Chapter four The experimental design

4)One way & repeated measures.

4.1)model:

$$X_{ij}=\mu+\alpha_i+\beta_j+\varepsilon_{ij}, i=1,2,...,A, j=1,2,...,B,$$

$$\varepsilon_{ij}\overset{iid}{\sim}N(0,\sigma^2), \sum_{i=1}^{A}\alpha_i=0, E(\beta_j)=0, Var(\beta_j)=\sigma_\beta^2,$$

Note : The error distribution is normal distribution only.

4.2)The analysis:

$$X_{ij}=\hat{\mu}+\hat{\alpha}_i+\hat{\beta}_j+e_{ij},$$

$$\Rightarrow \sum_i\sum_j X_{ij}=\sum_i\sum_j\hat{\mu}+\sum_i\sum_j\hat{\alpha}_i+\sum_i\sum_j\hat{\beta}_j+\sum_i\sum_j e_{ij},$$

$$when \quad \sum_j e_{ij}=0, \sum_j\hat{\beta}_j=0, \hat{\mu}+\hat{\alpha}_i=\frac{\sum_j X_{ij}}{B}=\overline{X}_{i\bullet},$$

$$when \quad \sum_i e_{ij}=0, \sum_i\hat{\alpha}_i=0, \hat{\mu}+\hat{\beta}_j=\frac{\sum_i X_{ij}}{A}=\overline{X}_{\bullet j},$$

$$\Rightarrow \sum_i\sum_j X_{ij}=\sum_i\sum_j\hat{\mu}+B\times\sum_i\hat{\alpha}_i+A\times\sum_j\hat{\beta}_j,$$

$$when \quad \sum_i\hat{\alpha}_i=0, \sum_j\hat{\beta}_j=0,$$

$$\Rightarrow \sum_i\sum_j X_{ij}=n_T\times\hat{\mu}, \hat{\mu}=\frac{\sum_i\sum_j X_{ij}}{n_T}=\overline{X}_{\bullet\bullet}, n_T=A\times B,$$

$$\hat{\alpha}_i=\overline{X}_{i\bullet}-\overline{X}_{\bullet\bullet},$$

$$\hat{\beta}_j=\overline{X}_{\bullet j}-\overline{X}_{\bullet\bullet},$$

$$e_{ij}=X_{ij}-\overline{X}_{i\bullet}-\overline{X}_{\bullet j}+\overline{X}_{\bullet\bullet},$$

$$X_{ij}=\hat{\mu}+\hat{\alpha}_i+\hat{\beta}_j+e_{ij}\Rightarrow X_{ij}-\overline{X}_{\bullet\bullet}=(\overline{X}_{i\bullet}-\overline{X}_{\bullet\bullet})+(\overline{X}_{\bullet j}-\overline{X}_{\bullet\bullet})+e_{ij},$$

$$\sum_i\sum_j(X_{ij}-\overline{X}_{\bullet\bullet})X_{ij}=\sum_i\sum_j\hat{\alpha}_i\times X_{ij}+\sum_i\sum_j\hat{\beta}_j\times X_{ij}+\sum_i\sum_j e_{ij}\times X_{ij},$$

$$SST=\sum_i\sum_j(X_{ij})^2-n_T\times(\overline{X}_{\bullet\bullet})^2, df=n_T-1,$$

$$SSE=\sum_i\sum_j e_{ij}\times(\hat{\mu}+\hat{\alpha}_i+\hat{\beta}_j+e_{ij})=\sum_i\sum_j(e_{ij})^2, df=(A-1)\times(B-1),$$

$$SSA=\sum_i\sum_j\hat{\alpha}_i\times X_{ij}=\sum_i\sum_j\hat{\alpha}_i\times(\hat{\mu}+\hat{\alpha}_i+\hat{\beta}_j+e_{ij})=\sum_i\sum_j(\hat{\alpha}_i)^2$$

$$=\sum_i\sum_j(\overline{X}_{i\bullet}-\overline{X}_{\bullet\bullet})^2, df=A-1,$$

ANOVA

Soruce	df	SS	MS
Subject	$B-1$	SSB	
Between	$A-1$	SSA	MSA=SSA/ (A-1)
Error	$(A-1)(B-1)$	SSE	MSE=SSE/ $((A-1)(B-1))$
Total	$A\times B-1$	SST	

$H_0 : \alpha_1 = ... = \alpha_A = 0, H_1$:against H_0

$$test \quad statistic = \frac{SSA/(A-1)}{SSE/((A-1)\times(B-1))},$$

4.3)Three basic assumption test:

4.3.1).The error probability distribution: Pearson chi square test.

