



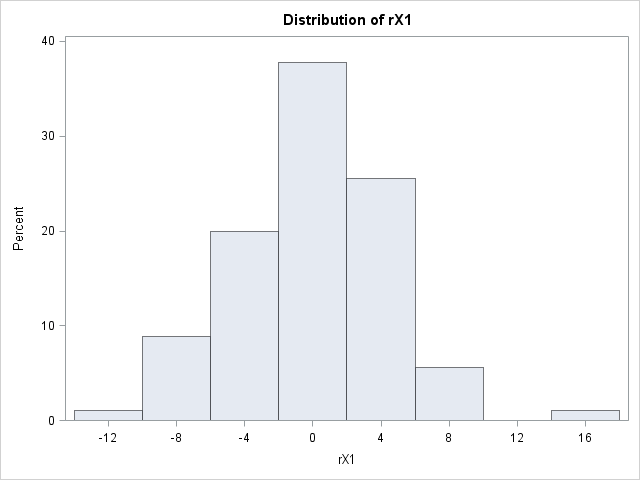
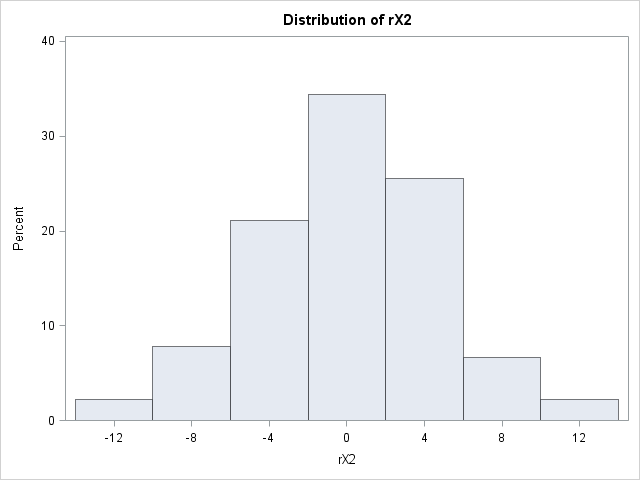
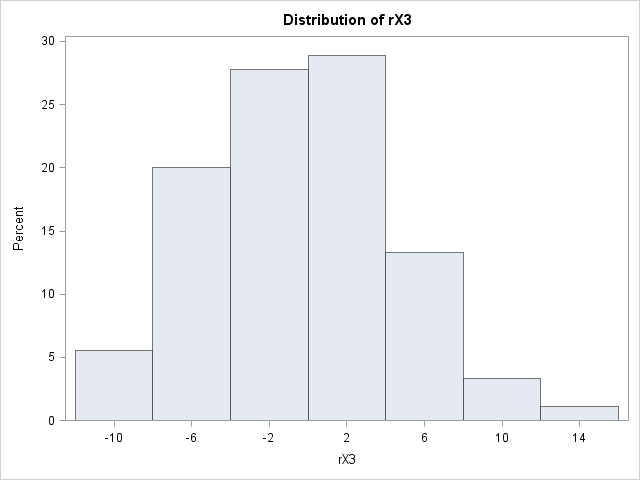
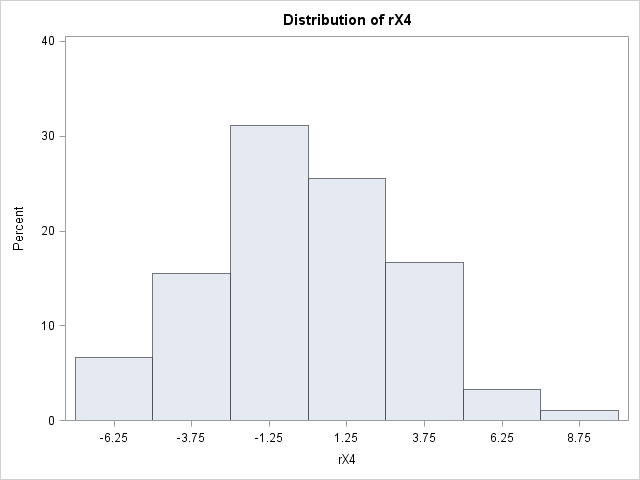
Assumption 1: The data from group i has common mean vector μi

