



ANÁLISIS DE VARIANZA


UNIDAD N° 5

OBJETIVO DEL ANÁLISIS DE VARIANZA


Explicar (controlar) las variaciones de una variable aleatoria X continua (numérica), mediante factores (variables cualitativas que definen categorías) que controlamos (no aleatorios)

Este análisis permite poner en evidencia eventuales relaciones entre X y estos factores.

La teoría de muestreo muestra la importancia de la diferencia entre dos medias muestrales, suponiendo que las poblaciones de las que provenían las muestras tenían la misma varianza



Puede ser necesario probar la importancia de la diferencia entre tres o más medias muestrales



Esto equivale a probar la hipótesis nula de que todas estas medias muestrales son iguales



Utilizaremos la técnica de *análisis de varianza*, desarrollada por Fisher, empleando la distribución F

CLASIFICACIÓN EN UN SENTIDO O EXPERIMENTOS CON UN FACTOR

En un experimento de un factor, las mediciones (u observaciones) se hace de ***a*** grupos independientes de muestras, y ***b*** es la cantidad de mediciones en cada muestra

Se habla de ***a*** *tratamientos*, cada uno con ***b*** *repeticiones*

Los resultados de un experimento de un factor se acostumbra presentarlos en una tabla con ***a*** filas y ***b*** columnas

CLASIFICACIÓN EN UN SENTIDO O EXPERIMENTOS CON UN FACTOR

- AQUÍ X_{jk} DENOTA LA MEDICIÓN DE LA FILA j Y COLUMNA k , DONDE $j = 1, 2, \dots, a$ Y DONDE $k = 1, 2, \dots, b$. POR EJEMPLO X_{35} , SIGNIFICA LA QUINTA MEDICIÓN DEL TERCER TRATAMIENTO

Tratamiento 1	$X_{11}, X_{12}, \dots, X_{1b}$	\bar{X}_1
Tratamiento 2	$X_{21}, X_{22}, \dots, X_{2b}$	\bar{X}_2
\vdots	\vdots	
Tratamiento a	$X_{a1}, X_{a2}, \dots, X_{ab}$	\bar{X}_a

- LA MEDIA DE LAS MEDICIONES EN LA FILA j SE DENOTA

$$\bar{X}_j = \frac{1}{b} \sum_{k=1}^b X_{jk} \qquad j = 1, 2, \dots, a$$

- EL PUNTO EN \bar{X}_j SE UTILIZA PARA INDICAR QUE EL ÍNDICE k SE HA SUMADO. LOS VALORES \bar{X}_j SE DENOMINAN *MEDIAS DE GRUPO, MEDIAS DE TRATAMIENTO O MEDIAS DE FILA*.
- LA GRAN MEDIA O MEDIA TOTAL ES LA MEDIA DE TODAS LAS MEDIAS EN TODOS LOS GRUPOS Y SE DENOTA POR \bar{X} , ESTO ES:

$$\bar{X} = \frac{1}{ab} \sum_{j,k} X_{jk} = \frac{1}{ab} \sum_{j=1}^a \sum_{k=1}^b X_{ij}$$

MODELO MATEMÁTICO PARA EL ANÁLISIS DE VARIANZA

Cada fila de la tabla anterior, se considera como una muestra aleatoria de tamaño b tomada de la población de ese determinado tratamiento.

