Analysis of Nested Repeated Measures Model with applications

Dr. Abdul - Hussein S . Al - Mouel and Ban M . Batto
Dept. of Mathematics
College of Education
University
of Basrah
Basrah – Iraq

Abstract

In this paper we study the balanced 4 – way mixed effects model that has one random effect interacting with one fixed effects and this interaction interacts with each of other two fixed effects as a nested repeated measures model .

As studies of water in the present time of the basic necessities, which must be carried out, the fact that the quality of water has become specific agricultural production, both quantitatively and qualitatively, in addition, the use of the kinds of poor water is reflected not only damage to crop productivity but may lie in the land itself so that result, over the years to a deterioration in the overall qualities good soil and the emergence of the problems of salinization and alkalinity and toxin crops planted.

From here we discussed the application of data on agricultural the experience of the study as well as types or levels of different fertilization to reduce existing salinity in the waters of subterranean wells for three sites in the Zubayr region and the impact on the production rate recipes growth(the level of higher plant) a crop also targeted the experience in the same time creating overlap between the quality of water used, as well as chemical fertilization and their impact on production quantity and quality through the application of the model of repeated measurements.

1. Introduction :

Many situations arise when a particular individual (person , field , animal , etc.) receives several treatments at various levels and statistician is asked to analysis the effects of these treatments . In such situations observations can not be assumed to be independent of each other . such design is called a repeated measures design (RMM) . Now a days this design is one of the most widely used experimental designs in biological , agricultural , educational . Hence a lot of statistical and experimental work have been done in this filed .

Hypotheses tests are interest in these situations generally fall into one of two categories : tests about the means and tests about the covariance structure . In this work our primary concern is in testing hypotheses about the means , but we also derive sensible tests for many hypotheses about the covariance structure . We often assume that we have sequence of independent p - vectors of measurements , Y_1 ,…., Y_n , where Y_i is taken to have a multivariate normal distribution with mean μ_i and covariance matrix $\sum$. In the most general case when $\sum$ is taken to be arbitrary we have to determine p(p+1) / 2 parameters , so in order to get more powerful procedures , we must reduce the number of parameter in $\sum$ to be estimated . For this reason , the covariance matrix is often assumed to have a known structure.

In section 2 we set up the model and write ANOVA table where as in section 3 we derive the estimators of the parameters as a nested repeated measures model of Al . Mouel , in section 4 we give an applications of this study .

2. Setting up the model :

We assume throughout this paper that each individual has the same number , m , of sub individuals and each sub individual receives the same number , rc , of treatments . Let Y_{ijkl} be the $(k,l)^{th}$ observations on the j^{th} sub individual from the i^{th} individual

$(1 \le i \le n, 1 \le j \le m, 1 \le k \le r, 1 \le l \le c)$ and let :

$$\boldsymbol{Y}_{ijk} = \begin{bmatrix} \boldsymbol{Y}_{ijk1} \\ \vdots \\ \boldsymbol{Y}_{ijkc} \end{bmatrix}, \boldsymbol{Y}_{ij} = \begin{bmatrix} \boldsymbol{Y}_{ij1} \\ \vdots \\ \boldsymbol{Y}_{ijr} \end{bmatrix}, \boldsymbol{Y}_{i} = \begin{bmatrix} \boldsymbol{Y}_{i1} \\ \vdots \\ \boldsymbol{Y}_{im} \end{bmatrix}, \boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_{1} \\ \vdots \\ \boldsymbol{Y}_{n} \end{bmatrix}$$

Then Y_{ijk} is the vector of observations on the j^{th} sub individual from the i^{th} individual when the row treatment fixed at level k , Y_{ij} is the vector of observations of the j^{th} sub individual from i^{th} individual , and Y_i is the vector of observations on sub individuals of the i^{th} individual.

