

C.6 Analysis of Covariance

Uric acid is measured in three groups of subjects, namely, normal (group-1), diabetic (group-2) and cardiovascular disease (group-3). Sample size is 30 in each group. The purpose is to find if three groups have same mean level of uric acid. However, age pattern was different in the three groups. It is suspected that age would influence the uric acid level. Analysis of covariance (ANCOVA) is the method that would eliminate the linear effect of age when age is considered as a covariate.

First confirm that uric acid and age are indeed correlated. If not correlated, there is no need to run ANCOVA.

Command for Pearson correlation:

```
CORRELATIONS VARIABLES = uricacid age
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

This command produces the Pearson correlation with two-tailed P-value after excluding the pair value if either uric acid or age is missing.

Correlations

		uric acid	age
uric acid	Pearson Correlation	1	.522(**)
	Sig. (2-tailed)		.000
	N	90	90
age	Pearson Correlation	.522(**)	1
	Sig. (2-tailed)	.000	
	N	90	90

** Correlation is significant at the 0.01 level (2-tailed).

Correlation between level of uric acid and age is 0.522 ($P < 0.001$). This is statistically significant (cannot be considered zero). Thus it is in order to run ANCOVA and adjust for effect of age. However, ANCOVA is valid only if (i) error variance of uric acid across groups is nearly equal (homoscedasticity), and (ii) the gradient of relationship between uric acid and age is the same across the groups.

To check the above two assumptions, run the following ANCOVA command:

```
UNIANOVA uricacid BY group WITH age
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PRINT = HOMOGENEITY
/CRITERIA = ALPHA(.05)
/DESIGN = group age age*group.
```

The DESIGN command includes age*group interaction also to test equality of gradients.

Levene Test of Equality of Error Variances(a)

Dependent Variable: uric acid

F	df1	df2	Sig.
3.003	2	87	.055

a Design: Intercept+group+age+group*age

The above table shows the result of Levene test for equality of error variance using F -test. Since P -value >0.05 , evidence is not enough to reject the null hypothesis. We are not wrong in assuming homoscedasticity. The following table shows that the interaction term between the age and group is not significant ($P = 0.961$). Thus there is no evidence of violation of the equal gradient assumption.

Tests of Between-Subjects Effects

Dependent Variable: uric acid

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	53.669(a)	5	10.734	12.439	.000
Intercept	95.862	1	95.862	111.09	.000
group	1.673	2	.836	.969	.384
age	8.461	1	8.461	9.805	.002
group * age	0.069	2	.035	.040	.961
Error	72.486	84	0.863		
Total	2344.666	90			
Corrected Total	126.155	89			

a. R Squared = .425 (Adjusted R Squared = .391)

Equal gradient can be checked graphically also as follows (Figure C.6).

```
GRAPH
  /SCATTERPLOT(BIVAR)=age WITH uricacid BY group
  /MISSING=LISTWISE.
```

[Note: For drawing the lines in the following figure, we have used the chart editor in SPSS output]

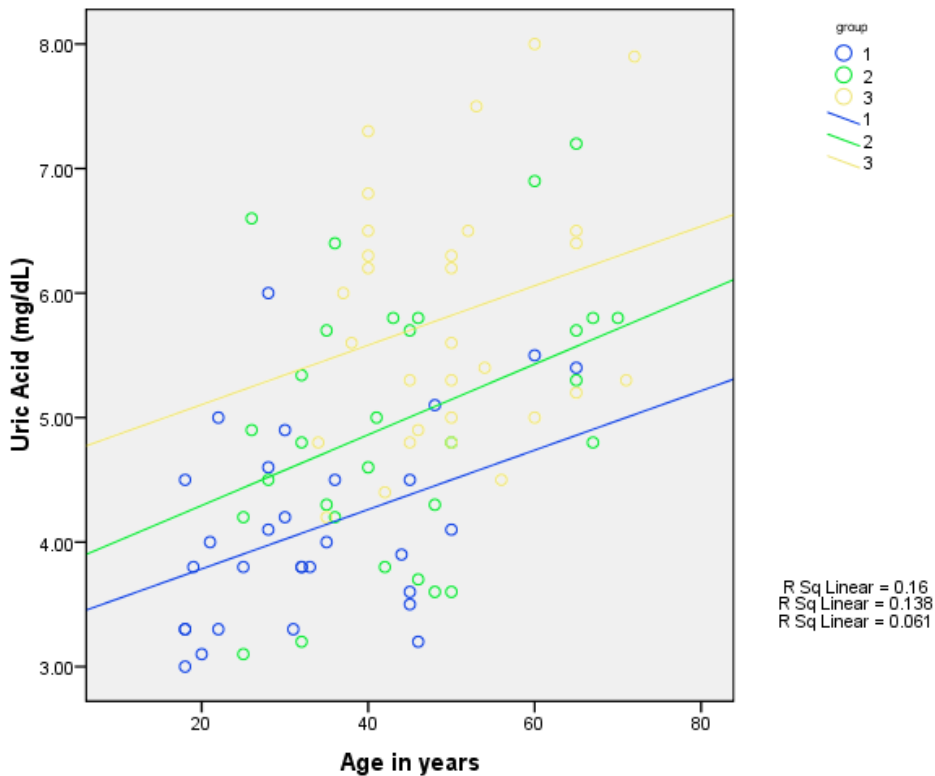


FIGURE C.6 Relationship between uric acid and age in the three groups

Assumptions of analysis of covariance are fulfilling and we can run ANCOVA.

Command to run analysis of covariance:

```
UNIANOVA uricacid BY group WITH age /METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE /EMMEANS = TABLES(group) WITH(age=MEAN) COMPARE ADJ(BONFERRONI)
/PRINT = DESCRIPTIVE OPOWER PARAMETER HOMOGENEITY /CRITERIA = ALPHA(.05)
/DESIGN = group age.
```

The DESIGN command now excludes interaction as it is not significant. The following output gives the group means and SDs.

Descriptive Statistics

Dependent Variable: uric acid

group	Mean	Std. Deviation	N
1	4.097	.769	30
2	4.981	1.081	30
3	5.817	1.029	30
Total	4.965	1.191	90

Some of these differences across groups could be due to differences in age structure of the subjects in the three groups. The following tests the null hypothesis that the error variance of the dependent variable is equal across groups, now for ANCOVA model.

Levene's Test of Equality of Error Variances(a)

Dependent Variable: uric acid

F	df1	df2	Sig.
3.105	2	87	.050

a Design: Intercept+group+age

Note: The results of Levene's test is different in SPSS and R due to design consideration. R used `uricacid~group` whereas in SPSS `uricacid~age + group`.

Since P -value of Levene test is = 0.05, which is marginal, we assume for simplicity that the variances are not significantly different. Under this assumption, we get the following output for ANCOVA model.

Tests of Between-Subjects Effects

Dependent Variable: uric acid

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power(a)
Corrected Model	53.600(b)	3	19.960	21.178	.000	63.533	1.000
Intercept	107.089	1	100.193	126.934	.000	126.934	1.000
group	19.289	2	9.644	11.431	.000	22.863	.992
age	9.212	1	9.212	10.919	.001	10.919	.904
Error	72.555	86	0.844				
Total	2344.666	90					
Corrected Total	126.155	89					

a Computed using alpha = .05

b R Squared = .425 (Adjusted R Squared = .405)

Effect of both age and group are significant, indicating that mean uric acid is different in different groups even after adjusting for age differentials, and age effect by itself is also significant. However R^2 (= 0.425) is not high and the model is not sufficiently good for prediction.

Parameter Estimates

Dependent Variable:uricacid

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	3.246	.307	10.561	.000	2.635	3.857
age	.026	.008	3.304	.001	.010	.041
[group=3.00]	1.288	.271	4.759	.000	.750	1.827
[group=2.00]	.598	.253	2.369	.020	.096	1.100
[group=1.00]	0 ^a

a. This parameter is set to zero because it is redundant.

The above table shows regression parameter estimate of uric acid on age. This is $B = 0.026$. B-coefficients for groups are used to estimate mean uric acid in the three groups after removing the effect of age. This is calculated at mean age as shown below. Reference group in this case is group = 1 (Normal).

Estimates

Dependent Variable:uricacid

group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	4.336 ^a	.183	3.973	4.699
2	4.934 ^a	.168	4.600	5.269
3	5.624 ^a	.177	5.272	5.977

a. Covariates appearing in the model are evaluated at the following values: age = 42.37.

At the average age = 42.37 years, mean uric acid is (5.624 – 4.934 =) 0.690 mg/dL higher in cardiovascular disease group than diabetic group and 1.288 mg/dL higher than control group. Mean of which group is significantly different from the others is found by the following pairwise comparisons using Bonferroni adjustment.

Pairwise Comparisons

Dependent Variable:uricacid

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
1	2	-.598	.253	.060	-1.215	.018
	3	-1.288*	.271	.000	-1.950	-.627
2	1	.598	.253	.060	-.018	1.215
	3	-.690*	.241	.016	-1.279	-.101
3	1	1.288*	.271	.000	.627	1.950
	2	.690*	.241	.016	.101	1.279

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

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

The results show that there is a significant difference in diabetic and cardiovascular groups compared with normal after the linear effect of age is eliminated. However, there no significant difference between the diabetic and cardiovascular subjects.