When the sample size less than 8, this test will be not done.

4.3.2).The homogenous of variances:

$H_0 : \sigma_{11}^2 = ... = \sigma_{AB}^2, H_1 :$ against H_0

The test statistic is $\frac{Max(e_{11}^2,....,e_{AB}^2)}{SSE}$ (The critical value is gotten by simulation)

4.3.3).Whether error is randomly or not:

H_0 :Error is random H_1 :against H_0

The run test.

4.4)The multiple comparison:

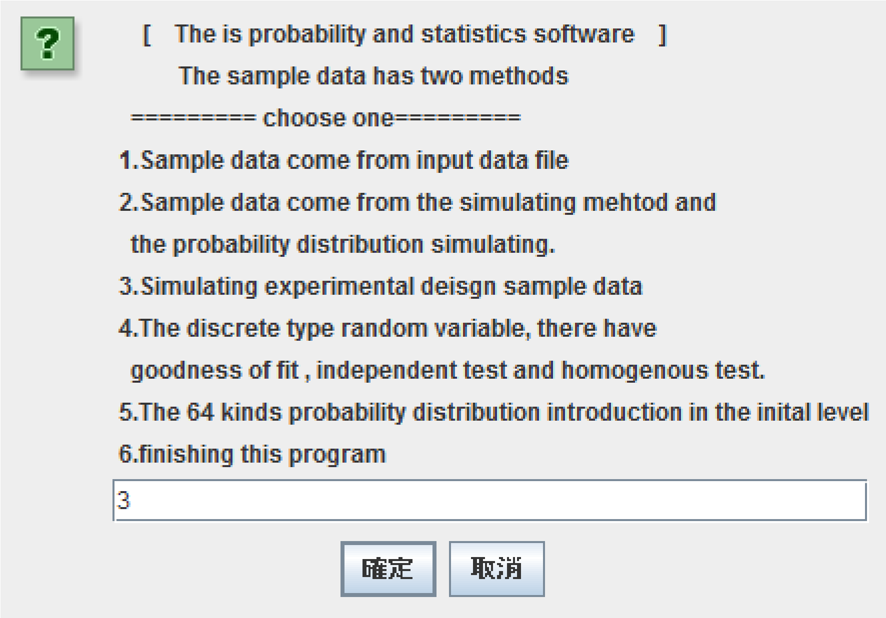
$(1-\alpha)\times 100\% \quad C.I. \quad for \quad \mu_{i\bullet} - \mu_{j\bullet} (i \neq j, i = 1,2,.., A, j = 1,2,.., A)$

$$(\overline{Y}_{i\bullet} - \overline{Y}_{j\bullet}) \pm critical \quad value \times \sqrt{MSE} \times \sqrt{\frac{1}{B} + \frac{1}{B}}$$

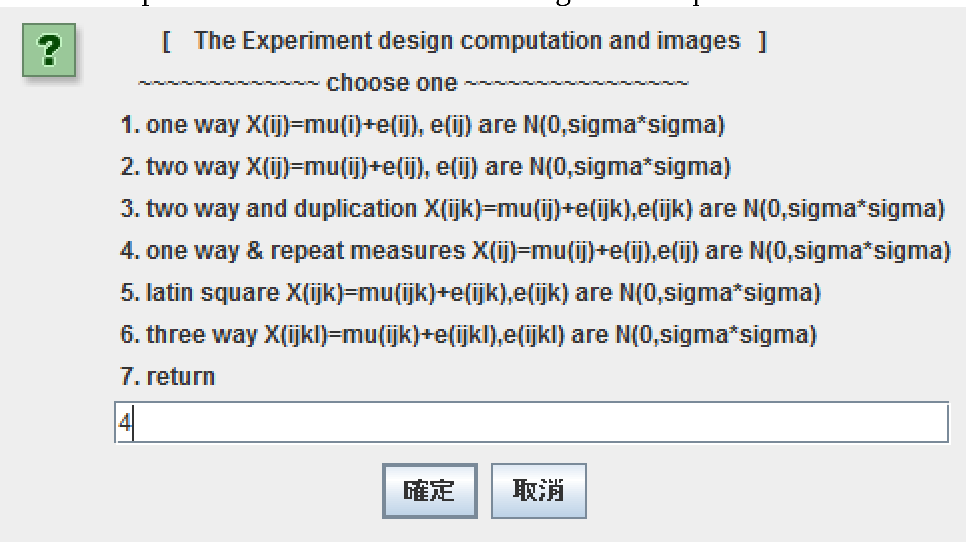
Error is normal is LSD, it is $t_{\alpha/2,(A-1)(B-1)}$,

