

ANÁLISIS DE VARIANCIA DE DOS VIAS CON DATOS QUE SE REGISTRAN DURANTE UN PERIODO DE TIEMPO

ANALYSIS OF VARIANCE TWO-WAY DATA RECORDED DURING A PERIOD OF TIME

¹Raphael Félix Valencia Chacón

Resumen

En el desarrollo de diseños experimentales se suele encontrar casos en los que al evaluar a un grupo de tratamientos, las unidades experimentales pueden registrar observaciones durante un período de tiempo. Estas observaciones suelen estar altamente correlacionadas debido al efecto del tiempo y esto enmascara la información recogida alterando los resultados que se puedan obtener. En el presente trabajo de investigación se desarrollará una metodología adecuada para realizar el análisis de variancia de dos vías con datos que se registran durante un periodo de tiempo.

Palabras claves: variancia.

Summary

In the development of experimental designs, it is usual to find group of treatments, where the experimental data have to be registered in periods of time. These observations are usually correlated due to the time effect, and this effect masks the information that is collected and alters the results. The present research develops an adequate methodology to proceed with an appropriate Analysis of Variance that registers data with the periods of time.

Key words: Variance.

1. Introducción

En algunas investigaciones es habitual encontrar situaciones en las que una variable es observada durante un determinado periodo de tiempo, recogiendo valores de dicha variable en momentos puntuales, que pueden estar prefijados antes de comenzar el ensayo.

Es necesario destacar el carácter especial de este tipo de datos, ya que las observaciones registradas durante un periodo de tiempo suelen estar altamente correlacionadas y esto enmascara la información recogida alterando los resultados que se puedan obtener.

El propósito de este trabajo de investigación es desarrollar una metodología estadística adecuada cuando se tiene información de individuos o de unidades experimentales y de quienes a su vez se recoge información durante un periodo de tiempo determinado.

2. Revisión de literatura

Algunos experimentos suelen contener mediciones repetidas de una o más variables durante un periodo de tiempo. En muchos casos esto es deseable para comparar dos o más grupos de medias de estas medidas repetidas. Una aproximación común para este problema es fijarse muy bien que los datos han surgido de un diseño anidado, Steel and Torrie (8).

Hay experimentos en que debido a la naturaleza de las unidades experimentales, estas a veces son muy grandes;

en este caso es necesario sacar sub-unidades de cada unidad. Mediante los resultados de un experimento analizado previamente con unidades y sub-unidades, se puede estudiar la relación existente entre las unidades y las sub-unidades para reducir el error experimental, en estos casos se obtendrán dos fuentes de variabilidad: la variabilidad entre las sub-unidades de una misma unidad experimental y la variabilidad entre las unidades experimentales de un mismo tratamiento, Calzada (3).

Una serie de tiempo es un conjunto de valores registrados u observados durante periodos de tiempo. En el análisis de una serie temporal se intenta encontrar aquellos factores que ejercen influencia sobre cada uno de los valores periódicos de la serie; Box and Jenkins (2).

Para verificar el supuesto de homogeneidad de varianzas, se utiliza la prueba de Esfericidad de Mauchly, la cual indica que las varianzas de las diferencias entre todos los pares de medidas repetidas sean iguales (Prueba de Esfericidad de Mauchly)

En el enfoque clásico al analizar una serie de tiempo mediante un procedimiento llamado descomposición, se identifica cuatro de estos factores llamados componentes: tendencia (T), fluctuaciones cíclicas (C), variaciones estacionales (E) y variaciones irregulares (I).

El modelo en que se basa el análisis clásico de una serie de tiempo se apoya en la suposición de que, para cualquier periodo de la serie de tiempo, el valor de la variable está determinado por el efecto de estos cuatro componentes. La

¹Facultad de Economía y Planificación. Universidad Nacional Agraria la Molina Av. La Molina S/N, Lima, Perú. Email: raphael@lamolina.edu.pe

componente de tendencia son movimientos de largo plazo en una serie histórica que se puede describir mediante una línea recta o curva y la variable independiente es el tiempo, Bjornsson (1).

