10.0	11.5	14.1
H	9.3	10.4	14.4

- Do the randomization of the three blocks
- Build the ANOVA table for the observed data
- Multiple test by LSD
- Normality test by Q-Q and P-P plot

- What else do we want?
- Orthogonal contrasts