Error is non-normal :The critical value is gotten by simulation.

4.5)



The sample data is simulated in according to the requirements.



4.5.1)The number of factor A is 4 and the number of factor B is 3.

(A=4,B=5). $X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, i = 1,.., A, j = 1,..., B,$

$\alpha_1 = \alpha_2 = ... = \alpha_A = 0, \beta_1 = \beta_2 = ... = \beta_B = 0,$

4.5.1.1) (A=4,B=5). $X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, i = 1,.., A, j = 1,..., B,$

$\alpha_1 = \alpha_2 = ... = \alpha_A = 0,$

$\mu = -10,\ \varepsilon_{ij}$ ~ Normal distribution, $E(\varepsilon_{ij}) = 0, Var(\varepsilon_{ij}) = 625,$

X(1,1)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(1,2)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(1,3)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(1,4)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(1,5)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(2,1)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(2,2)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(2,3)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(2,4)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(2,5)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(3,1)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(3,2)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(3,3)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(3,4)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(3,5)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(4,1)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(4,2)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(4,3)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(4,4)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5
X(4,5)~Normal(mu=-10.000000,sigma*sigma=625.000000), sample size=5

The one way & repeated measures

X(ij)=mu+alpha(i)+beta(j)+e(ij),i=1,..,4, j=1,..,5
mu=-10.000000,
alpha(1)=0.000000, alpha(2)=0.000000, alpha(3)=0.000000, alpha(4)=0.000000,
eij iid ~Normal(0,sigma*sigma=25.000000)

	A1	A2	A3	A4
B1 |	-3.3190210134	-12.5548779657	-8.7045503611	-12.7473242440
B2 |	-18.1506306991	-12.0601782934	-11.2226933395	-9.7923320936
B3 |	-11.3719232517	-12.5265421798	-1.0400630358	-4.6733034163
B4 |	-3.5981593613	-1.4816464077	-12.1618264368	-13.2529463412
B5 |	-5.3891936464	-14.9758269381	-10.5608350087	-4.4394103641

One way & repeated measures
X(ij)=mu+alpha(i)+beta(j)+e(ij), i=1,2...,4, j=1,2...,5

	A1	A2	A3	A4
factor A sample mean	-8.36579	-10.71981	-8.73799	-8.98106
alpha estimate value	0.83538	-1.51865	0.46317	0.22010

summation of alpha(i)=-0.000000

ANOVA

Source	df	SS	MS	F
Factor A	3	16.3356354732	5.4452118244	0.1861543633
between repeated	4	75.4668976613		
Error	12	351.0126796303	29.2510566359	
Total	19	442.8152127648		

H0:alpha(1)=...=alpha(4)=0
F(3,12) test value=0.186154

The F test p value=0.903800

class	[1]	[2]	[3]	[4]	[5]
lower limit		-4.55180	-1.37004	1.37009	4.55147
upper limit	-4.55180	-1.37004	1.37009	4.55147	
observed no	5.00000	4.00000	2.00000	6.00000	3.00000
probability	0.20000	0.20000	0.20000	0.20000	0.20000
expected no	4.00000	4.00000	4.00000	4.00000	4.00000
chi square	0.25000	0.00000	1.00000	1.00000	0.25000

degree of freedom=3

H0: residual~Normal(0,sigma(error)*sigma(error)), sigma(error) are unknown
pearson chi-square test statistic =2.500000
p-value=0.475200