El método más ampliamente utilizado para describir tendencias de línea recta es el denominado método de mínimos cuadrados. Este enfoque calcula la línea que mejor se ajusta a un grupo de puntos, de conformidad matemática con un criterio establecido.

Para este caso la ecuación de tendencia sería:

$$Y = b_0 + bX$$

Donde Y es el valor de la predicción de la tendencia lineal, b_0 es el valor de la tendencia lineal en el periodo cero, b es el incremento o disminución promedio de Y para incrementos de un periodo de X.

La ecuación lineal o modelo que representa un diseño completamente aleatorizado de dos vías se muestra de la siguiente forma:

$$Y_{kj} = \mu + \alpha_k + \beta_j + \epsilon_{kj} \quad k = 1, 2, \dots, t \quad j = 1, 2, \dots, r$$

Donde μ es la media global de todas las poblaciones sometidas al tratamiento k, α_k es el efecto del tratamiento de un grupo particular k, β_j es el efecto del bloque de un grupo particular j y ϵ_{kj} es el error aleatorio.

Implícita a ésta aproximación es la suposición que las medias individuales son independientes, idénticas y normalmente distribuidas con media cero.

El procedimiento del análisis de variancia con dos criterios de clasificación se ocupa de probar la existencia de diferencias significativas entre las k medias muestrales cuando se asignan los elementos en forma aleatoria a cada uno de los diversos grupos de tratamientos y de bloques.

Más específicamente, sea Y_{ijt} las observaciones en el tiempo t del j-ésimo individuo del grupo i y suponga que los Y_{ijt} pueden ser representados por la siguiente expresión:

$$Y_{ijt} = \mu_i + \alpha_{ij} + \beta_j + \epsilon_{ijt} \\ i = 1, 2, \dots, m \quad j = 1, 2, \dots, n \quad t = 1, 2, \dots, t$$

Donde μ_i son grupos de medias corregidos, α_{ij} representan los efectos individuales, β_j representan los efectos de bloques y ϵ_{ijt} son los residuos. Asumiremos adicionalmente que ϵ_{ijt} , α_{ij} y β_j son independientes, α_{ij} y β_j son variables aleatorias independientes con media cero y variancia σ_α^2 y variancia σ_β^2 respectivamente; ϵ_{ijt} son series de tiempo para i y j corregidos.

El problema de probar la hipótesis nula de igualdad de medias de grupos será considerado. En muchos problemas prácticos el usual Análisis de Variancia para la Prueba de F, representada sobre promedios de observaciones en el tiempo, brinda un estadístico de prueba eficiente.

Antes de aplicar el modelo de análisis de la variancia a los datos de un determinado estudio, es necesario evaluar los supuestos: Normalidad y Homogeneidad de varianzas.

Un supuesto implícito en este modelo (medidas repetidas), es que para que los valores de F con los grados de libertad especificados sean válidos, **debe cumplirse la condición denominada de esfericidad**, que viene a decir que las covarianzas entre cada par de tratamientos son las mismas (de ocasión a ocasión el cambio es idéntico).

Cuando se mide en varias veces sucesivas a los mismos sujetos (y siempre que tengamos muestras relacionadas) como es frecuente en muchos diseños experimentales (o en estudios exploratorios), estas medidas están correlacionadas; en este caso bajan los cuadrados medios del término del error (el denominador de la razón F) y se obtiene con mayor facilidad un valor de F significativo.

