ASG2 Solutions

Ex 3-1:

a). $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

$H_1$: Not all means are equal

**ANOVA**

STRENGTH

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 489740.2 | 3 | 163246.729 | 12.728 | .000 |
| Within Groups | 153908.3 | 12 | 12825.688 | | |
| Total | 643648.4 | 15 | | | |

F-test statistic value is 12.728.

From F-table with 3 numerator df and 12 denominator df at 5% level is 3.49

Since, F-statistic value is larger than 3.49, reject $H_0$ at $\alpha=0.05$

Or

P-value is less than 0.001 which is small compared 0.05 hence reject the null hypothesis.

b).

**Multiple Comparisons**

Dependent Variable: STRENGTH

| | (I) TECHNIQU | (J) TECHNIQU | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval Lower Bound | 95% Confidence Interval Upper Bound |
|---|---|---|---|---|---|---|---|
| LSD | 1.00 | 2.00 | -185.2500* | 80.0802 | .039 | -359.7298 | -10.7702 |
| | | 3.00 | 37.2500 | 80.0802 | .650 | -137.2298 | 211.7298 |
| | | 4.00 | 304.7500* | 80.0802 | .003 | 130.2702 | 479.2298 |
| | 2.00 | 1.00 | 185.2500* | 80.0802 | .039 | 10.7702 | 359.7298 |
| | | 3.00 | 222.5000* | 80.0802 | .017 | 48.0202 | 396.9798 |
| | | 4.00 | 490.0000* | 80.0802 | .000 | 315.5202 | 664.4798 |
| | 3.00 | 1.00 | -37.2500 | 80.0802 | .650 | -211.7298 | 137.2298 |
| | | 2.00 | -222.5000* | 80.0802 | .017 | -396.9798 | -48.0202 |
| | | 4.00 | 267.5000* | 80.0802 | .006 | 93.0202 | 441.9798 |
| | 4.00 | 1.00 | -304.7500* | 80.0802 | .003 | -479.2298 | -130.2702 |
| | | 2.00 | -490.0000* | 80.0802 | .000 | -664.4798 | -315.5202 |
| | | 3.00 | -267.5000* | 80.0802 | .006 | -441.9798 | -93.0202 |

*. The mean difference is significant at the .05 level.

Duncan's Test

**STRENGTH**

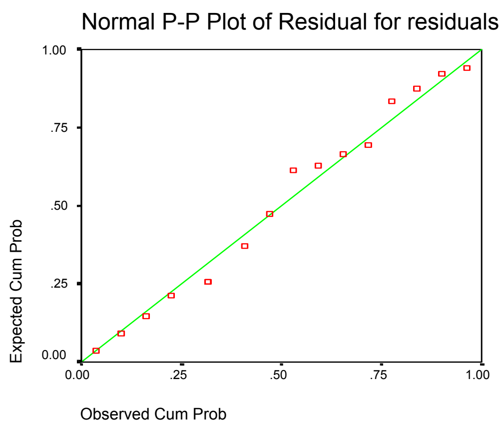
| | TECHNIQU | N | Subset for alpha = .05 | | |
|---|---|---|---|---|---|
| | | | 1 | 2 | 3 |
| Duncan[a] | 4.00 | 4 | 2666.2500 | | |
| | 3.00 | 4 | | 2933.7500 | |
| | 1.00 | 4 | | 2971.0000 | |
| | 2.00 | 4 | | | 3156.2500 |
| | Sig. | | 1.000 | .650 | 1.000 |

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 4.000.

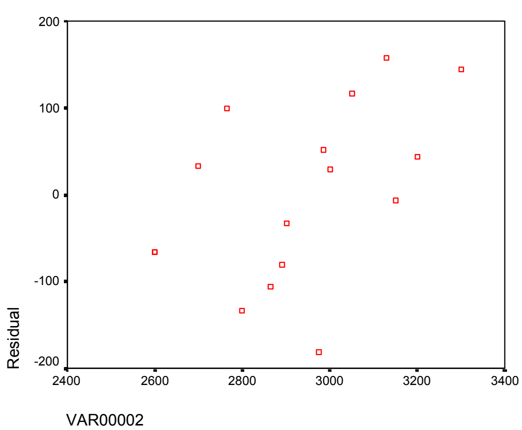
**d).**

Normal P-P plot




Normality assumptions is reasonable.