H0: Variances are equal
Max(residual(ij)*residual(ij)/SSE=test value=0.167189
p value=0.866613

~~~~~ The run test of residual~~~~~~~~~~~~~
number of the negative of residual=9
number of the positive ofresidual=11
Run=11
H0: residualis random , H1: Increasing line or decreasing line
Z=0.046437, p-value=0.518600
H0: residual is random , H1: Oscillation
Z=0.046437, p-value=0.481400
H0: residual is random , H1: Increasing line or decreasing line or Oscillation
Z=0.046437, p-value=0.962800

multiple comparison of population means

Factor A ,there has 4 cateogries
. LSD(least significant difference)
The confidence coefficietn=0.95
95% C.I. for mu(1)-mu(2)
[-5.0983267467, 9.80638427180]
mu(1)=mu(2)

95% C.I. for mu(1)-mu(3)
[-7.0801474673, 7.82456355120]
mu(1)=mu(3)
~~~~~

95% C.I. for mu(1)-mu(4)
[-6.8370778118, 8.06763320670]
mu(1)=mu(4)

95% C.I. for mu(2)-mu(3)
[-9.4341762298, 5.47053478870]
mu(2)=mu(3)

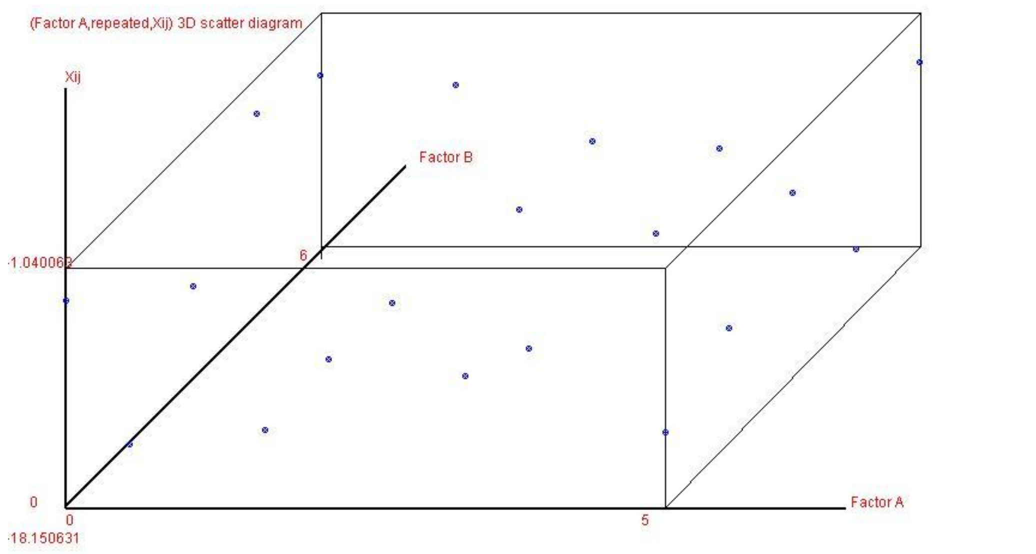
95% C.I. for mu(2)-mu(4)
[-9.1911065743, 5.71360444410]
mu(2)=mu(4)

