# TMA 4255
# Forsøksplanlegging og anvendte statistiske metoder

Våren 2007

Del 2



# One-Way ANOVA

**TABLE 13.1** Absorption of Moisture in Concrete Aggregates

| Aggregate: | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| | 551 | 595 | 639 | 417 | 563 | |
| | 457 | 580 | 615 | 449 | 631 | |
| | 450 | 508 | 511 | 517 | 522 | |
| | 731 | 583 | 573 | 438 | 613 | |
| | 499 | 633 | 648 | 415 | 656 | |
| | 632 | 517 | 677 | 555 | 679 | |
| Total | 3320 | 3416 | 3663 | 2791 | 3664 | 16,854 |
| Mean | 553.33 | 569.33 | 610.50 | 465.17 | 610.67 | 561.80 |

# General setup – One-Way ANOVA

**TABLE 13.2** $k$ Random Samples

| **Treatment:** | **1** | **2** | ... | $i$ | ... | $k$ | |
|---|---|---|---|---|---|---|---|
| | $y_{11}$ | $y_{21}$ | ... | $y_{i1}$ | ... | $y_{k1}$ | |
| | $y_{12}$ | $y_{22}$ | ... | $y_{i2}$ | ... | $y_{k2}$ | |
| | ⋮ | ⋮ | | ⋮ | | ⋮ | |
| | $y_{1n}$ | $y_{2n}$ | ... | $y_{in}$ | ... | $y_{kn}$ | |
| Total | $Y_{1.}$ | $Y_{2.}$ | ... | $Y_{i.}$ | ... | $Y_{k.}$ | $Y_{..}$ |
| Mean | $\bar{y}_{1.}$ | $\bar{y}_{2.}$ | ... | $\bar{y}_{i.}$ | ... | $\bar{y}_{k.}$ | $\bar{y}_{..}$ |



**TABLE 13.3** Analysis of Variance for the One-Way ANOVA

| Source of variation | Sum of squares | Degrees of freedom | Mean square | Computed $f$ |
|---|---|---|---|---|
| Treatments | $SSA$ | $k-1$ | $s_1^2 = \frac{SSA}{k-1}$ | $\frac{s_1^2}{s^2}$ |
| Error | $SSE$ | $k(n-1)$ | $s^2 = \frac{SSE}{k(n-1)}$ | |
| Total | $SST$ | $nk-1$ | | |

One-Way ANOVA

# MINITAB analysis of Table 3.1

### Data Display

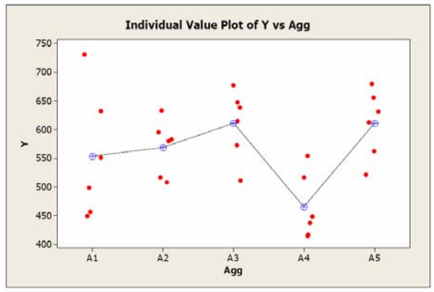
| Row | Y | Agg |
|---|---|---|
| 1 | 551 | A1 |
| 2 | 457 | A1 |
| 3 | 450 | A1 |
| 4 | 731 | A1 |
| 5 | 499 | A1 |
| 6 | 632 | A1 |
| 7 | 595 | A2 |
| 8 | 580 | A2 |
| 9 | 508 | A2 |
| 10 | 583 | A2 |
| 11 | 633 | A2 |
| 12 | 517 | A2 |
| 13 | 639 | A3 |
| 14 | 615 | A3 |
| 15 | 511 | A3 |
| 16 | 573 | A3 |
| 17 | 648 | A3 |
| 18 | 677 | A3 |
| 19 | 417 | A4 |
| 20 | 449 | A4 |
| 21 | 517 | A4 |
| 22 | 438 | A4 |
| 23 | 415 | A4 |
| 24 | 555 | A4 |
| 25 | 563 | A5 |
| 26 | 631 | A5 |
| 27 | 522 | A5 |
| 28 | 613 | A5 |
| 29 | 656 | A5 |
| 30 | 679 | A5 |