Las X_{jk} difieren de la media poblacional μ_j correspondiente al tratamiento j en un *error aleatorio* que se denota ε_{jk} ; por lo tanto:

$$X_{jk} = \mu_j + \varepsilon_{jk}$$

Se supone que estos errores están distribuidos de manera normal con media 0 y varianza σ^2 . Si μ es la media de la población de todos los tratamientos y si se denota $\alpha_j = \mu_j - \mu$, entonces $\mu_j = \mu + \alpha_j$, que es la cantidad que depende del tratamiento usado (descompuesta en una media global y un efecto del tratamiento) y ε_{jk} es la cantidad que depende solamente de la unidad experimental y que se identifica con el error experimental, de esta manera tenemos:

$$X_{jk} = \mu + \alpha_j + \varepsilon_{jk} \quad \text{donde} \quad \sum_j \alpha_j = 0$$

Según lo explicado anteriormente, se concluye que X_{jk} se pueden considerar como variables aleatorias distribuidas en forma normal, con media μ y varianza σ^2

MODELO MATEMÁTICO PARA EL ANÁLISIS DE VARIANZA

La hipótesis de que los distintos tratamientos no producen ningún efecto (o la de que las medias de todas las poblaciones son iguales) se contrasta mediante el análisis de la varianza de una vía, comparando la variabilidad entre grupos con la variabilidad dentro de los grupos.

$$H_0: \alpha_j = 0 \quad j = 1, 2, \dots, a \quad \text{o} \quad H_0: \mu_j = \mu \quad j = 1, 2, \dots, a$$

Si H_0 es verdadera, todas la poblaciones de los tratamientos tendrán la misma distribución normal (es decir, con la misma media y varianza). En estos casos, sólo hay un tratamiento poblacional (es decir, todos los tratamientos son estadísticamente idénticos); en otras palabras, no hay diferencia significativa entre los tratamientos

ANÁLISIS DE VARIANZA

El análisis de la varianza se basa en la descomposición de la variabilidad total en dos partes, una parte debida a la variabilidad entre las distintas poblaciones o tratamientos (variabilidad entre grupos o variabilidad explicada por el diseño) y otra parte que puede considerarse como la variabilidad intrínseca de las observaciones (variabilidad dentro de los grupos o residual)

VARIACIÓN TOTAL

Definimos la variación total, denotada por v , como la suma de los cuadrados de las desviaciones de cada medida de la gran media \bar{X} , es decir: $V = \sum_{j,k} (X_{jk} - \bar{X})^2$

VARIACIÓN DENTRO DE LOS TRATAMIENTOS

Se trata de los cuadrados de las desviaciones de las X_{jk} respecto a las medias de los tratamientos \bar{X}_j y se denota por V_W : $V_W = \sum_{j,k} (X_{jk} - \bar{X}_j)^2$
mide la variabilidad intrínseca de las observaciones, es decir, si el experimento está bien diseñado y no se incluyen factores de variación distintos al estudiado, debe ser error puramente aleatorio producido como resultado de la variabilidad biológica del material experimental

VARIACIÓN ENTRE LOS TRATAMIENTOS

Se trata de los cuadrados de las desviaciones de las medias de los tratamientos \bar{X}_j respecto a la gran media \bar{X} , y se denota por V_B : $V_B = \sum_{j,k} (\bar{X}_j - \bar{X})^2 = b \sum_j (\bar{X}_j - \bar{X})^2$
mide la discrepancia entre los grupos y la media global, de forma que si no hay diferencias entre ellos (la hipótesis nula es cierta) obtendremos variabilidades pequeñas. Si, por el contrario, la hipótesis nula es falsa, cabe esperar que la variabilidad entre grupos sea grande

RELACIÓN ENTRE VARIACIONES: $V = V_W + V_B$

MÉTODOS ABREVIADOS PARA OBTENER LAS VARIACIONES

Para simplificar el cálculo de las variaciones anteriores se emplean las siguientes fórmulas:

$$V = \sum_{j,k} X_{jk}^2 - \frac{T^2}{ab}$$

$$V_B = \frac{1}{b} \sum_{j,k} T_j^2 - \frac{T^2}{ab}$$

$$V_W = V - V_B$$

Donde T es la suma de todos los valores X_{jk} y donde T_j es la suma de todos los valores del tratamiento j -ésimo:

$$T = \sum_{j,k} X_{jk}$$

$$T_j = \sum_k X_{jk}$$

VALORES ESPERADOS DE LAS VARIACIONES

Los valores esperados de V_W , V_B , y V están dados por:

$$E(V_W) = a(b-1)\sigma^2$$

$$E(V_B) = (a-1)\sigma^2 + b \sum_j \alpha_j^2$$

$$E(V) = (ab-1)\sigma^2 + b \sum_j \alpha_j^2$$

De acuerdo con lo anterior tenemos los siguientes estimadores de σ^2 , sin importar si H_0 es o no verdadera.