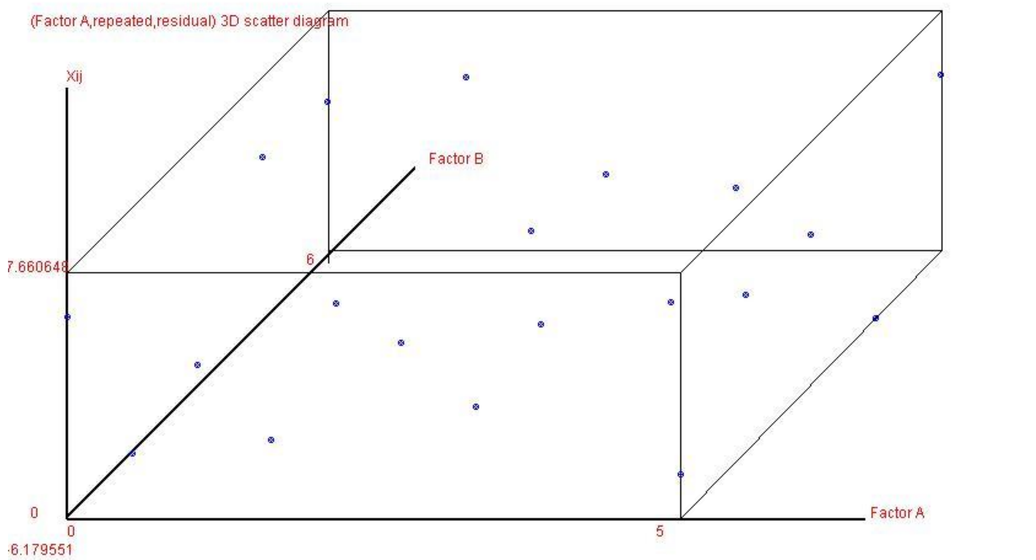
95% C.I. for mu(3)-mu(4)
[-7.2092858537, 7.69542516470]
mu(3)=mu(4)

~~~~~~~~~~~ error ~~~~~~~~~~~

5.17704	-1.70478	0.16372	-3.63598
-6.17955	2.26493	1.12059	2.79403
-4.80434	-3.60493	5.89972	2.50955
3.19011	7.66065	-5.00135	-5.84940
2.61674	-4.61586	-2.18269	4.18181

The common population standard deviation and variance confidence interval
90% confidence interval for population variance
[16.694230 , 67.166788]
90% confidence interval for population standard deviation
[4.085857 , 8.195535]
95% confidence interval for population variance
[15.041023 , 79.706914]
95% confidence interval for population standard deviation
[3.878276 , 8.927873]
99% confidence interval for population variance
[12.403121 , 114.192317]
99% confidence interval for population standard deviation
[3.521807 , 10.686081]
~~~~~~~~~~~


(Factor A,repeated,Xij) 3D scatter diagram
Xij
Factor B
-1.040062
6
0
0
5
Factor A
-18.150631


(Factor A,repeated,residual) 3D scatter diagram
Xij
Factor B
7.660648
6
0
0
5
Factor A
-6.179551

4.5.1.2) (A=4,B=5). $X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, i = 1,.., A, j = 1,..., B,$

$\alpha_1 = 4, \alpha_2 = 5, .\alpha_2 = 10, .\alpha_A = 12,$

$\mu = -10,\ \varepsilon_{ij}$ ~ Uniform distribution, $E(\varepsilon_{ij}) = 0, Var(\varepsilon_{ij}) = 25,$

X(1,1)~Normal(mu=-6.000000,sigma*sigma=25.000000), sample size=5
X(1,2)~Normal(mu=-6.000000,sigma*sigma=25.000000), sample size=5
X(1,3)~Normal(mu=-6.000000,sigma*sigma=25.000000), sample size=5
X(1,4)~Normal(mu=-6.000000,sigma*sigma=25.000000), sample size=5
X(1,5)~Normal(mu=-6.000000,sigma*sigma=25.000000), sample size=5
X(2,1)~Normal(mu=-5.000000,sigma*sigma=25.000000), sample size=5
X(2,2)~Normal(mu=-5.000000,sigma*sigma=25.000000), sample size=5
X(2,3)~Normal(mu=-5.000000,sigma*sigma=25.000000), sample size=5
X(2,4)~Normal(mu=-5.000000,sigma*sigma=25.000000), sample size=5
X(2,5)~Normal(mu=-5.000000,sigma*sigma=25.000000), sample size=5
X(3,1)~Normal(mu=0.000000,sigma*sigma=25.000000), sample size=5
X(3,2)~Normal(mu=0.000000,sigma*sigma=25.000000), sample size=5
X(3,3)~Normal(mu=0.000000,sigma*sigma=25.000000), sample size=5
X(3,4)~Normal(mu=0.000000,sigma*sigma=25.000000), sample size=5
X(3,5)~Normal(mu=0.000000,sigma*sigma=25.000000), sample size=5
X(4,1)~Normal(mu=2.000000,sigma*sigma=25.000000), sample size=5
X(4,2)~Normal(mu=2.000000,sigma*sigma=25.000000), sample size=5
X(4,3)~Normal(mu=2.000000,sigma*sigma=25.000000), sample size=5
X(4,4)~Normal(mu=2.000000,sigma*sigma=25.000000), sample size=5
X(4,5)~Normal(mu=2.000000,sigma*sigma=25.000000), sample size=5