Let $E(Y_{ijkl}) = \mu_{ijkl}$, $\mu ijk = E(Y_{ijk}), \mu_{ij} = E(Y_{ij})$ and $\mu_i = E(Y_i)$. It assumed that the Y_i are independently normally distributed with mean μ_i and covariance matrix $\sum$.

Now we study the balanced four – way mixed effects model that has one random effect interacting with one fixed effects and this interaction interacts with each of the other two fixed effects .

We consider the model in which we observe

$$Y_{ijkl} = \theta + \alpha_i + \beta_j + \gamma_k + \xi_l + (\alpha\beta)_{ij} + (\alpha\beta\gamma)_{ijk} + (\alpha\beta\xi)_{ijl} + e_{ijkl}$$

$i = 1,.....,a$, $j = 1,.....,b$, $k = 1,.....,c$, $l = 1,.....,d$, where

θ , β_j , γ_k , ξ_l are unknown parameters such that :

$\sum_j \beta_j = 0, \sum_k \gamma_k = 0, \sum_l \xi_l = 0$, *the* $\alpha_i, (\alpha\beta)_{ij}, (\alpha\beta\gamma)_{ijk}, (\alpha\beta\xi)_{ijl}, e_{ijkl}$ are unobserved independent random variables such that:

$$\alpha_i \sim N(0,\sigma_\alpha^2), (\alpha\beta)_{ij} \sim N(0,\sigma_{\alpha\beta}^2), (\alpha\beta\gamma)_{ijk} \sim N(0,\sigma_{\alpha\beta\gamma}^2), (\alpha\beta\xi)_{ijl} \sim N(0,\sigma_{\alpha\beta\xi}^2)$$

and $e_{ijkl} \sim N(0,\sigma_e^2)$.

The parameter space for this model is given by $-\infty < \theta < \infty$,

$$\sum_j \beta_j = 0, \sum_k \gamma_k = 0, \sum_l \xi_l = 0, \sigma_\alpha^2 \geq 0, \sigma_{\alpha\beta}^2 \geq 0, \sigma_{\alpha\beta\gamma}^2 \geq 0, \sigma_{\alpha\beta\xi}^2 \geq 0, \sigma_e^2 \geq 0$$

We write the analysis of variance table :

Model Term	Sources	Sum of squares	Degrees of freedom
α_i	A	$bcd\sum_i (\bar{Y}_{i...} - \bar{Y}_{....})^2$	a-1
β_j	B	$acd\sum_j (\bar{Y}_{.j..} - \bar{Y}_{....})^2$	b-1
γ_k	C	$abd\sum_k (\bar{Y}_{..k.} - \bar{Y}_{....})^2$	c-1
ξ_l	D	$abc\sum_l (\bar{Y}_{...l} - \bar{Y}_{....})^2$	d-1
$(\alpha\beta)_{ij}$	AB	$cd\sum_i\sum_j (\bar{Y}_{ij..} - \bar{Y}_{i...} - \bar{Y}_{.j..} + \bar{Y}_{....})^2$	(a-1)(b-1)
$(\alpha\beta\gamma)_{ijk}$	ABC	$d\sum_i\sum_j\sum_k (\bar{Y}_{ijk.} - \bar{Y}_{ij..} - \bar{Y}_{..k.} + \bar{Y}_{....})^2$	(ab-1)(c-1)
$(\alpha\beta\xi)_{ijl}$	ABD	$c\sum_i\sum_j\sum_l (\bar{Y}_{ij.l} - \bar{Y}_{ij..} - \bar{Y}_{...l} + \bar{Y}_{....})^2$	(ab-1)(d-1)
e_{ijkl}	E	$\sum_i\sum_j\sum_k\sum_l (\bar{Y}_{ijkl} - \bar{Y}_{iijk.} - \bar{Y}_{ij.l} + \bar{Y}_{....})^2$	ab(c-1)(d-1)

3. Estimators of Parameters in the model :

In this section we find the estimators of the unknown parameters in 4- way mixed effects model .