En consecuencia, el incumplimiento de este supuesto repercute gravemente en los resultados arrojados por la prueba F, ya que no es robusta ante observaciones correlacionadas. Concretamente, la correlación entre las puntuaciones provoca un sesgo positivo en la prueba F, aumentando la probabilidad de cometer errores tipo I. Es decir, aumenta la probabilidad de encontrar resultados estadísticamente significativos cuando, realmente, no lo son. Por lo tanto, para analizar los datos procedentes de un diseño MR mediante el modelo univariado de Análisis de Varianza (ANOVA), será necesario comprobar si en dichos datos se cumple el supuesto de esfericidad, siempre que la(s) VI(s) implicada(s) tengan más de dos niveles

Uno de los test más utilizados para evaluar el supuesto de esfericidad es la W de Mauchly (1940). Esta prueba la aporta por defecto la salida que el SPSS.

Cuando hay esfericidad entonces, los datos se pueden analizar mediante el modelo univariado de ANOVA sin correr el riesgo de obtener un valor sesgado de la prueba F. Sin embargo, si se rechaza la H_0 , el supuesto de esfericidad no se puede mantener y, por tanto, el modelo univariado de ANOVA no es apropiado para analizar los datos.

3. Materiales y métodos

Modelo Aditivo Lineal

Para el presente trabajo de investigación se plantea el siguiente modelo aditivo lineal:

$$Y_{ijt} = \mu_i + \alpha_{ij} + \beta_j + \epsilon_{ijt}$$

$$i = 1, 2, \dots, m \quad j = 1, 2, \dots, n \quad t = 1, 2, \dots, t$$

Donde:

Y_{ijt} : representa el resultado observado para el tiempo t, del individuo j al que se le aplicó el tratamiento i.

μ_i : grupos de medias corregidos.

α_{ij} : representan el efecto individual del i-ésimo tratamiento para el j-ésimo individuo..

β_j : representan el efecto del j-ésimo bloque.

ϵ_{ijt} : Error Total.

Supuestos del modelo:

1. El modelo estadístico es aditivo y lineal.
2. Los ε_{ijt} se distribuyen normal e independientemente con media μ_{ε} cero y variancia σ_{ε}^2 .
3. α_{ij} son variables aleatorias independientes con media cero y variancia σ_{α}^2
4. β_j son variables aleatorias independientes con media cero y variancia σ_{β}^2
5. Existe Homogeneidad de Variancias.

Estimación de los Parámetros

Assumiendo el siguiente modelo estadístico muestral:

$$Y_{ijt} = \hat{\mu}_i + \hat{\alpha}_{ij} + \hat{\beta}_j + \hat{\varepsilon}_{ijt}$$

$i = 1, 2, \dots, m \quad j = 1, 2, \dots, n \quad t = 1, 2, \dots, t$

Tenemos que el efecto $\hat{\mu}_i$ se puede descomponer en

$$Q = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^t \varepsilon_{ijk}^2 = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^t \left(Y_{ijk} - \mu - \tau_i - \alpha_{ij} - \beta_j \right)^2$$

$$\sum_{i=1}^m \tau_i = 0 \quad \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} = 0 \quad \sum_{j=1}^n \beta_j = 0$$

Se obtienen los estimadores de los parámetros:

$$\hat{\mu} = \frac{Y_{...}}{m n t}$$

$$\hat{\tau}_i = \frac{Y_{i..}}{n t} - \frac{Y_{...}}{m n t}$$

$$\hat{\alpha}_{ij} = \frac{Y_{ij.}}{t} - \frac{Y_{i..}}{n t} - \frac{Y_{.j.}}{m t} + \frac{Y_{...}}{m n t}$$

$$\hat{\beta}_j = \frac{Y_{.j.}}{m t} - \frac{Y_{...}}{m n t}$$

Luego de descomponer la variabilidad total en cada una de las fuentes de variación se llega a obtener las sumas de cuadrados.

