ANÁLISIS DE VARIANZA. EXPERIMENTOS DE UN SOLO FACTOR.

Experimentos de un solo factor.

Número de tratamientos: t.

Número de repeticiones por cada tratamiento: r_i

N = Número total de observaciones.

Modelo de los efectos.

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad \begin{cases} i = 1, 2, \ldots, t \\ j = 1, 2, \ldots, r \end{cases}$$

y_{ij} es la observación ij-ésima μ es un parámetro común a todos los tratamientos al que se le llama la media global, τ_i es un parámetro único del tratamiento i-ésimo al que se le llama el efecto del tratamiento i-ésimo y y_{ij} es un componente del error aleatorio que incorpora todas las demás fuentes de variabilidad del experimento, incluso las mediciones.

El interés se encuentra en probar la igualdad de las medias de los t tratamientos. Las hipótesis apropiadas son:

1.- Planteamiento de las hipótesis.

Hipótesis nula.

H_0: $\mu_1 = \mu_2 = \ldots = \mu_t$

Hipótesis alternativa.

H_1: $\mu_i \neq \mu_j$ para al menos un par (i, j).

2.- Nivel de significancia.

El nivel de significancia para la prueba es α (Una cola).

3.- Regla de decisión.

Diseño balanceado: No rechazar H_0 si $F < F_{\alpha,\, t-1,\, t(r-1)}$ Rechazar H_0 si $F > F_{\alpha,\, t-1,\, t(r-1)}$

Diseño no balanceado: No rechazar H_0 si $F < F_{\alpha,\, t-1,\, N-t}$ Rechazar H_0 si $F > F_{\alpha,\, t-1,\, N-t}$

4.- Estadístico de prueba.

$$F = \frac{CM_{Tratamientos}}{CM_E}$$

5.- Datos típicos de un experimento de un solo factor.

Tratamiento (nivel)	Observaciones				Totales ($y_{i.}$)	$y_{i.}^2$	Promedios ($\bar{y}_{i.}$)	$\sum_{j=1}^{r_i} y_{ij}^2$
1	y_{11}	y_{12}	...	y_{1r}	$y_{1.}$		$\bar{y}_{1.}$	
2	y_{21}	y_{22}	...	y_{2r}	$y_{2.}$		$\bar{y}_{2.}$	
⋮	⋮	⋮	...	⋮	⋮		⋮	
t	y_{t1}	y_{t2}	...	y_{tr}	$y_{t.}$		$\bar{y}_{t.}$	
					$y_{..}$	$\sum_{i=1}^{t} y_{i.}^2$	$\bar{y}_{..}$	$\sum_{i=1}^{t}\sum_{j=1}^{r_i} y_{ij}^2$

$y_{i.} = \sum_{j=1}^{r} y_{ij}$ Total de las observaciones bajo el tratamiento i-ésimo.

$\bar{y}_{i.} = \frac{y_{i.}}{r}$ Promedio de las observaciones bajo el tratamiento i-ésimo.

$y_{..} = \sum_{i=1}^{t}\sum_{j=1}^{r} y_{ij}$ Gran total de todas las observaciones.

$\bar{y}_{..} = \frac{y_{..}}{N}$ Gran promedio de todas las observaciones.

6.- Suma de cuadrados.

Total.

$$SC_T = \sum_{i=1}^{t}\sum_{j=1}^{r} (y_{ij} - \bar{y}_{..})^2 \qquad SC_T = \sum_{i=1}^{t}\sum_{j=1}^{r} y_{ij}^2 - \frac{y_{..}^2}{N}$$

Tratamientos.

$$SC_{Tratamientos} = r\sum_{i=1}^{t} (\bar{y}_{i.} - \bar{y}_{..})^2 \qquad SC_{Tratamientos} = \frac{1}{r}\sum_{i=1}^{t} y_{i.}^2 - \frac{y_{..}^2}{N}$$

Experimento no balanceado: $$SC_{Tratamientos} = \sum_{i=1}^{t} \frac{y_{i.}^2}{r_i} - \frac{y_{..}^2}{N}$$

Error:

$$SC_E = \sum_{i=1}^{t}\sum_{j=1}^{r} (y_{ij} - \bar{y}_{i.})^2 \qquad SC_E = SC_T - SC_{Tratamientos}$$

7.- Grados de libertad.