$$\mu_{ijkl} = E(Y_{ijkl}) = \theta + \beta_j + \gamma_k + \xi_l$$

Since $\beta_1^{(i)} = \bar{\mu}_{i...}$ from [2]

$$\therefore \mu_{i...} = \sum_j \sum_k \sum_l (\mu_{ijkl}) = \sum_j \sum_k \sum_l (\theta + \beta_j + \gamma_k + \xi_l) = bcd\theta$$

$$\therefore \bar{\mu}_{i...} = \theta \qquad \therefore \quad \beta_1^{(i)} = \theta \qquad (1)$$

Since $\beta_2^{(ij)} = \bar{\mu}_{ij..} - \bar{\mu}_{i...}$ from [2]

$$\therefore \mu_{ij..} = \sum_k \sum_l (\mu_{ijkl}) = \sum_k \sum_l (\theta + \beta_j + \gamma_k + \xi_l) = cd\theta + cd\beta_j$$

$$\therefore \bar{\mu}_{ij..} = \theta + \beta_j \qquad \therefore \quad B_2^{(ij)} = \theta + \beta_j - \theta = B_j$$

$$\therefore \quad \beta_2^{(ij)} = \beta_j \qquad (2)$$

Since $\beta_3^{(ijk)} = \bar{\mu}_{ijk.} - \bar{\mu}_{ij..}$ from [2]

$$\therefore \mu_{ijk.} = \sum_l (\mu_{ijkl}) = \sum_l (\theta + \beta_j + \gamma_k + \xi_l) = d\theta + d\beta_j + d\gamma_k$$

$$\therefore \bar{\mu}_{ijk.} = \theta + \beta_j + \gamma_k$$

$$\therefore \beta_3^{(ijk)} = \theta + \beta_j + \gamma_k - \beta_j - \theta = \theta + \gamma_k - \theta = \gamma_k$$

$$\therefore \beta_3^{(ijk)} = \gamma_k \qquad (3)$$

Since $\beta_4^{(ijl)} = \bar{\mu}_{ij.l} - \bar{\mu}_{ij..}$ from [2]

$$\therefore \mu_{ij.l} = \sum_k (\mu_{ijkl}) = \sum_k (\theta + \beta_j + \gamma_k + \xi_l$$

$$\therefore \bar{\mu}_{ij.l} = \theta + \beta_j + \xi_l$$

$$\therefore \beta_4^{(ijl)} = \xi_l \qquad (4)$$

Now let

$$Q=\sum_i\sum_j\sum_k\sum_l (Y_{ijkl}-\mu_{ijkl})^2=\sum_i\sum_j\sum_k\sum_l (Y_{ijkl}-\theta-\beta_j-\gamma_k-\xi_l)^2$$

$$\frac{\partial Q}{\partial\theta}\Big|_{\theta=\hat{\theta}}=0 \qquad \Rightarrow Y_{....}=\hat{\theta}$$

From (1) we get $\hat{\beta}_1^{(i)}=\bar{Y}_{....}$

$$\frac{\partial Q}{\partial\beta_j}\Big|_{\beta_j=\hat{\beta}_j}=0 \qquad \Rightarrow \hat{\beta}_j=\bar{Y}_{.j..}-\bar{Y}_{....}$$

From (2) we get $\hat{\beta}_2^{(ij)}=\bar{Y}_{.j..}-\bar{Y}_{....}$

$$\frac{\partial Q}{\partial\gamma_k}\Big|_{\gamma_k=\hat{\gamma}_k}=0 \qquad \Rightarrow \hat{\gamma}_k=\bar{Y}_{..k.}-\bar{Y}_{....}$$

From (3) we get $\hat{\beta}_3^{(ijk)}=\bar{Y}_{..k.}-\bar{Y}_{....}$

$$\frac{\partial Q}{\partial\xi_l}\Big|_{\xi_l=\hat{\xi}_l}=0 \qquad \Rightarrow \hat{\xi}_l=\bar{Y}_{..l}-\bar{Y}_{....}$$