SUMA DE CUADRADOS:

Suma de Cuadrados de Total (S.C. Total)

$$S.C. Total = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^t Y_{ijk}^2 - \frac{Y_{...}^2}{m n t}$$

Suma de Cuadrados de Tratamientos (S.C. Tratamientos)

$$S.C. Tratamientos = \sum_{i=1}^m \frac{Y_{i..}^2}{n t} - \frac{Y_{...}^2}{m n t}$$

Suma de Cuadrados de Bloques (S.C. Bloques)

$$S.C. Bloques = \sum_{j=1}^n \frac{Y_{.j.}^2}{m t} - \frac{Y_{...}^2}{m n t}$$

$\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i$ entonces el modelo estadístico muestral se podría representar como:

$$Y_{ijt} = \hat{\mu} + \hat{\tau}_i + \hat{\alpha}_{ij} + \hat{\beta}_j + \hat{\varepsilon}_{ijt}$$

$i = 1, 2, \dots, m \quad j = 1, 2, \dots, n \quad t = 1, 2, \dots, t$

donde: $\hat{\mu}, \hat{\tau}_i, \hat{\alpha}_{ij}, \hat{\beta}_j, \hat{\varepsilon}_{ijt}$ son los estimadores de los correspondientes parámetros:

$\mu, \tau_i, \alpha_{ij}, \beta_j, \varepsilon_{ijt}$ respectivamente.

Para obtener las ecuaciones normales y los estimadores de los parámetros se aplicará el método de los mínimos cuadrados, que para este caso consiste en minimizar la suma de cuadrados de Error Experimental (SC(EE)) con respecto a cada uno de los estimadores. La suma de cuadrados del Error Experimental (SC(EE)) está dada por la siguiente expresión:

Suma de Cuadrados del Error Experimental (S.C. (E.E.))

$$S.C. (E.E.) = \sum_{i=1}^m \sum_{j=1}^n \frac{Y_{ij.}^2}{t} - \sum_{i=1}^m \frac{Y_{i..}^2}{n t} - \sum_{j=1}^n \frac{Y_{.j.}^2}{m t} + \frac{Y_{...}^2}{m n t}$$

Suma de Cuadrados del Error Total (S.C. (E.T.))

$$S.C. (E.T.) = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^t Y_{ijk}^2 - \sum_{i=1}^m \sum_{j=1}^n \frac{Y_{ij.}^2}{t}$$

Para probar la hipótesis si existen diferencias entre las unidades experimentales tratadas con el mismo tratamiento; se establece la siguiente Hipótesis planteada:

Ho: $\mu_1 = \mu_2 = \dots = \mu_n$ contra la hipótesis alternante H1: Ho no es verdadera.

El estadístico de prueba a realizar sería:

$$F_c = \frac{SC (Error Experimental) / (m - 1)(n - 1)}{S.C. (Error Total) / mn (t - 1)}$$

y F_c tendrá una distribución F de Fisher con $((m - 1)(n - 1), mn (t - 1))$ grados de libertad.

Para probar la hipótesis si existen diferencias entre los tratamientos; se establece la siguiente Hipótesis planteada:

Ho: $\mu_1 = \mu_2 = \dots = \mu_m$ contra la hipótesis alternante H1: Ho no es verdadera.

El estadístico de prueba a realizar sería:

$$F_c = \frac{SC (Trat) / (m - 1)}{S.C. (E.E.) / (m - 1)(n - 1)}$$

y F_c tendrá una distribución F de Fisher con $(m - 1, (m - 1)(n - 1))$ grados de libertad.

Luego se tendrá el siguiente análisis de variancia.

Fuente	G.L.	S.C.	C.M.
Tratamientos	m - 1	$\sum_{i=1}^m \frac{Y_{i.}^2}{n t} - \frac{Y_{...}^2}{m n t}$	C.M. (TRAT)
Bloques	n - 1	$\sum_{j=1}^n \frac{Y_{.j}^2}{m t} - \frac{Y_{...}^2}{m n t}$	C.M. (BLOQ)
Error Experimental	(m - 1) (n - 1)	$\sum_{i=1}^m \sum_{j=1}^n \frac{Y_{ij.}^2}{t} - \sum_{i=1}^m \frac{Y_{i.}^2}{n t} - \sum_{j=1}^n \frac{Y_{.j}^2}{m t} + \frac{Y_{...}^2}{m n t}$	C.M. (EE.)
Error Total	m n (t - 1)	$\sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^t Y_{ijk}^2 - \sum_{i=1}^m \sum_{j=1}^n \frac{Y_{ij.}^2}{t}$	C.M. (ET.)
Total	m n t - 1	$\sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^t Y_{ijk}^2 - \frac{Y_{...}^2}{m n t}$	