### One-way ANOVA: Y versus Agg

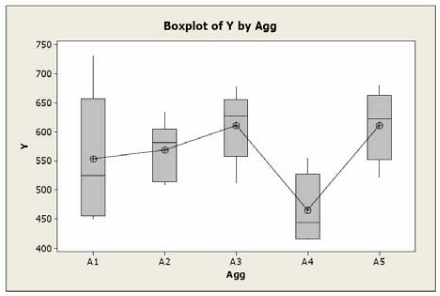
| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Agg | 4 | 85356 | 21339 | 4,30 | 0,009 |
| Error | 25 | 124020 | 4961 | | |
| Total | 29 | 209377 | | | |

S = 70,43 R-Sq = 40,77% R-Sq(adj) = 31,29%

Individual Value Plot of Y vs Agg

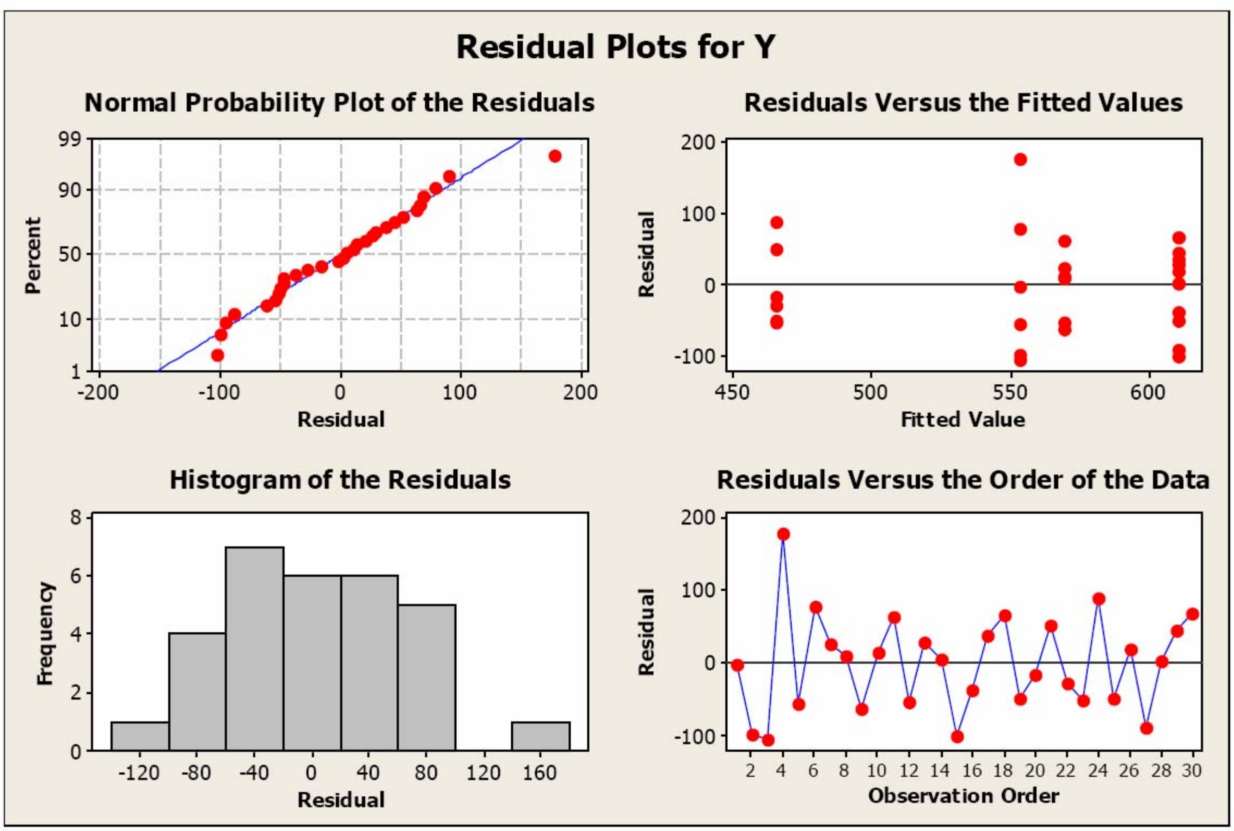



Boxplot of Y by Agg






## Residual Plots for Y



**Example 13.2** Part of a study conducted at the Virginia Polytechnic Institute and State University was designed to measure serum alkaline phosphatase activity levels (Bessey-Lowry Units) in children with seizure disorders who were receiving anticonvulsant therapy under the care of a private physician. Forty-five subjects were found for the study and categorized into four drug groups:

G-1: Control (not receiving anticonvulsants and having no history of seizure disorders)

G-2: Phenobarbital

G-3: Carbamazepine

G-4: Other anticonvulsants

From blood samples collected on each subject the serum alkaline phosphatase activity level was determined and recorded in Table 13.5. Test the hypothesis at the 0.05 level of significance that the average serum alkaline phosphatase activity level is the same for the four drug groups.