$$E\left[\frac{V_W}{a(b-1)}\right] = \sigma^2 \quad \text{de manera que}$$
$$\hat{S}_W^2 = \frac{V_W}{a(b-1)}$$

Si H_0 es verdadera (es decir, $\alpha_j = 0$) se tendrá:

$$E\left[\frac{V_B}{a-1}\right] = \sigma^2 \quad \text{entonces} \quad \hat{S}_B^2 = \frac{V_B}{a-1}$$

$$E\left[\frac{V}{ab-1}\right] = \sigma^2 \quad \text{entonces} \quad \hat{S}^2 = \frac{V}{ab-1}$$

Si H_0 no es verdadera, se tendrá:

$$E(\hat{S}_B^2) = \sigma^2 + \frac{b}{a-1} \sum_j \alpha_j^2$$

TEOREMAS

Teorema 1: $\frac{V_W}{\sigma^2}$ tienen una distribución ji cuadrada con $a(b - 1)$ grados de libertad. Este teorema es válido, ya sea que H_0 sea o no verdadera.

Teorema 2: Bajo la hipótesis nula H_0 , $\frac{V_B}{\sigma^2}$ y $\frac{V}{\sigma^2}$ tienen distribuciones ji cuadrada con $a - 1$ y $ab - 1$ grados de libertad, respectivamente. Este teorema es válido bajo la suposición de que H_0 es verdadera.

Teorema 3: El estadístico $F = \frac{\hat{S}_B^2}{\hat{S}_W^2}$ tiene una distribución F con $a - 1$ y $a(b - 1)$ grados de libertad. Este teorema permite probar la hipótesis nula a determinado nivel de significancia, empleando la distribución F mediante una prueba de una cola.

ANÁLISIS DE VARIANZA

El contraste del Análisis de la varianza se basa en la comparación de la variabilidad entre y la variabilidad dentro, rechazaremos la hipótesis nula siempre que la variabilidad “entre” sea grande, pero utilizando como patrón de comparación la variabilidad “dentro”. Es decir, aceptaremos un efecto de los tratamientos siempre que estos produzcan mayores diferencias en las unidades experimentales que las que habría sin la aplicación de los mismos.

Antes de proceder a la comparación hemos de dividir las sumas de cuadrados por sus correspondientes grados de libertad, relacionados con el número de observaciones con las que se realiza el cálculo.

De esta forma obtenemos los cuadrados medios o estimadores de las variabilidades.

TABLAS PARA EL ANÁLISIS DE VARIANZA

Variación	Grados de libertad	Cuadrado medio (Estimador)	F
Entre tratamientos $V_B = b \sum_j (\bar{X}_j - \bar{X})^2$	$a - 1$	$\hat{S}_B^2 = \frac{V_B}{a - 1}$	$\frac{\hat{S}_B^2}{\hat{S}_W^2}$ Con $a - 1$ y $a(b - 1)$ grados de libertad
Dentro de tratamientos $V_W = V - V_B$ $= \sum_{j,k} (X_{jk} - \bar{X}_j)^2$	$a(b - 1)$	$\hat{S}_W^2 = \frac{V_W}{a(b - 1)}$	
Total $V = V_W + V_B$ $= \sum_{j,k} (X_{jk} - \bar{X})^2$	$ab - 1$		