4. Resultados

Para la aplicación práctica de la metodología descrita se tomó un experimento en el que se desea comparar si existen diferencias significativas en el rendimiento de cinco dietas alimenticias al ser aplicadas en 4 marranos de diferentes especies durante 3 diferentes periodos de tiempo (primera semana, segunda semana y tercera semana). Los incrementos de peso (en kgr.) se muestran a continuación:

Individuo (j)	Periodo de Tiempo (k)	Dieta (i)					
		1	2	3	4	5	
1	1	3.5	5.0	5.0	8.5	6.0	78
	2	4.0	5.5	4.5	6.0	5.5	
	3	3.0	4.0	5.0	9.0	3.5	
	Totales	10.5	14.5	14.5	23.5	15.0	
2	1	2.5	3.5	5.5	6.5	6.0	82.5
	2	4.5	3.5	6.0	7.0	8.5	
	3	5.5	3.0	5.0	8.0	7.5	
	Totales	12.5	10.0	16.5	21.5	22.0	
3	1	3.0	4.5	5.5	7.0	6.5	80
	2	3.0	4.0	4.5	7.0	6.5	
	3	2.5	4.0	6.5	7.0	8.5	
	Totales	8.5	12.5	16.5	21.0	21.5	
4	1	4.5	3.5	4.5	8.5	7.0	86
	2	5.0	4.5	5.0	6.5	7.5	
	3	5.0	4.5	5.5	7.0	7.5	
	Totales	14.5	12.5	15.0	22.0	22.0	
Total Dietas		46.0	49.5	62.5	88.0	80.5	

Interpretación : El efecto de la dieta 1 (t₁) es 1.6087 kgs. inferior al promedio general.

$$\hat{t}_2 = \frac{49.5}{4 \times 3} - \frac{326.5}{5 \times 4 \times 3} = -1.3166$$

Interpretación : El efecto de la dieta 2 (t₂) es 1.3166 kgs. inferior al promedio general.

$$\hat{t}_3 = \frac{62.5}{4 \times 3} - \frac{326.5}{5 \times 4 \times 3} = -0.2333$$

Interpretación : El efecto de la dieta 3 (t₃) es 0.2333 kgs. inferior al promedio general.

$$\hat{t}_4 = \frac{88.0}{4 \times 3} - \frac{326.5}{5 \times 4 \times 3} = 1.8917$$

Interpretación : El efecto de la dieta 4 (t₄) es 1.8917 kgs. superior al promedio general.

$$\hat{t}_5 = \frac{80.5}{4 \times 3} - \frac{326.5}{5 \times 4 \times 3} = 1.2667$$

Interpretación : El efecto de la dieta 5 (t₅) es 1.2667 kgs. superior al promedio general.

$$\hat{\alpha}_{11} = \frac{10.5}{3} - \frac{46.0}{4 \times 3} = -0.333$$

$$\hat{\alpha}_{12} = \frac{12.5}{3} - \frac{46.0}{4 \times 3} = 0.333$$

$$\hat{\alpha}_{54} = \frac{22.0}{3} - \frac{80.5}{4 \times 3} = 0.6250$$

Interpretación : El efecto de la dieta 5 en el individuo 4 (α₅₄) es 0.625 kgs. superior al promedio de la dieta 5.

Verificación de supuestos de Normalidad y Homogeneidad de Varianzas.

Supuesto de Normalidad de Errores.