**TABLE 13.5** Serum Alkaline Phosphatase Activity Level

| G-1 | | G-2 | G-3 | G-4 |
|---|---|---|---|---|
| 49.20 | 97.50 | 97.07 | 62.10 | 110.60 |
| 44.54 | 105.00 | 73.40 | 94.95 | 57.10 |
| 45.80 | 58.05 | 68.50 | 142.50 | 117.60 |
| 95.84 | 86.60 | 91.85 | 53.00 | 77.71 |
| 30.10 | 58.35 | 106.60 | 175.00 | 150.00 |
| 36.50 | 72.80 | 0.57 | 79.50 | 82.90 |
| 82.30 | 116.70 | 0.79 | 29.50 | 111.50 |
| 87.85 | 45.15 | 0.77 | 78.40 | |
| 105.00 | 70.35 | 0.81 | 127.50 | |
| 95.22 | 77.40 | | | |

**Solution**

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

$H_1$: At least two of the means are not equal.

$\alpha = 0.05$

Critical region: $f > 2.836$, by interpolating in Table A.6.

Computations: $Y_{1.} = 1460.25$, $Y_{2.} = 440.36$, $Y_{3.} = 842.45$, $Y_{4.} = 707.41$, and $Y_{..} = 3450.47$. The analysis of variance is shown in the MINITAB output of Table 13.6.

Decision: Reject $H_0$ and conclude that the average serum alkaline phosphatase activity levels for the four drug groups are not all the same. The $P$-value is 0.02.

**TABLE 13.6** MINITAB Analysis of Table 13.5

```
MTB > Oneway C1 C2

One-Way Analysis of Variance

Analysis of Variance for C1
Source        DF        SS         MS          F         P
C2             3     13939       4646       3.57     0.022
Error         41     53376       1302
Total         44     67315

Level          N      Mean      StDev
1             20     73.01      25.75
2              9     48.93      47.11
3              9     93.61      46.57
4              7    101.06      30.76

Pooled StDev = 36.08
MTB >
```