From (4) we get $\hat{\beta}_4^{(ijl)}=\bar{Y}_{..l}-\bar{Y}_{....}$

We note that Y_{ijkl} and $Y_{i'j'k'l'}$ are not independent for the mixed effects model and we are interested in testing that $\beta_j=0$, $\gamma_k=0$, $\xi_l=0$ (i.e. no fixed main effects) , that $\sigma_\alpha^2=0$ (i.e. no random main effects) and that the $\sigma_{\alpha\beta}^2=0$, $\sigma_{\alpha\beta\gamma}^2=0$, $\sigma_{\alpha\beta\xi}^2=0$ (i.e. no interaction effects)

$$Cov(Y_{ijkl}, Y_{i'j'k'l'}) = \begin{cases} \sigma^2 & if\ i = i', j = j', k = k', l = l' \\ \sigma^2_{\rho_4} & if\ i = i', j = j', k = k', l \neq l' \\ \sigma^2_{\rho_3} & if\ i = i', j = j', k \neq k', l = l' \\ \sigma^2_{\rho_2} & if\ i = i', j = j', k \neq k', l \neq l' \\ \sigma^2_{\rho_1} & if\ i = i', j \neq j' \\ 0 & if\ i \neq i' \end{cases}$$

Where $$\sigma^2 = \sigma^2_{\alpha} + \sigma^2_{\alpha\beta} + \sigma^2_{\alpha\beta\gamma} + \sigma^2_{\alpha\beta\xi} + \sigma^2_{e}$$,

$$\rho_4 = \frac{\sigma^2_{\alpha} + \sigma^2_{\alpha\beta} + \sigma^2_{\alpha\beta\gamma}}{\sigma^2} \ , \ \rho_3 = \frac{\sigma^2_{\alpha} + \sigma^2_{\alpha\beta} + \sigma^2_{\alpha\beta\xi}}{\sigma^2} \ , \ \rho_2 = \frac{\sigma^2_{\alpha} + \sigma^2_{\alpha\beta}}{\sigma^2} \ , \ \rho_1 = \frac{\sigma^2_{\alpha}}{\sigma^2}$$

4. Applications :

A synopsis of sample Research:

data taken from a researcher at the Faculty of Agriculture and dealt with the study of groundwater quality in the Zubayr region and their suitability for irrigation under levels of various fertilization was the choice of three wells by the representative of the quality of water prevailing in the region for the implementation of the experience of farming in the three sites to show the impact of the quality of water as well as overlap between the quality of water, fertilization , the characteristics of growth and production a crop.

Data:

The following table shows the impact of three sites of underground water levels and five fertilization rate on crop production a recipes growth (the average rise plant) to him.

third site	**second site**	**first site**	**The sites of wells** **fertilization levels**
28.03 85.08	32.19 87.49	46.80 91.46	1 1
32.29 88.36	37.04 94.00	61.65 100.05	2 2
34.84 91.92	40.67 99.58	71.86 131.26	3 3
35.69 94.56	43.70 102.22	74.05 141.92	4 4
38.61 100.50	44.78 107.80	75.52 151.97	5 5

The statistical analysis:

For the purpose of testing and on scratch and alternative:

$$H_o : \beta_1 = \beta_2 = \beta_3$$

$$H_1 : \beta_1 \neq \beta_2 \neq \beta_3$$

or in other words, is there the impact of the quality of groundwater, compared with levels of fertilization, the rate of production quality and attributes the rise in plant) crop according to a system of SPSS in computer we got what follows:

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
STATE	Sphericity Assumed	5740.844	2	2870.422	107.745	.000
	Greenhouse-Geisser	5740.844	1.015	5655.278	107.745	.000
	Huynh-Feldt	5740.844	2.000	2870.422	107.745	.000
	Lower-bound	5740.844	1.000	5740.844	107.745	.000
STATE * LEVEL	Sphericity Assumed	927.438	8	115.930	4.352	.017
	Greenhouse-Geisser	927.438	4.061	228.404	4.352	.068
	Huynh-Feldt	927.438	8.000	115.930	4.352	.017
	Lower-bound	927.438	4.000	231.859	4.352	.069
Error(STATE)	Sphericity Assumed	266.410	10	26.641		
	Greenhouse-Geisser	266.410	5.076	52.488		
	Huynh-Feldt	266.410	10.000	26.641		
	Lower-bound	266.410	5.000	53.282		

The table above, we find that the value of F= 107.745 and private sites and the level of moral Sig. =0.000 indicating that there are differences between moral sites wells also said that the value of F=4.352 special overlap between sites and levels of fertilization and the level of moral Sig. =0.017 any differences there are significant moral any impact of the levels of fertilization on the quality of groundwater.

Table:

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	STATE	Type III Sum of Squares	df	Mean Square	F	Sig.
STATE	Linear	5077.720	1	5077.720	124.226	.000
	Quadratic	663.125	1	663.125	53.446	.001
STATE * LEVEL	Linear	775.808	4	193.952	4.745	.059
	Quadratic	151.629	4	37.907	3.055	.126
Error(STATE)	Linear	204.374	5	40.875		
	Quadratic	62.036	5	12.407		

Special analog comparisons within the groups it is better to use the linear relationship the variable (on the location of the well, as well as levels of fertilization) because the level of moral was Sig. = 0.000 , Sig. = 0.059 it represents the best.

As for the such table :

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	171439.2	1	171439.2	33.708	.002
LEVEL	2360.290	4	590.073	.116	.971
Error	25429.917	5	5085.983		

which showed tests as a result of influences between groups exacting the acceptance of the hypothesis that there are no differences with significance between statistical production rate growth recipes a crop, where the level of significance was Sig. = 0.971, the largest of the level of the dependent study (0.05).

The discussion of the results:

There is an increase in the rate of production and high intentions down salty water irrigation increased levels of fertilization, where it was stated that under the circumstances of irrigation salt water (water subterranean wells (to be a factor Salinity is the determining factor in the rate of production and growth can be attributed this influence to that plant because of high **alasmosey** pressure, an increase of salty water irrigation spend additional capacity to absorb water from intravenous saline, which may benefit plant in the building of its cells new tissues and thereby increase its production.

References:

1. Al-Hiello, A.A. (1987), Groundwater Quality In Zubair Area and It's Suitability For Irrigation Under Different Fertilizer Levels, A Thesis Submitted to college of Agriculture university of Basrah.
2. Al-Mouel , A.S. (2004) " Nested Reported Measures Model Sufficiency And Estimation ", Chinese Journal of Applied probability and statistics, Vol. 20, No.2, 133-146 .
3. Al-Mouel , A.S and Wany, J.L. (2004) " one – way Multivariate Repeated Measurements analysis of Variance Model ". Applied Mathematics of Chinese Universities . Vol. 19, No.4, 435-448.
4. Al-Mouel , A.S. and Jassim, J.M. (2006) " Two – Way Multivariate Repeated Measurements analysis of variance Model " . J. Basrah Researches (science) Vol. 32, Part 1, 17-31.
5. Anderson, T.W. (1984) " An Introduction to Multivariate Statistical Analysis " Second Edition, New York, John Wiley.
6. Arnold, S.F. (1979) " Liner Models with exchangebly distributed errors " Journal of the American statistical Association 74 , 194-199.
7. Arnold, S.F. (1979) "A Coordinate – free approach to finding optimal procedures for repeated measures designs " Annals of statistics 7, 812-822.
8. Timm, N.H., (2002) " Applied Multivariate Analysis " Springer – Verlag, New York .
9. Vonesh, E.F., and Chinchilli, V.M.,(1997) Linear and nonlinear Models for the Analysis of Repeated Measurements " Marcel Dakker, Inc., New York .