Ho: Los errores tienen distribución Normal.

H1: Los errores NO tienen distribución Normal

Alfa: 0.05

Prueba Estadística de Anderson Darling

AD = 0.634

Criterios de Decisión

Pvalor = 0.094

Conclusión: No se rechaza H_0 , no existe evidencia estadística para afirmar que los errores no tienen distribución Normal.

Supuesto de Esfericidad de Mauchly

H_0 : La matriz de varianza-covarianza cumple con el

supuesto de esfericidad.

H_1 : La matriz de varianza-covarianza NO cumple con el supuesto de esfericidad.

Alfa: 0.05

Prueba Estadística de Mauchly

Test = 0.984

Prueba de esfericidad de Mauchly^a

Medida: MEASURE_1

Efecto inter sujetos	W de Mauchly	Aprox. Chi-cuadrado	gl	Sig.	Épsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Límite inferior
Tiempo	,984	,286	2	,867	,984	1,000	,500

Prueba la hipótesis nula que la matriz de covarianzas de error de las variables dependientes con transformación ortogonalizada es proporcional a una matriz de identidad.

a. Diseño : Interceptación

Diseño dentro de sujetos: Tiempo

b. Se puede utilizar para ajustar los grados de libertad para las pruebas promedio de significación. Las pruebas corregidas se visualizan en la tabla de pruebas de efectos dentro de sujetos.

Criterios de Decisión

Pvalor = 0.867

Conclusión: No se rechaza H_0 , La matriz de varianza-covarianza cumple con el supuesto de esfericidad.

Análisis de Varianza

Calculo de las sumas de cuadrados:

Sumas de Cuadrados Total:

$$S.C. Total = 1944.25 - \frac{(326.5)^2}{5 \times 4 \times 3} = 167.5458$$

Sumas de Cuadrados de las Dietas:

$$S.C. Tratamientos = \frac{(46.0)^2}{4 \times 3} + \dots + \frac{(80.5)^2}{4 \times 3} - \frac{(326.5)^2}{5 \times 4 \times 3}$$

$$S.C. Tratamientos = 1891.3958 - 1776.7042 = 114.6916$$

Sumas de Cuadrados de los Bloques:

$$S.C. Bloques = \frac{(78.0)^2}{5 \times 3} + \dots + \frac{(86)^2}{5 \times 3} - \frac{(326.5)^2}{5 \times 4 \times 3}$$

$$S.C. Bloques = 1779.0833 - 1776.7042 = 2.3792$$

Suma de Cuadrados del Error Experimental:

$$S.C. (E.E.) = \frac{(10.5)^2}{3} + \frac{(12.5)^2}{3} + \dots + \frac{(22.0)^2}{3} - 1891.3958 - 1779.0833 + 1776.7042$$

$$S.C. (E.E.) = 1915.4167 - 1891.3958 - 1779.0833 + 1776.7042 = 21.6418$$

Suma de Cuadrados del Error Total:

$$S.C. (E.T.) = 1944.25 - 1915.4167 = 28.8333$$

Luego, se tendría el siguiente cuadro de análisis de variancia:

Prueba de hipótesis

H_0 : No hay diferencias significativas entre las unidades experimentales tratadas con el mismo tratamiento.

H_1 : Si hay diferencias significativas en al menos una de las unidades experimentales tratadas con el mismo tratamiento.

Nivel de Significación: $\alpha = 0.05$

El estadístico de prueba a realizar sería:

$$F_c = \frac{21.6418/12}{28.8333/40} = 2.50$$

y F_c tendrá una distribución $F(12,40)$ gl. = 2.0035

Las evidencias estadísticas muestrales nos indican que si existen diferencias significativas en al menos una de las unidades experimentales tratadas con el mismo tratamiento.

H_0 : Las medias de las 5 dietas alimenticias son iguales al evaluar el incremento de peso (en Kgrs.)