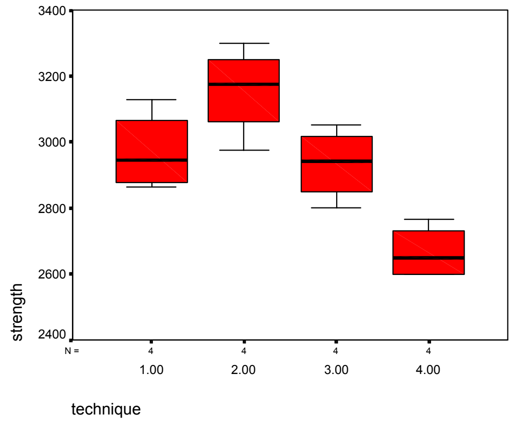
**e).**




There is some systematic pattern exist. Smaller value of residual correspond to smaller values of the outcome variable (var0002).

f). Box Plot:

1. Techniques 1 and 3 look 'alike'
2. Technique 2 is 'different' from other three techniques.
3. Technique 4 is different from other three.
4. Technique 2 has highest mean and technique 4 has the lowest mean.




Ex 3-4

a).

**ANOVA**

DENSITY

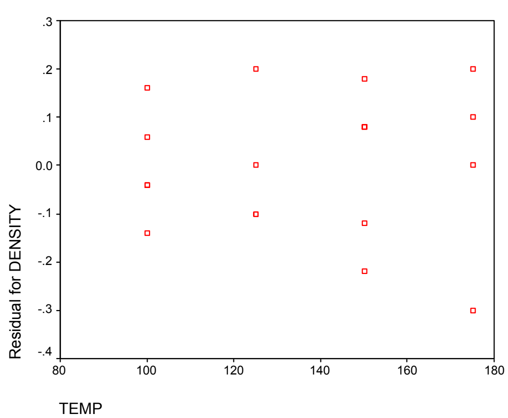
| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | .156 | 3 | .052 | 2.024 | .157 |
| Within Groups | .360 | 14 | .026 | | |
| Total | .516 | 17 | | | |

Since the p-value is 0.157 for testing difference among group means, we do no reject $H_0$: $\mu_1=\mu_2=\mu_3=\mu_4$ .

No statistical differences among mean density for 4 temp. settings.

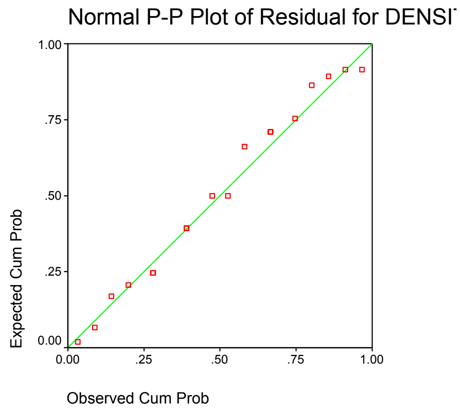
c).

Plot of residuals vs groups (Temp).



Spread (variation) of residuals across temp groups looks approximately equal and hence the constant variance assumption for the error term looks to be true.

P-P plot to validate normality assumption.




Since, all the points approximately lie on the diagonal line, the normal assumption on the error term looks to be true.

d). $\sqrt{MS_E / n} = \sqrt{0.026 / n}$

EX 3-6

a). $H_0$: $\mu_1=\mu_2=\mu_3=\mu_4$.

$H_1$: Not all means are same.

**ANOVA**

CONDUCTI

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 844.688 | 3 | 281.563 | 14.302 | .000 |
| Within Groups | 236.250 | 12 | 19.688 | | |
| Total | 1080.938 | 15 | | | |

At 5% level of significance, we reject the null hypothesis since the p-value is smaller than 0.001 and conclude that mean conductivity differ statistically across 4 coating types.

b).

**Descriptives**

CONDUCTI

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| 1.00 | 4 | 145.0000 | 3.91578 | 1.95789 | 138.7691 | 151.2309 | 141.00 | 150.00 |
| 2.00 | 4 | 145.2500 | 6.65207 | 3.32603 | 134.6651 | 155.8349 | 137.00 | 152.00 |
| 3.00 | 4 | 132.2500 | 3.86221 | 1.93111 | 126.1044 | 138.3956 | 127.00 | 136.00 |
| 4.00 | 4 | 129.2500 | 2.06155 | 1.03078 | 125.9696 | 132.5304 | 127.00 | 132.00 |
| Total | 16 | 137.9375 | 8.48896 | 2.12224 | 133.4141 | 142.4609 | 127.00 | 152.00 |

**95% CI for coating type 4 is [125.9696, 132.5304]**
**99% CI for the mean difference between coating types 1 and 4 is given by**
**[145-129.25]+/-(2.545*** $\sqrt{MS_E / n}$ ) which is equivalent to

**[145-129.25]+/-(2.545*** $\sqrt{19.688/4}$ ).

d).