# (MINITAB analysis of Table 13.1)

**Test for Equal Variances: Y versus Agg**

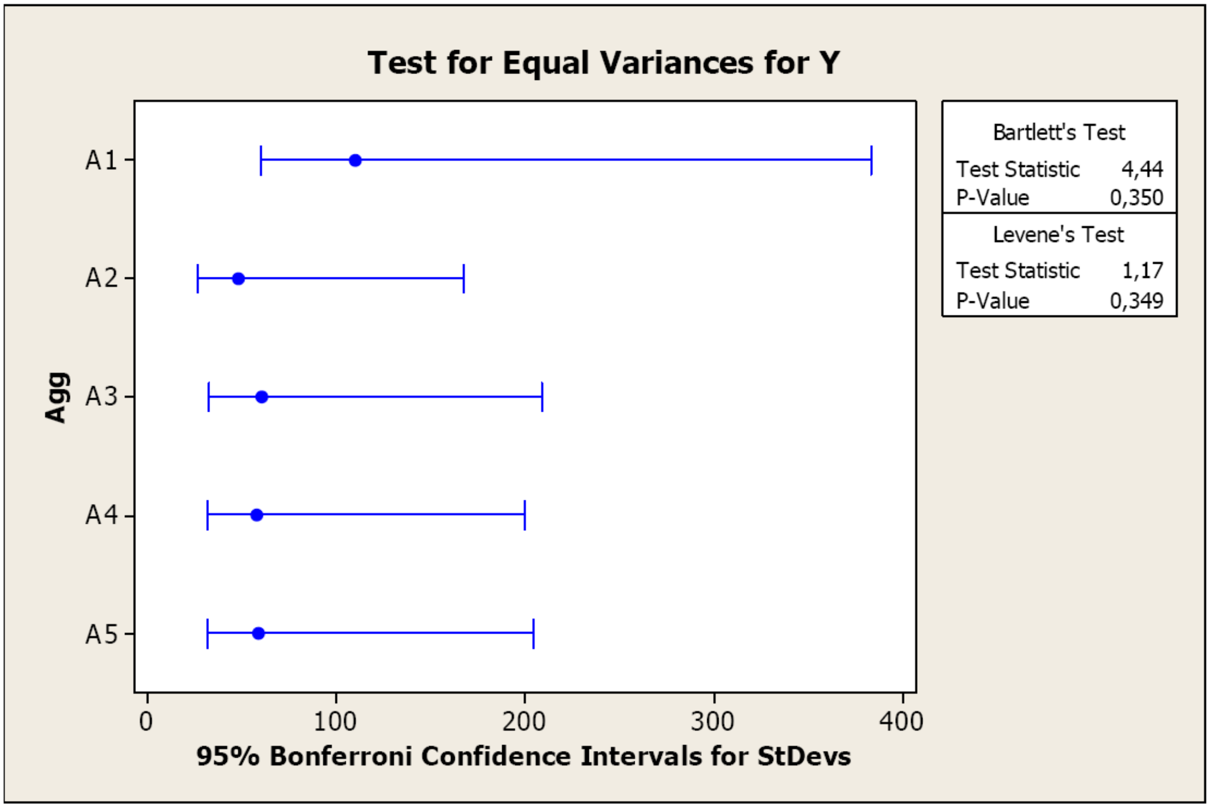
```
95% Bonferroni confidence intervals for standard deviations

Agg  N    Lower    StDev    Upper
 A1  6  60,1842  110,154  383,859
 A2  6  26,2179   47,986  167,220
 A3  6  32,7523   59,946  208,897
 A4  6  31,4744   57,607  200,746
 A5  6  32,1171   58,783  204,845


Bartlett's Test (normal distribution)
Test statistic = 4,44; p-value = 0,350


Levene's Test (any continuous distribution)
Test statistic = 1,17; p-value = 0,349
```

## Test for Equal Variances: Y versus Agg






### 1.3.5.7. Bartlett's Test (Used in MINITAB)

*Purpose: Test for Homogeneity of Variances*

Bartlett's test ( Snedecor and Cochran, 1983) is used to test if $k$ samples have equal variances. Equal variances across samples is called homogeneity of variances. Some statistical tests, for example the analysis of variance, assume that variances are equal across groups or samples. The Bartlett test can be used to verify that assumption.

Bartlett's test is sensitive to departures from normality. That is, if your samples come from non-normal distributions, then Bartlett's test may simply be testing for non-normality. The Levene test is an alternative to the Bartlett test that is less sensitive to departures from normality.

*Definition*

The Bartlett test is defined as:

$H_0$: $\sigma_1 = \sigma_2 = \ldots = \sigma_k$

$H_a$: $\sigma_i \neq \sigma_j$ for at least one pair $(i,j)$.

Test Statistic: The Bartlett test statistic is designed to test for equality of variances across groups against the alternative that variances are unequal for at least two groups.

$$T = \frac{(N-k)\ln s_p^2 - \sum_{i=1}^{k}(N_i - 1)\ln s_i^2}{1 + (1/(3(k-1)))((\sum_{i=1}^{k} 1/(N_i - 1)) - 1/(N-k))}$$

In the above, $s_i^2$ is the variance of the ith group, $N$ is the total sample size, $N_i$ is the sample size of the $i$th group, $k$ is the number of groups, and $s_p^2$ is the pooled variance. The pooled variance is a weighted average of the group variances and is defined as:

$$s_p^2 = \sum_{i=1}^{k}(N_i - 1)s_i^2/(N-k)$$

Significance Level: $\alpha$

Critical Region: The variances are judged to be unequal if,

$$T > \chi^2_{(\alpha,k-1)}$$

where $\chi^2_{(\alpha,k-1)}$ is the upper critical value of the chi-square distribution with $k$ - 1 degrees of freedom and a significance level of $\alpha$.