MODIFICACIONES PARA NÚMEROS DISTINTOS DE OBSERVACIONES

- EN CASO DE QUE LOS TRATAMIENTOS 1, ... A TENGAN NÚMEROS DISTINTOS DE OBSERVACIONES –IGUALES A N_1, \dots, N_A , RESPECTIVAMENTE– LOS RESULTADOS ANTERIORES PUEDEN MODIFICARSE:

$$V = \sum_{j,k} (X_{jk} - \bar{X})^2 = \sum_{j,k} X_{jk}^2 - \frac{T^2}{ab}$$

$$V_B = \sum_{j,k} (\bar{X}_j - \bar{X})^2 = \sum_j N_j (\bar{X}_j - \bar{X})^2 = \sum_j \frac{T_j^2}{N_j} - \frac{T^2}{N}$$

$$V_W = V - V_B$$

DONDE $\sum_{j,k}$ DENOTA LA SUMATORIA, PRIMERO SOBRE K DESDE 1 HASTA N_j Y DESPUÉS LA SUMATORIA SOBRE J DESDE 1 HASTA A.

MODIFICACIONES PARA NÚMEROS DISTINTOS DE OBSERVACIONES

En caso de que los tratamientos $1, \dots, a$ tengan números distintos de observaciones –iguales a N_1, \dots, N_a , respectivamente– los resultados anteriores pueden modificarse

$$V = \sum_{j,k} (X_{jk} - \bar{X})^2 = \sum_{j,k} X_{jk}^2 - \frac{T^2}{ab}$$

$$V_B = \sum_{j,k} (\bar{X}_j - \bar{X})^2 = \sum_j N_j (\bar{X}_j - \bar{X})^2 = \sum_j \frac{T_j^2}{N_j} - \frac{T^2}{N}$$

$$V_W = V - V_B$$

Donde $\sum_{j,k}$ denota la sumatoria, primero sobre k desde 1 hasta N_j y después la sumatoria sobre j desde 1 hasta a

TABLAS PARA EL ANÁLISIS DE VARIANZA (Nº DISTINTO DE OBSERVACIONES)

Variación	Grados de libertad	Cuadrado medio (Estimador)	F
Entre tratamientos $V_B = \sum_j N_j (\bar{X}_j - \bar{X})^2$	$a - 1$	$\hat{S}_B^2 = \frac{V_B}{a - 1}$	$\frac{\hat{S}_B^2}{\hat{S}_W^2}$ Con $a - 1$ y $N - a$ grados de libertad
Dentro de tratamientos $V_W = V - V_B$	$N - a$	$\hat{S}_W^2 = \frac{V_W}{N - a}$	
Total $V = V_W + V_B$ $= \sum_{j,k} (X_{jk} - \bar{X})^2$	$N - 1$		

CLASIFICACIÓN PARA EXPERIMENTOS CON DOS FACTORES

En muchas situaciones prácticas las unidades experimentales no son homogéneas por lo que conviene agruparlas en distintos conjuntos de observaciones homogéneas. A tales conjuntos se les denomina **bloques**

Los tratamientos se aplican dentro de cada bloque siguiendo las mismas técnicas de aleatorización expuestas previamente

Se procurará que los tratamientos estén representados de la misma manera en todos los bloques

NOTACIÓN PARA EXPERIMENTOS CON DOS FACTORES

- CUANDO SE TIENEN A TRATAMIENTOS Y B BLOQUES, SE SUPONE QUE PARA CADA TRATAMIENTO Y PARA CADA BLOQUE HAY UN VALOR EXPERIMENTAL. X_{JK} DENOTA EL TRATAMIENTO J Y EL BLOQUE K .