**Multiple Comparisons**

Dependent Variable: CONDUCTI
LSD

| (I) TYPE | (J) TYPE | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 1.00 | 2.00 | -.2500 | 3.13748 | .938 | -7.0860 | 6.5860 |
| | 3.00 | 12.7500* | 3.13748 | .002 | 5.9140 | 19.5860 |
| | 4.00 | 15.7500* | 3.13748 | .000 | 8.9140 | 22.5860 |
| 2.00 | 1.00 | .2500 | 3.13748 | .938 | -6.5860 | 7.0860 |
| | 3.00 | 13.0000* | 3.13748 | .001 | 6.1640 | 19.8360 |
| | 4.00 | 16.0000* | 3.13748 | .000 | 9.1640 | 22.8360 |
| 3.00 | 1.00 | -12.7500* | 3.13748 | .002 | -19.5860 | -5.9140 |
| | 2.00 | -13.0000* | 3.13748 | .001 | -19.8360 | -6.1640 |
| | 4.00 | 3.0000 | 3.13748 | .358 | -3.8360 | 9.8360 |
| 4.00 | 1.00 | -15.7500* | 3.13748 | .000 | -22.5860 | -8.9140 |
| | 2.00 | -16.0000* | 3.13748 | .000 | -22.8360 | -9.1640 |
| | 3.00 | -3.0000 | 3.13748 | .358 | -9.8360 | 3.8360 |

*. The mean difference is significant at the .05 level.

Significant differences exist between types 1 and 3 (p-value =0.002), between types 1 and 4 (with p-value <0.001), between types 2 and 3 (p-value=0.001), between types 2 and 4 (p-value<0.001). Hence, statistically (at 5% level) we can conclude that types 1 and 2 as one group and types 3 and 4 as another group with respect to mean conductivity.

f).Since minimum value for the mean conductivity is required and type 4 being currently used which has a minimum value for the mean conductivity the recommendation to the manufacturer is to continue with the current type and the current type (type 4 ) is not statistically different from type 3.

Ex 3-9:

a).

$H_0$: $\mu_1=\mu_2=\mu_3=\mu_4=\mu_5=\mu_6$ .

$H_1$: Not all means are same.

**ANOVA**

RADON

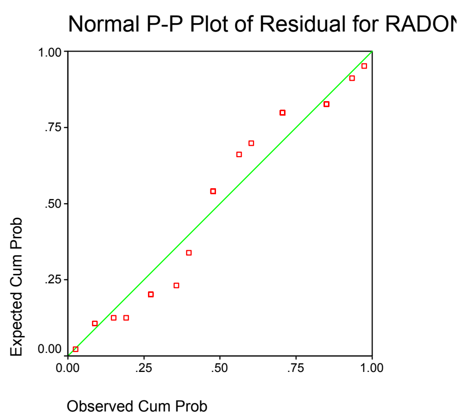
| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 1133.375 | 5 | 226.675 | 30.852 | .000 |
| Within Groups | 132.250 | 18 | 7.347 | | |
| Total | 1265.625 | 23 | | | |

Reject the null hypothesis since the p-value (<0.001) is smaller than 0.05 and hence conclude that size of the orifice statistically (at 5% level of significance) affect the mean percentage of radon released.

b).P-value is small the probability that F-statistics is larger than 30.852 when the null hypothesis is assumed to be true. In this case the p-value is <0.001 (In the exam no need to compute the exact p-value, just give reasonable upper boundary if the value is small, or lower boundary if the value is larger than 0.10)

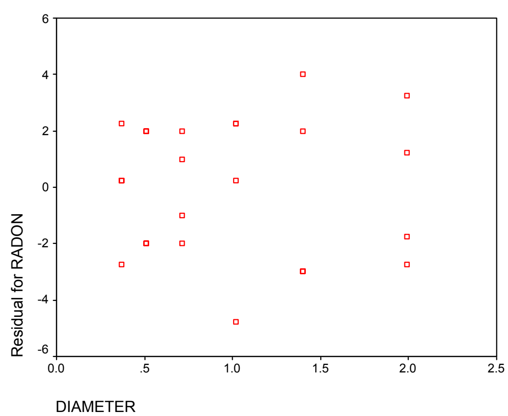
c). Residual analysis

i). P-P plot to test normality




Since, all the points approximately lie on the diagonal line, the normal assumption on the error term looks to be true.

ii). To test constant variance across treatment groups.