Diseño balanceado: Tratamientos: $t - 1$ Error: $t(r - 1)$ Total: $tr - 1$

Diseño no balanceado: Tratamientos: $t - 1$ Error: $N - t$ Total: $N - 1$

8.- Cuadrado medio.

Diseño balanceado.

Tratamientos: $$CM_{Tratamientos} = \frac{SC_{Tratamientos}}{t-1}$$ Error: $$CM_E = \frac{SC_E}{t(r-1)}$$

Diseño no balanceado.

Tratamientos: $$CM_{Tratamientos} = \frac{SC_{Tratamientos}}{t-1}$$ Error: $$CM_E = \frac{SC_E}{N-t}$$

9.- Tabla ANOVA.

Fuente de Variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	F
Tratamientos	$SC_{Tratamientos}$	$t - 1$	$CM_{Tratamientos}$	$CM_{Tratamientos}/CM_E$
Error	SC_E	$t(r - 1)$ ó $N - t$	CM_E	
Total	SC_T	$tr - 1$ ó $N - 1$		

10.- Coeficiente de determinación. $$R^2 = \frac{SC_{Tratamientos}}{SC_T}$$

11.- Coeficiente de variación. $$CV = \frac{\sqrt{CM_E}}{\bar{y}_{..}}$$

12.- Intervalos de confianza.

Para la media del i-ésimo tratamiento.

Diseño balanceado: $$\bar{y}_{i.} - t_{\alpha/2,\, t(r-1)}\sqrt{\frac{CM_E}{r}} \le \mu_i \le \bar{y}_{i.} + t_{\alpha/2,\, t(r-1)}\sqrt{\frac{CM_E}{r}}$$

Diseño no balanceado: $$\bar{y}_{i.} - t_{\alpha/2,\, N-t}\sqrt{\frac{CM_E}{r_i}} \le \mu_i \le \bar{y}_{i.} + t_{\alpha/2,\, N-t}\sqrt{\frac{CM_E}{r_i}}$$

Para la diferencia de las medias de dos tratamientos.

Diseño balanceado.

$$\bar{y}_{i.} - \bar{y}_{j.} - t_{\alpha/2,\, t(r-1)}\sqrt{\frac{2CM_E}{r}} \le \mu_i - \mu_j \le \bar{y}_{i.} - \bar{y}_{j.} + t_{\alpha/2,\, t(r-1)}\sqrt{\frac{2CM_E}{r}}$$

Diseño no balanceado.

$$\bar{y}_{i.} - \bar{y}_{j.} - t_{\alpha/2,\, N-t}\sqrt{CM_E\left(\frac{1}{r_i} + \frac{1}{r_j}\right)} \le \mu_i - \mu_j \le \bar{y}_{i.} - \bar{y}_{j.} + t_{\alpha/2,\, N-t}\sqrt{CM_E\left(\frac{1}{r_i} + \frac{1}{r_j}\right)}$$

13.- Pruebas sobre medias de tratamientos individuales.

Hipótesis nula.

H_0: $\mu_i = \mu_j$

Hipótesis alternativa.

H_1: $\mu_i \neq \mu_j$

14.- Diferencia mínima significativa (DMS) de Fisher.

Diseño balanceado.

$$DMS = t_{\alpha/2,\, t(r-1)}\sqrt{\frac{2CM_E}{r}}$$

Diseño no balanceado.

$$DMS_{i,j} = t_{\alpha/2,\, N-t}\sqrt{CM_E\left(\frac{1}{r_i} + \frac{1}{r_j}\right)}$$

Regla de decisión. Rechazar H_0 si $\left|\bar{y}_{i.} - \bar{y}_{j.}\right| > DMS$ No rechazar H_0 si $\left|\bar{y}_{i.} - \bar{y}_{j.}\right| < DMS$

15.- Prueba de Tukey.