	Bloque				
	1	2	...	b	
Tratamiento 1	X_{11}	X_{12}	...	X_{1b}	\bar{X}_1
Tratamiento 2	X_{21}	X_{22}	...	X_{2b}	\bar{X}_2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
Tratamiento a	X_{a1}	X_{a2}	...	X_{ab}	\bar{X}_a
	\bar{X}_1	\bar{X}_2		\bar{X}_b	

- LA MEDIA DE LAS ENTRADAS EN EL RENGLÓN J SE DENOTA \bar{X}_j , DONDE $J = 1, \dots, A$, Y LA MEDIA DE LAS ENTRADAS EN LA COLUMNA K SE DENOTA \bar{X}_k , DONDE $K = 1, \dots, B$
- LA MEDIA GENERAL, SE DENOTA \bar{X}

$$\bar{X}_j = \frac{1}{b} \sum_{k=1}^b X_{jk} \quad \bar{X}_k = \frac{1}{a} \sum_{j=1}^a X_{jk} \quad \bar{X} = \frac{1}{ab} \sum_{j,k} X_{jk}$$

VARIACIONES EN LOS EXPERIMENTOS CON DOS FACTORES

$$V = \sum_{j,k} (X_{jk} - \bar{X})^2 = V_E + V_R + V_C$$

$$V_E = \text{variación debida al error o a la casualidad} = \sum_{j,k} (X_{jk} - \bar{X}_j - \bar{X}_k + \bar{X})^2$$

(V_E variación residual o variación aleatoria)

$$V_R = \text{variación entre renglones (tratamientos)} = b \sum_{j=1}^a (\bar{X}_j - \bar{X})^2$$

$$V_C = \text{variación entre columnas (bloques)} = a \sum_{k=1}^b (\bar{X}_k - \bar{X})^2$$

- FÓRMULAS DE CÁLCULO ABREVIADAS

$$V = \sum_{j,k} X_{jk}^2 - \frac{T^2}{ab} \quad V_R = \frac{1}{b} \sum_{j=1}^a T_j^2 - \frac{T^2}{ab} \quad V_C = \frac{1}{a} \sum_{k=1}^b T_k^2 - \frac{T^2}{ab} \quad V_E = V - V_R - V_C$$

DONDE T_j ES EL TOTAL (LA SUMA) DE LAS ENTRADAS EN EL RENGLÓN J -ÉSIMO, T_k ES EL TOTAL (LA SUMA) DE LAS ENTRADAS EN LA COLUMNA K , Y T ES EL TOTAL (LA SUMA) DE TODAS LAS ENTRADAS

ANÁLISIS DE VARIANZA PARA EXPERIMENTOS CON DOS FACTORES

El modelo matemático para experimentos con dos factores está dado por:

$$X_{jk} = \mu + \alpha_j + \beta_k + \varepsilon_{jk}$$

Donde $\sum \alpha_j = 0$ y $\sum \beta_k = 0$

μ es la media de la población

α_j es la parte de X_{jk} atribuida a los diferentes tratamientos (efectos del tratamiento)

β_k es la parte de X_{jk} atribuida a los diferentes bloques (efectos de los bloques)

ε_{jk} es la parte de X_{jk} atribuida a la casualidad o al error

Las esperanzas de las variaciones están dadas por:

$$E(V_E) = (a - 1)(b - 1)\sigma^2$$

$$E(V_R) = (a - 1)\sigma^2 + b \sum_j \alpha_j^2$$

$$E(V_C) = (b - 1)\sigma^2 + a \sum_k \beta_k^2$$

$$E(V) = (ab - 1)\sigma^2 + b \sum_j \alpha_j^2 + a \sum_k \beta_k^2$$

ANÁLISIS DE VARIANZA PARA EXPERIMENTOS CON DOS FACTORES

Las hipótesis nulas que se quieren probar son dos:

$H_0^{(1)}$: todas las medias de los tratamientos (renglones) son iguales, es decir, $\alpha_j = 0$ y $j = 1, \dots, a$

$H_0^{(2)}$: todas las medias de los bloques (columnas) son iguales, es decir, $\beta_k = 0$ y $k = 1, \dots, b$