Spread (variation) of residuals across diameter groups looks approximately equal and hence the constant variance assumption for the error term looks to be true.

d).

**Descriptives**

RADON

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean Lower Bound | 95% Confidence Interval for Mean Upper Bound | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| .37 | 4 | 82.7500 | 2.06155 | 1.03078 | 79.4696 | 86.0304 | 80.00 | 85.00 |
| .51 | 4 | 77.0000 | 2.30940 | 1.15470 | 73.3252 | 80.6748 | 75.00 | 79.00 |
| .71 | 4 | 75.0000 | 1.82574 | .91287 | 72.0948 | 77.9052 | 73.00 | 77.00 |
| 1.02 | 4 | 71.7500 | 3.30404 | 1.65202 | 66.4925 | 77.0075 | 67.00 | 74.00 |
| 1.40 | 4 | 65.0000 | 3.55903 | 1.77951 | 59.3368 | 70.6632 | 62.00 | 69.00 |
| 1.99 | 4 | 62.7500 | 2.75379 | 1.37689 | 58.3681 | 67.1319 | 60.00 | 66.00 |
| Total | 24 | 72.3750 | 7.41803 | 1.51420 | 69.2426 | 75.5074 | 60.00 | 85.00 |

95%CI for the mean radon when the diameter is 1.40 is given by
[59.3368, 70.6632]

Ex 3-10

a). $H_0$: $\mu_1=\mu_2=\mu_3$ .
$H_1$: Not all means are same.

**ANOVA**

RESP_TIM

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 543.600 | 2 | 271.800 | 16.083 | .000 |
| Within Groups | 202.800 | 12 | 16.900 | | |
| Total | 746.400 | 14 | | | |

Reject the null hypothesis at 1% level of significance since the p-value is smaller than 0.001 and hence conclude that mean response time differ statistically (at 1% level of significance) across 3 circuit types.

b). Tukey’s test

**Multiple Comparisons**

Dependent Variable: RESP_TIM
Tukey HSD

| (I) CIRC_TYP | (J) CIRC_TYP | Mean Difference (I-J) | Std. Error | Sig. | 99% Confidence Interval Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|
| 1.00 | 2.00 | -11.4000* | 2.60000 | .002 | -20.6768 | -2.1232 |
| | 3.00 | 2.4000 | 2.60000 | .637 | -6.8768 | 11.6768 |
| 2.00 | 1.00 | 11.4000* | 2.60000 | .002 | 2.1232 | 20.6768 |
| | 3.00 | 13.8000* | 2.60000 | .001 | 4.5232 | 23.0768 |
| 3.00 | 1.00 | -2.4000 | 2.60000 | .637 | -11.6768 | 6.8768 |
| | 2.00 | -13.8000* | 2.60000 | .001 | -23.0768 | -4.5232 |

*. The mean difference is significant at the .01 level.

Types 1 and 2, and types 2 and 3 are statistically different with respect to mean response time at 1 % level of signficance. However, types 1 and 3 are not statistically different.

e).

**Descriptives**

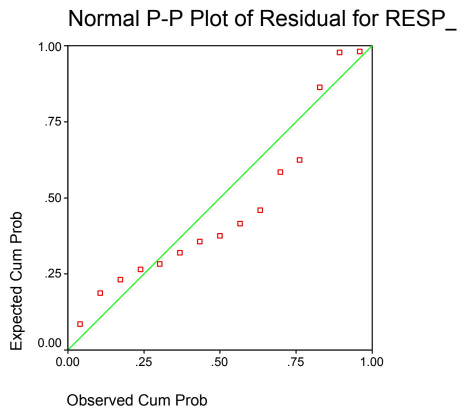
RESP_TIM

| | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean Lower Bound | Upper Bound | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| 1.00 | 5 | 10.8000 | 2.77489 | 1.24097 | 7.3545 | 14.2455 | 8.00 | 15.00 |
| 2.00 | 5 | 22.2000 | 4.86826 | 2.17715 | 16.1553 | 28.2447 | 17.00 | 30.00 |
| 3.00 | 5 | 8.4000 | 4.39318 | 1.96469 | 2.9452 | 13.8548 | 5.00 | 16.00 |
| Total | 15 | 13.8000 | 7.30166 | 1.88528 | 9.7565 | 17.8435 | 5.00 | 30.00 |

Since, mean response time is required type 3 has the least mean response time. Further, type 1 and type 3 are not statistically different, but both types are different from type 2, we recommend type 3 or type 1.