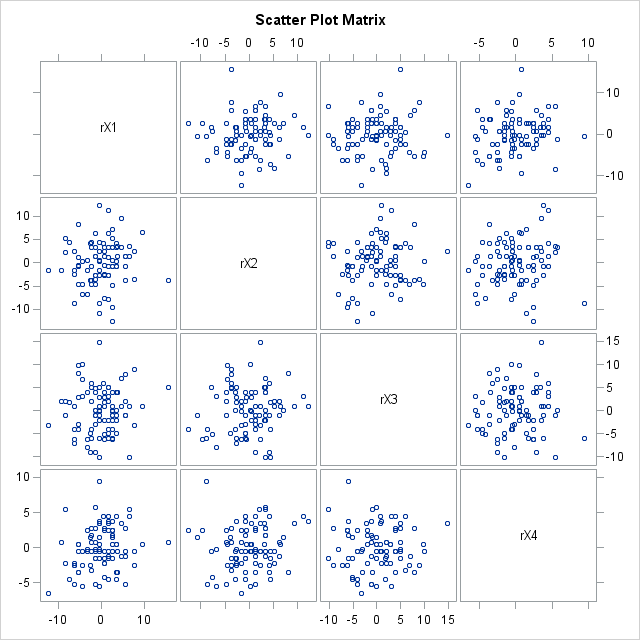
This assumption says that there are no subpopulations with different mean vectors. Here, this assumption might be violated if skulls collected from a given time period has inconsistency due to some reason.

Assumption 3: Independence: The subjects are independently sampled. Even though we see that the samples were selected over a period of time, time is a part of the hypothesis.

Assumption 4: Normality: The data are multivariate normally distributed.

We have from SAS:



Histograms suggest that the distributions are relatively symmetric.

Assumption 2: The data from all groups have common variance-covariance matrix Σ.

|  |
| --- |
| Bartlett's Test |

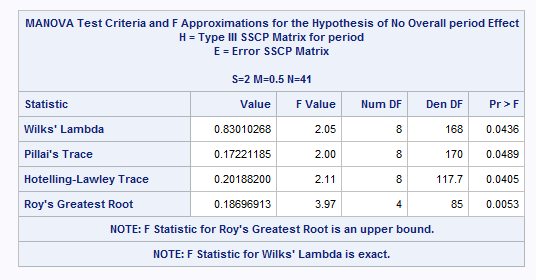
The DISCRIM Procedure

Test of Homogeneity of Within Covariance Matrices

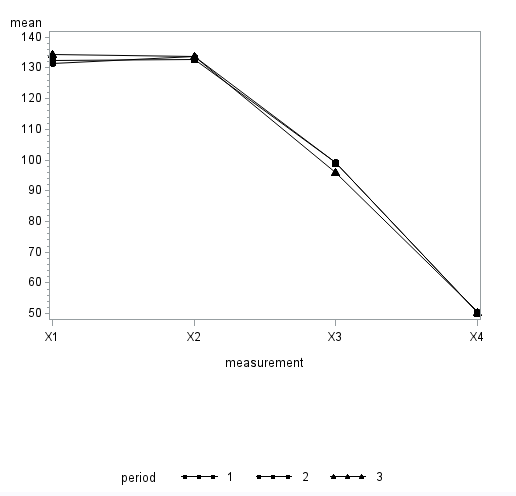
| **Chi-Square** | **DF** | **Pr > ChiSq** |
| --- | --- | --- |
| 21.048436 | 20 | 0.3943 |

We find no statistically significant evidence against the null hypothesis that the variance-covariance matrices are homogeneous (L' = 21.05; d.f. = 20; p = 0.39).

Manova output:



The measurements of the skull depends on the period from which the sample was obtained ( Λ\* = 0.8301; F = 2.05; d.f. = 8, 168; p = 0.0436 < 0.05). It was found therefore, that there are differences in the measurements of the skull of at least one dimension between at least one pair of periods.



|  |  |  |  |
| --- | --- | --- | --- |
| **Skull Dimension** | **F** | **df** | **SAS *p*-value** |
| **X1** | **3.66** | **2, 87** | **0.0298** |
| **X2** | **0.47** | **2, 87** | **0.6293** |
| **X3** | **3.84** | **2, 87** | **0.0251** |
| **X4** | **0.1** | **2, 87** | **0.9007** |

**Analysis of Individual Chemical Elements – Naïve approach**

F(2, 87, 0.05) = 3.1013

No we see that F-value is > Fcritical for X1 and X3 dimensions. Also for these variables p < 0.05 and therefore we see that for X1 and X3 we have significant results

**Analysis of Individual Chemical Elements – Bonferroni correction**

Here, p = 4 variables, g = 3 groups, and a total of N = 90 observations. So, for an α = 0.05 level test, we reject

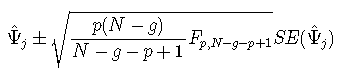
H0:μ1k=μ2k=⋯=μgk if F > F(2, 87, 0.05/4) = 4.61

Since none of the F-statistics exceed the critical value of 4.61, or equivalently, since the SAS p-values all fall above 0.0125 (= 0.05/4), we can see that none of the dimensions are significant at the 0.05 level under the Bonferroni correction.