The one way & repeated measures

X(ij)=mu+alpha(i)+beta(j)+e(ij),i=1,..,4, j=1,..,5
mu=-10.000000,
alpha(1)=4.000000, alpha(2)=5.000000, alpha(3)=10.000000, alpha(4)=12.000000,
eij iid ~Normal(0,sigma*sigma=5.000000)

	A1	A2	A3	A4
B1 |	-5.2548236901	-6.5104277192	-3.1013078945	5.2836194915
B2 |	-10.9002114413	-5.1519132808	0.9326308555	1.2226114453
B3 |	-6.0224612809	-0.4133497419	0.9281081007	2.3512276751
B4 |	-4.1543495200	-8.2652057448	-1.2734966540	2.0019190065
B5 |	-7.9356213969	-3.9933408012	2.8363629840	-0.9107740844

One way & repeated measures

X(ij)=mu+alpha(i)+beta(j)+e(ij), i=1,2...,4, j=1,2...,5

	A1	A2	A3	A4
factor A sample mean	-6.85349	-4.86685	0.06446	1.98972
alpha estimate value	-4.43695	-2.45031	2.48100	4.40626

summation of alpha(i)=0.000000

ANOVA

Source	df	SS	MS	F
Factor A	3	256.3052726117	85.4350908706	11.6688117695
between repeated	4	16.1240403734		
Error	12	87.8599390154	7.3216615846	
Total	19	360.2892520005		

H0:alpha(1)=...=alpha(4)=0
F(3,12) test value=11.668812

The F test p value=0.000800

class	[1]	[2]	[3]	[4]	[5]
lower limit		-2.27729	-0.68544	0.68546	2.27712
upper limit	-2.27729	-0.68544	0.68546	2.27712	
observed no	4.00000	6.00000	2.00000	4.00000	4.00000
probability	0.20000	0.20000	0.20000	0.20000	0.20000
expected no	4.00000	4.00000	4.00000	4.00000	4.00000
chi square	0.00000	1.00000	1.00000	0.00000	0.00000

degree of freedom=3

H0: residual~Normal(0,sigma(error)*sigma(error)), sigma(error) are unknown
pearson chi-square test statistic =2.000000
p-value=0.572400

H0: Variances are equal
Max(residual(ij)*residual(ij)/SSE=test value=0.121934
p value=0.995150

~~~~~ The run test of residual~~~~~~~~~~~~~
number of the negative of residual=10
number of the positive ofresidual=10
Run=12
H0: residualis random , H1: Increasing line or decreasing line
Z=0.459468, p-value=0.677100
H0: residual is random , H1: Oscillation
Z=0.459468, p-value=0.322900
H0: residual is random , H1: Increasing line or decreasing line or Oscillation
Z=0.459468, p-value=0.645800

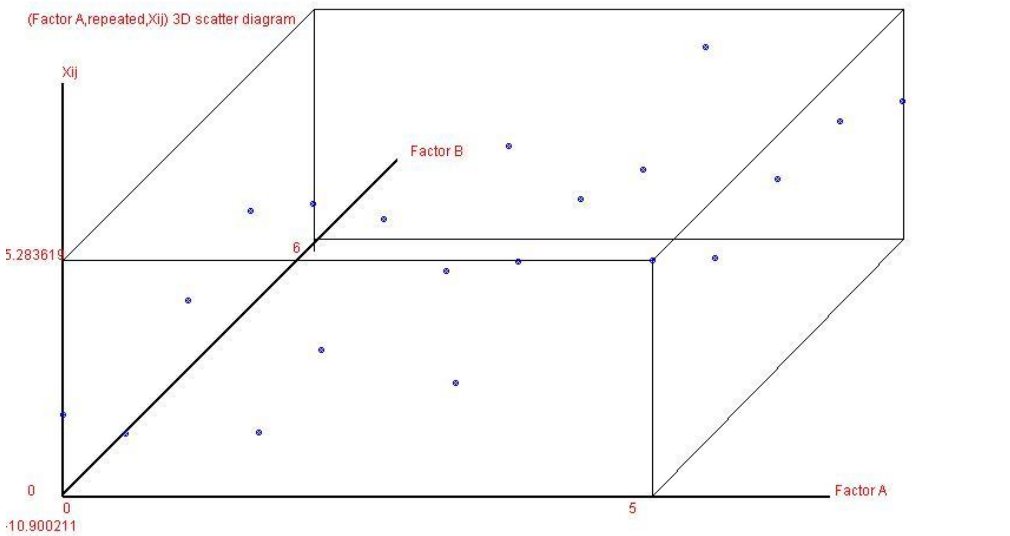
multiple comparison of population means

Factor A ,there has 4 cateogries
. LSD(least significant difference)
The confidence coefficietn=0.95
95% C.I. for mu(1)-mu(2)
[-5.7150898903, 1.74179787380]
mu(1)=mu(2)
~~~~~