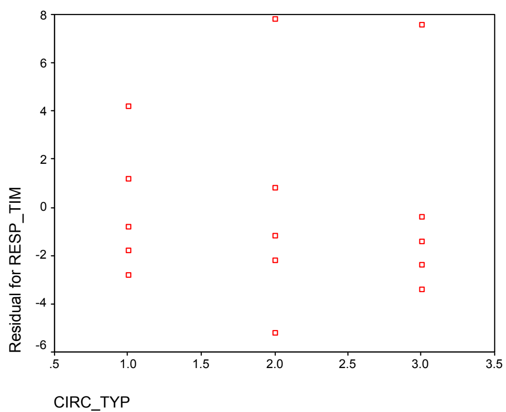
f).

i). Normality assumption validation.




Since all point are not (at least roughly) on the diagonal line we suspect the normality assumption.

ii). Constant variance assumption validation.




Residuals for Types 2 and 3 tend to have more variations than residuals for type I we can suspect constant variance assumption.

Ex 4-2

$H_0$: $\mu_1=\mu_2=\mu_3$ .

$H_1$: Not all means are same.

**Tests of Between-Subjects Effects**

Dependent Variable: GROWTH

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Corrected Model | 1810.417[a] | 5 | 362.083 | 41.913 | .000 |
| Intercept | 4218.750 | 1 | 4218.750 | 488.344 | .000 |
| BLOCK | 1106.917 | 3 | 368.972 | 42.711 | .000 |
| SOLUTION | 703.500 | 2 | 351.750 | 40.717 | .000 |
| Error | 51.833 | 6 | 8.639 | | |
| Total | 6081.000 | 12 | | | |
| Corrected Total | 1862.250 | 11 | | | |

a. R Squared = .972 (Adjusted R Squared = .949)

Reject the null hypothesis (equality of means for 3 solutions) even at 1% level of significance since the p-value ( corresponding to solution) is <0.001 and hence conclude that mean effectiveness in retarding bacteria growth is significantly (at 1% level) different across 3 solutions.

Ex4-5

a). $H_0$: $\mu_1=\mu_2=\mu_3=\mu_4=\mu_5$ .

$H_1$: Not all means are same.

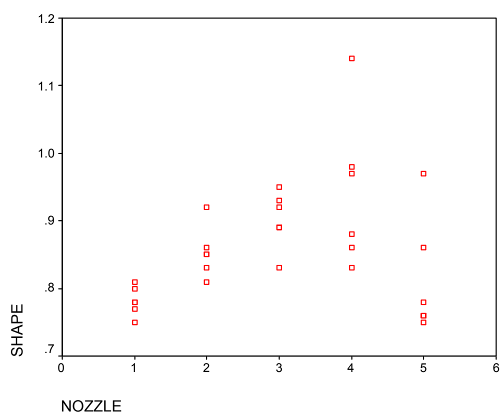
**Tests of Between-Subjects Effects**

Dependent Variable: SHAPE

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Corrected Model | .165[a] | 9 | 1.834E-02 | 6.401 | .000 |
| Intercept | 22.119 | 1 | 22.119 | 7720.507 | .000 |
| NOZZLE | .102 | 4 | 2.555E-02 | 8.916 | .000 |
| JET_VELO | 6.287E-02 | 5 | 1.257E-02 | 4.389 | .007 |
| Error | 5.730E-02 | 20 | 2.865E-03 | | |
| Total | 22.342 | 30 | | | |
| Corrected Total | .222 | 29 | | | |

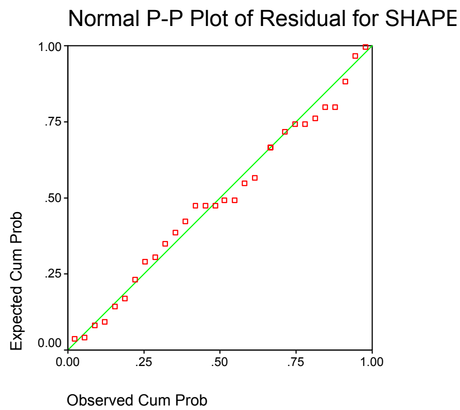
a. R Squared = .742 (Adjusted R Squared = .626)

Reject the null hypothesis (equality of means for 5 types of Nozzles) at 5% level of significance since the p-value (corresponding to NOZZLE) is <0.001 and hence conclude that nozzle design affect the shape factor.




By looking at the above plot, we can conclude that shape varies with nozzle design.

b). Since, all the points approximately lie on the diagonal line, the normal assumption on the error term looks to be true (see the plot below).




Spread (variation) of residuals across nozzle groups looks approximately equal (see the plot below) and hence the constant variance assumption for the error term looks to be true.