**Conclusion**: Means for all dimensions don’t differ significantly among the periods.

We set up contrasts as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| **Contrast** | **Period 1** | **Period 2** | **Period 3** |
| 1 – 1 vs 2 | 1 | -1 | 0 |
| 2 – 1 vs 3 | 1 | 0 | -1 |
| 3 – 2 vs 3 | 0 | 1 | -1 |

We have 

Recall that we have p = 4 dimensions, g = 3 periods, and a total of N = 90 observations. From the F-table, we have F[4,90-3-4+1,0.05] = F[4,84,0.05] = 2.48. Then our multiplier is:

M = sqrt( ((p \* (N-g)) / (N-g-p+1)) \* F[p, N-g-p+1]) = sqrt( ((4 \* (90-3)) / (90-3-4+1)) \* 2.48) = 3.205

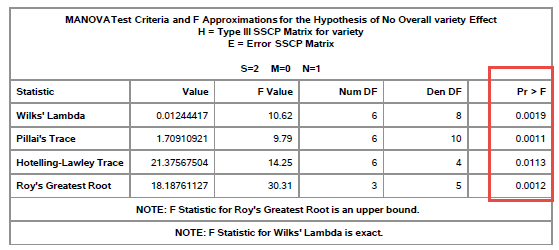
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | M | SQRT( ((4 \* (90-3)) / (90-3-4+1)) \* 2.48) | 3.205353 |  |
| **Contrast 1** | 1 vs 2 |  |  |  |  |
|  |  |  |  |  |  |
|  | **Ψ^** | **SE** | **M X SE** | **Lower CI** | **Upper CI** |
| **X1** | -1 | 1.17 | 3.750262619 | -4.75026 | 2.75026262 |
| **X2** | 0.9 | 1.214 | 3.891298137 | -2.9913 | 4.79129814 |
| **X3** | 0.1 | 1.284 | 4.115672823 | -4.31567 | 3.91567282 |
| **X4** | 0.3 | 0.802 | 2.570692838 | -2.27069 | 2.87069284 |
| **Contrast 2** | 1 vs 3 |  |  |  |  |
|  |  |  |  |  |  |
|  | **Ψ^** | **SE** | **M X SE** | **Lower CI** | **Upper CI** |
| **X1** | -3.1 | 1.17 | 3.750262619 | -6.85026 | 0.65026262 |
| **X2** | -0.2 | 1.214 | 3.891298137 | -4.0913 | 3.69129814 |
| **X3** | 3.133 | 1.284 | 4.115672823 | -0.98267 | 7.24867282 |
| **X4** | -0.033 | 0.802 | 2.570692838 | -2.60369 | 2.53769284 |
| **Contrast 3** | 2 vs 3 |  |  |  |  |
|  |  |  |  |  |  |
|  | **Ψ^** | **SE** | **M X SE** | **Lower CI** | **Upper CI** |
| **X1** | -2.1 | 1.17 | 3.750262619 | -5.85026 | 1.65026262 |
| **X2** | -1.1 | 1.214 | 3.891298137 | -4.9913 | 2.79129814 |
| **X3** | 3.033 | 1.284 | 4.115672823 | -1.08267 | 7.14867282 |
| **X4** | -0.333 | 0.802 | 2.570692838 | -2.90369 | 2.23769284 |



1. At alpha = 0.05 (no Bonferrorni correction) - The dependant variables depends on the location of the seeds ( Λ\* = 0.10651620; F = 11.18; d.f. = 3, 4; p = 0.0205 < 0.05). We reject the null hypothesis.

At alpha = 0.05 after Bonferrorni correction = 0.05/3 = 0.0167 - The dependant variables don’t depends on the location of the seeds ( Λ\* = 0.10651620; F = 11.18; d.f. = 3, 4; p = 0.0205 > 0.0167). We fail to reject the null hypothesis.

1. The p-values are:

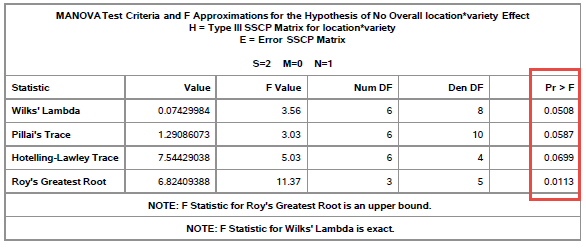


Making conclusions based on Wilks Lambda:

At alpha = 0.05 (no Bonferrorni correction) - The dependant variable means depends on the variety of the seeds ( Λ\* = 0.01244417; F = 10.62; d.f. = 6, 8; p = 0.0019 < 0.05). We reject the null hypothesis.