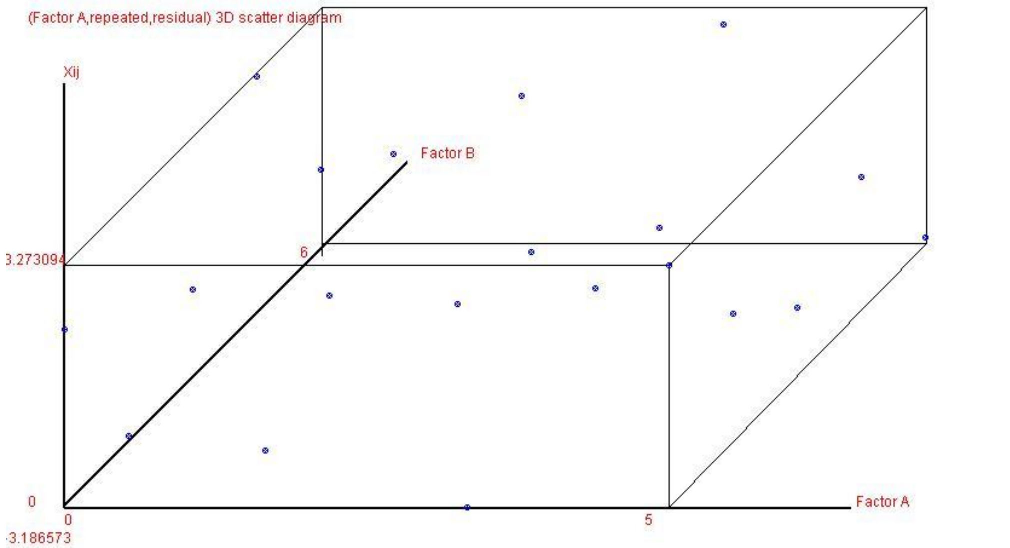
```
     95% C.I. for mu(1)-mu(3)
[          -10.6463968262,            -3.18950906210]
   mu(1)<mu(3)
     95% C.I. for mu(1)-mu(4)
[          -12.5716580547,            -5.11477029060]
   mu(1)<mu(4)
     95% C.I. for mu(2)-mu(3)
[           -8.6597508180,            -1.20286305380]
   mu(2)<mu(3)
     95% C.I. for mu(2)-mu(4)
[          -10.5850120465,            -3.12812428230]
   mu(2)<mu(4)
     95% C.I. for mu(3)-mu(4)
[           -5.6537051106,             1.80318265360]
   mu(3)=mu(4)


              ~~~~~~~~~~~~ error ~~~~~~~~~~~~
       1.57786      -1.66439      -3.18657       3.27309
      -2.98904       0.77261       1.92585       0.29057
      -0.79639       2.82608      -0.76377      -1.26591
       3.20539      -2.89212      -0.83171       0.51844
      -0.99782       0.95781       2.85621      -2.81619


  The common population standard deviation and variance confidence interval
   90% confidence interval for population variance
   [4.178636 ,    16.812127]
   90% confidence interval for population standard deviation
   [2.044171 ,    4.100259]
   95% confidence interval for population variance
   [3.764831 ,    19.950973]
   95% confidence interval for population standard deviation
   [1.940317 ,    4.466651]
   99% confidence interval for population variance
   [3.104553 ,    28.582814]
   99% confidence interval for population standard deviation
   [1.761974 ,    5.346290]
```



(Factor A,repeated,Xij) 3D scatter diagram
Xij
Factor B
5.283619
6
0
0
-10.900211
5
Factor A



(Factor A,repeated,residual) 3D scatter diagram
Xij
Factor B
3.273094
6
0
0
-3.186573
5
Factor A