**Homework No 2**

1. A manufacturer suspects that the batches of raw material furnished by his supplier differ significantly in calcium content. There are a large number of batches currently in the warehouse. Five of these are randomly selected for study. A chemist makes five determinations on each batch as obtains the following data:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Batch 1 | Batch 2 | Batch 3 | Batch 4 | Batch 5 |
| 23.46 | 23.59 | 23.51 | 23.28 | 23.29 |
| 23.48 | 23.46 | 23.64 | 23.40 | 23.46 |
| 23.56 | 23.42 | 23.46 | 23.37 | 23.37 |
| 23.39 | 23.49 | 23.52 | 23.46 | 23.32 |
| 23.40 | 23.50 | 23.49 | 23.39 | 23.38 |

1. Is there significant variation in the calcium content from batch to batch? Use *α*=0.05.
2. Estimate the components of variance.
3. Find a 95 percent confidence interval for 
4. Analyze the residuals from this experiment. Are the analysis of variance assumptions are satisfied?
5. Four chemists are asked to determine the percentage of methyl alcohol in a certain chemical compound. Each chemist makes three determinations, and the results are the following:

|  |  |  |  |
| --- | --- | --- | --- |
| Chemist | Percentage of Methyl Alcohol | | |
| 1 | 84.99 | 84.04 | 84.38 | |
| 2 | 85.15 | 85.13 | 84.88 | |
| 3 | 84.72 | 84.48 | 85.16 | |
| 4 | 84.20 | 84.10 | 84.55 | |

1. Do chemists differ significantly? Use *α* = 0.05.
2. Analyze the residuals from this experiment.
3. If chemist 2 is a new employee, construct a meaningful set of orthogonal contrasts that might have been useful at the start of the experiment.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Chemists | Total | C1 | C2 | C3 |
| 1 | 253.41 | 1 | -2 | 0 |
| 2 | 255.16 | -3 | 0 | 0 |
| 3 | 254.36 | 1 | 1 | -1 |
| 4 | 252.85 | 1 | 1 | 1 |

1. Consider testing the equality of the means of two normal populations, where the variances are unknown but are assumed to be equal. The appropriate test procedure is the pooled *t* test. Show that the pooled *t* test is equivalent to the single factor analysis of variance.
2. A chemist wishes to test the effect of four chemical agents on the strength of a particular type of cloth. Because there might be variability from one bolt to another, the chemist decides to use a randomized block design, with the bolts of cloth considered as blocks. She selects five bolts and applies all four chemicals in random order to each bolt. The resulting tensile strengths follow.
3. Analyze the data from this experiment (use *α* = 0.05) and draw appropriate conclusions.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | Bolt |  |  |
| Chemical | 1 | 2 | 3 | 4 | 5 |
| 1 | 73 | 68 | 74 | 71 | 67 |
| 2 | 73 | 67 | 75 | 72 | 70 |
| 3 | 75 | 68 | 78 | 73 | 68 |
| 4 | 73 | 71 | 75 | 75 | 69 |

1. Plot the mean tensile strengths observed for each chemical type and compare them to a scaled *t* distribution. What conclusions would you draw from the display?
2. An article in *Nature Genetics* (2003, Vol. 34, pp. 85-90) “Treatment-Specific Changes in Gene Expression Discriminate in vivo Drug Response in Human Leukemia Cells” studied gene expressionas a function of different treatments for leukemia. Three treatment groups are: mercaptopurine (MP) only; low-dose methotrexate (LDMTX) and MP; and high-dose methotrexate (HDMTX) and MP. Each group contained ten subjects. The responses from a specific gene are shown in the table below:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Project | | | | | | | | | |
| MP ONLY | 334.5 | 31.6 | 701 | 41.2 | 61.2 | 69.6 | 67.5 | 66.6 | 120.7 | 881.9 |
| MP + HDMTX | 919.4 | 404.2 | 1024.8 | 54.1 | 62.8 | 671.6 | 882.1 | 354.2 | 321.9 | 91.1 |
| MP + LDMTX | 108.4 | 26.1 | 240.8 | 191.1 | 69.7 | 242.8 | 62.7 | 396.9 | 23.6 | 290.4 |

(a) Is there evidence to support the claim that the treatment means differ?

(b) Check the normality assumption. Can we assume these samples are from normal populations?

(c) Take the logarithm of the raw data. Is there evidence to support the claim that the treatment means differ for the transformed data?

(d) Analyze the residuals from the transformed data and comment on model adequacy.

1. The effect of five different ingredients (*A, B, C, D, E*) on reaction time of a chemical process is being studied. Each batch of new material is only large enough to permit five runs to be made. Furthermore, each run requires approximately 1 1/2 hours, so only five runs can be made in one day. The experimenter decides to run the experiment as a Latin square so that day and batch effects can be systematically controlled. She obtains the data that follow. Analyze the data from this experiment (use = 0.05) and draw conclusions.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | Day |  |  |
| Batch | 1 | 2 | 3 | 4 | 5 |
| 1 | *A*=8 | *B*=7 | *D*=1 | *C*=7 | *E*=3 |
| 2 | *C*=11 | *E*=2 | *A*=7 | *D*=3 | *B*=8 |
| 3 | *B*=4 | *A*=9 | *C*=10 | *E*=1 | *D*=5 |
| 4 | *D*=6 | *C*=8 | *E*=6 | *B*=6 | *A*=10 |
| 5 | *E*=4 | *D*=2 | *B*=3 | *A*=8 | *C*=8 |