H_1 : Al menos una de las medias de las 5 dietas alimenticias es diferente a las demás al evaluar el incremento de peso (en Kgrs.)

Nivel de Significación: $\alpha = 0.05$

Cuadro ANVA

Fuente	G.L.	S.C.	C.M.	Fcalc.
Dietas	4	114.6916	28.6729	36.15
Bloques	3	2.3792	0.7931	
Error Experimental	12	21.6418	1.8035	
Error Total	40	28.8333	0.7208	2.50
Total	59	167.5458		

El estadístico de prueba a realizar sería:

$$F_c = \frac{114.6916 / 4}{21.6418 / 12} = 36.15$$

y F_c tendrá una distribución $F(4,12)$ gl. = 3.25917

Las evidencias estadísticas muestrales nos indican que si existen diferencias significativas entre al menos una de las dietas alimenticias al evaluar el incremento de peso (en kgrs.).

5. Conclusiones

La metodología presentada en el presente trabajo de investigación, se podrá utilizar cuando se tiene información de unidades experimentales y dentro de las cuales se recoge información durante un periodo de tiempo determinado.

Cuando se realiza el análisis de variancia se encontrará dos fuentes de variabilidad del error: una de ellas corresponderá al error experimental, debido a las diferencias encontradas entre las observaciones registradas dentro de los tratamientos sujetos a evaluación y la otra fuente corresponderá al error total, debido a las diferencias encontradas al interior de las observaciones registradas las cuales corresponderían a los datos registrados durante un periodo de tiempo.

Si se desea verificar la existencia de diferencias significativas entre las unidades experimentales tratadas con el mismo tratamiento, para esto se realizará una prueba de F dividiendo el Cuadrado Medio del Error Experimental con el Cuadrado Medio del Error Total; si esta prueba resulta significativa entonces se llega a la conclusión que existe variabilidad entre las unidades experimentales de un mismo tratamiento, pero si no resulta significativa la prueba de hipótesis, entonces la variabilidad se deberá a la variabilidad de las mediciones realizadas dentro de cada una de las unidades experimentales, es decir a la variabilidad de los registros realizados durante el periodo de tiempo establecido.

En este ultimo caso la variancia del error para la prueba de medias de tratamientos y las consiguientes pruebas de comparación de medias tendrá que volver a ser calculada y se obtiene sumando la suma de cuadrados del error experimental y la suma de cuadrados del error total dividido entre suma de los grados de libertad de la fuente del error experimental y los grados de libertad del error total.

Como las mediciones de las unidades experimentales para un periodo establecido de tiempo pueden estar correlacionadas, se recomienda que estas mediciones se realicen en el mismo instante para cada una de las observaciones experimentales en cada uno de los tratamientos sometidos a evaluación.

6. Literatura citada

Bjornsson, H. (1978). Analysis of a series of long-term grassland experiments with autocorrelated errors. *Biometrics* 34, 645 – 651.

Box, G. E. P. and Jenkins, G. M. (1970). Time Series Analysis, Forecasting and control. San Francisco: Holden-Day.

Calzada Benza, José (1970). Métodos Estadísticos para la Investigación.

International Mathematical Statistical Libraries (1980). IMSL Library Reference Manual. Houston: IMSL.

Johnson, N. L. and Kotz, S. (1970). Continuous Univariate Distributions, Vol 2. New York: Wiley.

Quenouille, M. H. (1958). The comparison of correlations in time series. *Journal of the Royal Statistical Society, Series B* 20, 158 – 164.

Searle, S. R. (1971). Linear Models. New York: Wiley.

Steel, R. C. D. and Torrie, J. H. (1960). Principles and Procedures of Statistics. New York: Mc. Graw-Hill.

Swamy, P.A.V.B. (1970). Statistical Inference in Random Coefficient Regression Models. Berlin: Springer – Verlag.

Timm, N. H. (1975). Multivariate Analysis with Application in Education and Psychology. Monterrey: Wadsworth.