Si las hipótesis nulas, son p no verdaderas, una estimación insesgada está dada por:

$$\hat{S}_E^2 = \frac{V_E}{(a-1)(b-1)} \quad \text{esto es} \quad E(\hat{S}_E^2) = \sigma^2$$

Si las hipótesis nulas son verdaderas, entonces:

$$\hat{S}_R^2 = \frac{V_R}{a-1} \quad \hat{S}_C^2 = \frac{V_C}{b-1} \quad \hat{S}^2 = \frac{V}{ab-1}$$

Si las hipótesis nulas no son verdaderas, entonces:

$$E(\hat{S}_R^2) = \sigma^2 \frac{b}{a-1} \sum_j \alpha_j^2 \quad E(\hat{S}_C^2) = \sigma^2 \frac{a}{b-1} \sum_k \beta_k^2$$

TEOREMAS

Teorema 4: $\frac{V_E}{\sigma^2}$ es una distribución ji cuadrada con $(a - 1)(b - 1)$ grados de libertad, independientemente de $H_0^{(1)}$ o bien $H_0^{(2)}$

Teorema 5: Si la hipótesis $H_0^{(1)}$ es verdadera, $\frac{V_R}{\sigma^2}$ tiene una distribución ji cuadrada con $a - 1$ grados de libertad. Si la hipótesis $H_0^{(2)}$ es verdadera, $\frac{V_C}{\sigma^2}$ tiene una distribución ji cuadrada con $b - 1$ grados de libertad. Si las dos hipótesis $H_0^{(1)}$ y $H_0^{(2)}$ son verdaderas, $\frac{V}{\sigma^2}$ es una distribución ji cuadrada con $ab - 1$ grados de libertad

Teorema 6: Si la hipótesis $H_0^{(1)}$ es verdadera, el estadístico $\frac{\hat{S}_R^2}{\hat{S}_E^2}$ tiene una distribución F con $a - 1$ y $(a - 1)(b - 1)$ grados de libertad. Si la hipótesis $H_0^{(2)}$ es verdadera, el estadístico $\frac{\hat{S}_C^2}{\hat{S}_E^2}$ tiene una distribución F con $b - 1$ y $(a - 1)(b - 1)$ grados de libertad.

TABLAS PARA EL ANÁLISIS DE VARIANZA (EXPERIMENTOS CON DOS FACTORES)

Variación	Grados de libertad	Cuadrado medio (Estimador)	F
<p>Entre tratamientos</p> $V_R = b \sum_j (\bar{X}_j - \bar{X})^2$	$a - 1$	$\hat{S}_R^2 = \frac{V_R}{a - 1}$	$\frac{\hat{S}_R^2}{\hat{S}_E^2}$ Con $a - 1$ y $(a - 1)(b - 1)$ grados de libertad
<p>Entre bloques</p> $V_C = a \sum_k (\bar{X}_k - \bar{X})^2$	$b - 1$	$\hat{S}_C^2 = \frac{V_C}{b - 1}$	$\frac{\hat{S}_C^2}{\hat{S}_E^2}$ Con $b - 1$ y $(a - 1)(b - 1)$ grados de libertad
<p>Residual o aleatoria</p> $V_E = V - V_R - V_C$	$(a - 1)(b - 1)$	$\hat{S}_E^2 = \frac{V_E}{(a - 1)(b - 1)}$	
<p>Total</p> $V = \sum_{j,k} (X_{jk} - \bar{X})^2$ $= V_E + V_R + V_C$	$ab - 1$		

EXPERIMENTOS CON DOS FACTORES CON REPLICACIÓN

Se puede obtener más información acerca de los factores repitiendo el experimento, proceso que se llama *replicación*. En esos casos habrá más de una entrada para cada tratamiento y para cada bloque. Se supondrá que en cada posición hay c entradas; en el caso que los números de repeticiones no sean iguales, se hacen las modificaciones apropiadas

El modelo necesario considerando la replicación es:
$$X_{jkl} = \mu + \alpha_j + \beta_k + \gamma_{jk} + \varepsilon_{jkl}$$