At alpha = 0.05 after Bonferrorni correction = 0.05/3 = 0.0167 - The dependant variable means depends on the variety of the seeds ( Λ\* = 0.01244417; F = 10.62; d.f. = 6, 8; p = 0.0019 < 0.0167). We reject the null hypothesis.

1. The p-values are:



Making conclusions based on Wilks Lambda:

Here we have H0: there is no interaction and HA: an interaction exists

At alpha = 0.05, we see that the interaction term is not significant. Since the interaction term is not significant we can interpret the individual factors as we did in parts a and b

1. ANOVA for the 3 variables gives us:

|  |  |  |  |
| --- | --- | --- | --- |
| **Dimension** | **F** | **df** | **SAS *p*-value** |
| **X1** | **4.63** | **5, 6** | **0.0446** |
| **X2** | **6.92** | **5, 6** | **0.0177** |
| **X3** | **5.6** | **5, 6** | **0.0292** |

**Analysis of individual variables – Naïve approach**

F(5, 6, 0.05) = 4.387

Now we see that F-value is > Fcritical for all dimensions. Also for these variables p < 0.05 and therefore we see that all X1, X2 and X3 we have significant results

**Analysis of individual variables – Bonferroni correction**

Here, p = 3 variables. So, for an α = 0.05 level test, we reject

H0:μ1k=μ2k=⋯=μgk if F > F(5, 6, 0.05/3) = 7.10

Now we see that F-value is < Fcritical for all dimensions. Therefore we see that all X1, X2 and X3 we don’t have significant results

1. For contrast “diff56”

Lets make the table for all the contrasts

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Contrast | https://onlinecourses.science.psu.edu/stat505/sites/onlinecourses.science.psu.edu.stat505/files/lesson13/formula_80.gif | F | Num d.f. | Denom d.f. | p |
| diff56 | 0.069559 | 17.83 | 3 | 4 | 0.0089 |
| diff58 | 0.23363 | 4.37 | 3 | 4 | 0.094 |
| diff68 | 0.06095 | 20.54 | 3 | 4 | 0.0068 |

The mean dimension of seeds from variety 5 differs in at least one variable from that of variety 6 (https://onlinecourses.science.psu.edu/stat505/sites/onlinecourses.science.psu.edu.stat505/files/lesson13/formula_80.gif = 0.069559; F = 17.83; d.f. = 3, 4; p = 0.0089 < 0.05). We reject the null.

1. Refer table above - There is no significant difference in the mean dimension of seeds from variety 5 and 8 (https://onlinecourses.science.psu.edu/stat505/sites/onlinecourses.science.psu.edu.stat505/files/lesson13/formula_80.gif = 0.23363; F = 4.37; d.f. = 3, 4; p = 0.094 > 0.05). We fail to reject the null.
2. Refer table above - The mean dimension of seeds from variety 6 differs in at least one variable from that of variety 8 (https://onlinecourses.science.psu.edu/stat505/sites/onlinecourses.science.psu.edu.stat505/files/lesson13/formula_80.gif = 0.06095; F = 20.54; d.f. = 3, 4; p = 0.0068 < 0.05). We reject the null.
3. Here we will use the following variables:

Variable: X1



Variable: X2



Variable: X3



Based on the above tables we see that for contrast diff56, we have significant differences for variable X1 (p=0.0154) and X2 (p=0.0186)

1. Based on the above tables we see that for contrast diff58, we have significant differences for variable X2 (p=0.0064) and X3 (p=0.0082)
2. Based on the above tables we see that for contrast diff68, we have significant differences for variable X3 (p=0.0141)
3. The null-hypothesis of this test is that the population is normally distributed. Thus if the p-value is less than the chosen alpha level, then the null hypothesis is rejected and there is evidence that the data tested are not from a normally distributed population. In other words, the data are not normal.

Residuals: res1



We have p-value > alpha = 0.05, we fail to reject the null hypothesis. We can conclude that the data came from a normal distribution.

Residuals: res2



We have p-value > alpha = 0.05, we fail to reject the null hypothesis. We can conclude that the data came from a normal distribution.

Residuals: res3



We have p-value > alpha = 0.05, we fail to reject the null hypothesis. We can conclude that the data came from a normal distribution.