Donde los subíndices j , k y l de X_{jkl} corresponden, respectivamente, al j -ésimo renglón (o tratamiento), a la k -ésima columna (o bloque) y a la l -ésima repetición (o replicación). μ es la media de la población, α_j es la parte de X_{jk} atribuida a los diferentes tratamientos (efectos del tratamiento), β_k es la parte de X_{jk} atribuida a los diferentes bloques (efectos de los bloques), ε_{jk} es un término aleatorio o un término de error, y γ_{jk} denota los efectos de la interacción renglón-columna (o tratamiento-bloque) que se conocen simplemente como interacciones. Se tienen las restricciones:

$$\sum_j \alpha_j = 0 \quad \sum_k \beta_k = 0 \quad \sum_j \gamma_{jk} = 0 \quad \sum_k \gamma_{jk} = 0$$

VARIACIONES EN LOS EXPERIMENTOS CON REPLICACIÓN

$$V = \text{variación total} = \sum_{j,k,l} (X_{jkl} - \bar{X})^2 = V_E + V_R + V_C + V_I$$

$$V_R = \text{variación debida a los renglones} = bc \sum_{j=1}^a (\bar{X}_j - \bar{X})^2$$

$$V_C = \text{variación debida a las columnas} = ac \sum_{k=1}^b (\bar{X}_k - \bar{X})^2$$

$$V_I = \text{interacciones} = c \sum_{j,k} (\bar{X}_{jk} - \bar{X}_j - \bar{X}_k + \bar{X})^2$$

$$V_E = \text{error aleatorio o residual} = \sum_{j,k,l} (X_{jkl} - \bar{X}_{jk})^2$$

- EL VALOR ESPERADO SE ENCUENTRA COMO ANTES. EMPLEANDO, PARA CADA FUENTE DE VARIACIÓN, EL NÚMERO QUE LE CORRESPONDE DE GRADOS DE LIBERTAD
- LA TABLA PARA ANÁLISIS DE VARIANZA QUE SE MUESTRA A CONTINUACIÓN SE USAN PARA PROBAR LAS HIPÓTESIS NULAS:
 - $H_0^{(1)}$: TODAS LAS MEDIAS DE LOS TRATAMIENTOS (RENGLONES) SON IGUALES, ES DECIR, $\alpha_j = 0$
 - $H_0^{(2)}$: TODAS LAS MEDIAS DE LOS BLOQUES (COLUMNAS) SON IGUALES, ES DECIR, $\beta_k = 0$
 - $H_0^{(3)}$: ENTRE TRATAMIENTOS Y BLOQUES NO HAY INTERACCIÓN, ES DECIR, $\gamma_{jk} = 0$

TABLAS PARA EL ANÁLISIS DE VARIANZA (EXPERIMENTOS CON DOS FACTORES CON REPLICACIÓN)

Variación	Grados de libertad	Cuadrado medio (Estimador)	F
Entre tratamientos, V_R	$a - 1$	$\hat{S}_R^2 = \frac{V_R}{a - 1}$	$\hat{S}_R^2 / \hat{S}_E^2$ Con $a - 1$ y $ab(c - 1)$ grados de libertad
Entre bloques, V_C	$b - 1$	$\hat{S}_C^2 = \frac{V_C}{b - 1}$	$\hat{S}_C^2 / \hat{S}_E^2$ Con $b - 1$ y $ab(c - 1)$ grados de libertad
Interacción, V_I	$(a - 1)(b - 1)$	$\hat{S}_I^2 = \frac{V_I}{(a - 1)(b - 1)}$	$\hat{S}_I^2 / \hat{S}_E^2$ Con $(a - 1)(b - 1)$ y $ab(c - 1)$ grados de libertad
Residual o aleatoria, V_E	$a(c - 1)$	$\hat{S}_E^2 = \frac{V_E}{ab(c - 1)}$	
Total, V	$abc - 1$		

ANÁLISIS DE VARIANZA (EXPERIMENTOS CON DOS FACTORES CON REPLICACIÓN)

- DESDE UN PUNTO DE VISTA PRÁCTICO, HAY QUE DECIDIR PRIMERO SI $H_0^{(3)}$ PUEDE O NO SER RECHAZADA A NIVEL DE SIGNIFICANCIA APROPIADO USANDO EL F-COCIENTE $\hat{S}_I^2 / \hat{S}_E^2$:
 - **$H_0^{(3)}$ NO PUEDE SER RECHAZADA.** EN ESTE CASO SE CONCLUYE QUE LAS INTERACCIONES NO SON MUY GRANDES. ENTONCES SE PUEDEN PROBAR $H_0^{(1)}$ Y $H_0^{(2)}$ EMPLEANDO, RESPECTIVAMENTE LOS F-COCIENTES $\hat{S}_R^2 / \hat{S}_E^2$ Y $\hat{S}_C^2 / \hat{S}_E^2$.
 - **$H_0^{(3)}$ PUEDE SER RECHAZADA.** EN ESTE CASO SE CONCLUYE QUE LAS INTERACCIONES SON SIGNIFICATIVAMENTE GRANDES. ENTONCES, LAS DIFERENCIAS ENTRE LOS FACTORES SÓLO SERÁN IMPORTANTES SI SON GRANDES EN COMPARACIÓN CON ESTAS INTERACCIONES. A ESTO SE DEBE QUE SE RECOMIENDE PROBAR $H_0^{(1)}$ Y $H_0^{(2)}$ EMPLEANDO LOS F-COCIENTES $\hat{S}_R^2 / \hat{S}_I^2$ Y $\hat{S}_C^2 / \hat{S}_I^2$.

EJEMPLO

- LOS MIEMBROS DE UN EQUIPO CICLISTA SE DIVIDEN AL AZAR EN TRES GRUPOS QUE ENTRENAN CON MÉTODOS DIFERENTES. EL PRIMER GRUPO REALIZA LARGOS RECORRIDOS A RITMO PAUSADO, EL SEGUNDO GRUPO REALIZA SERIES CORTAS DE ALTA INTENSIDAD Y EL TERCERO TRABAJA EN EL GIMNASIO CON PESAS Y SE EJERCITA EN EL PEDALEO DE ALTA FRECUENCIA. DESPUÉS DE UN MES DE ENTRENAMIENTO SE REALIZA UN TEST DE RENDIMIENTO CONSISTENTE EN UN RECORRIDO CRONOMETRADO DE 9 KM. LOS TIEMPOS EMPLEADOS FUERON LOS SIGUIENTES:

Método I	Método II	Método III
15	14	13
16	13	12
14	15	11
15	16	14
17	14	11

- A UN NIVEL DE CONFIANZA DEL 95% ¿PUEDE CONSIDERARSE QUE LOS TRES MÉTODOS PRODUCEN RESULTADOS EQUIVALENTES? O POR EL CONTRARIO ¿HAY ALGÚN MÉTODO SUPERIOR A LOS DEMÁS?

SOLUCIÓN

Análisis de varianza de un factor

RESUMEN						
Grupos	Cuenta	Suma	Promedio	Varianza		
Columna 1	5	77	15,4	1,3		
Columna 2	5	72	14,4	1,3		
Columna 3	5	61	12,2	1,7		
ANÁLISIS DE VARIANZA						
Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	26,8	2	13,4	9,348837	0,00356	3,88529383
Dentro de los grupos	17,2	12	1,4333333			
Total	44	14				

El valor de la F teórica con 2 y 12 grados de libertad, a un nivel de confianza del 95% es 3,89. Por consiguiente se rechaza la hipótesis nula y se concluye que los tres métodos de entrenamiento producen diferencias significativas.