Diseño balanceado.

$$T_{\alpha} = q_{\alpha,t,t(r-1)} \sqrt{\frac{CM_E}{r}}$$

Diseño no balanceado.

$$T_{\alpha,i,j} = \frac{q_{\alpha,t,N-t}}{\sqrt{2}} \sqrt{CM_E \left(\frac{1}{r_i} + \frac{1}{r_j} \right)}$$

Regla de decisión. Rechazar H_0 si $|y_{i.} - y_{j.}| > T_{\alpha}$ No rechazar H_0 si $|y_{i.} - y_{j.}| < T_{\alpha}$

Intervalo de confianza para la diferencia de las medias de dos tratamientos.

Diseño balanceado.

$$y_{i.} - y_{j.} - q_{\alpha,t,t(r-1)} \sqrt{\frac{CM_E}{r}} \le \mu_i - \mu_j \le y_{i.} - y_{j.} + q_{\alpha,t,t(r-1)} \sqrt{\frac{CM_E}{r}}$$

Diseño no balanceado.

$$y_{i.} - y_{j.} - \frac{q_{\alpha,t,N-t}}{\sqrt{2}} \sqrt{CM_E \left(\frac{1}{r_i} + \frac{1}{r_j} \right)} \le \mu_i - \mu_j \le y_{i.} - y_{j.} + \frac{q_{\alpha,t,N-t}}{\sqrt{2}} \sqrt{CM_E \left(\frac{1}{r_i} + \frac{1}{r_j} \right)}$$

16.- Comparación de medias de tratamientos con un control (Prueba de Dunnett).

Diseño balanceado.

$$D_{\alpha} = d_{\alpha,t-1,t(r-1)} \sqrt{\frac{2\,CM_E}{r}}$$

Diseño no balanceado.

$$D_{\alpha,i,j} = d_{\alpha,t-1,N-t} \sqrt{CM_E \left(\frac{1}{r_i} + \frac{1}{r_j} \right)}$$

Regla de decisión. Rechazar H_0 si $|y_{i.} - y_{j.}| > D_{\alpha}$ No rechazar H_0 si $|y_{i.} - y_{j.}| < D_{\alpha}$

17.- Prueba del rango múltiple de Duncan.

$$R_p = r_{\alpha,p,N-t} \sqrt{\frac{CM_E}{r_h}} \quad \text{para} \quad p = 2,3,\ldots,t$$

Diseño balanceado: $r_h = r$ Diseño no balanceado: $r_h = \dfrac{t}{\sum_{i=1}^{t} \frac{1}{r_i}}$

Regla de decisión. Rechazar H_0 si $|y_{i.} - y_{j.}| > R_p$ No rechazar H_0 si $|y_{i.} - y_{j.}| < R_p$

18. Contrastes.

En general, un contraste es una combinación lineal de parámetros de la forma $\Gamma = \sum_{i=1}^{t} c_i \mu_i$ donde las constantes de los contrastes $c_1, c_2, \ldots, c_a$ suma cero; es decir, $\sum_{i=1}^{t} c_i = 0$. Las dos hipótesis pueden expresarse en términos de contrastes:

H_0: $\sum_{i=1}^{t} c_i \mu_i = 0$ H_1: $\sum_{i=1}^{t} c_i \mu_i \neq 0$.

Las pruebas de hipótesis que incluyen contrastes pueden hacerse de dos maneras básicas.

En el primer método se utiliza la prueba *t*. El contraste de interés se escribe en términos de los totales de los tratamientos, obteniéndose $C = \sum_{i=1}^{t} c_i y_{i\cdot}$.

El estadístico de prueba es

Diseño balanceado.

$$t_0 = \frac{\sum_{i=1}^{t} c_i y_{i\cdot}}{\sqrt{r\,CM_E \sum_{i=1}^{t} c_i^2}}$$

Diseño no balanceado.

$$t_0 = \frac{\sum_{i=1}^{t} c_i y_{i\cdot}}{\sqrt{CM_E \sum_{i=1}^{t} r_i c_i^2}}$$

Regla de decisión. Rechazar H_0 si $|t_0| > t_{\alpha/2,N-t}$ No rechazar H_0 si $|t_0| < t_{\alpha/2,N-t}$

En el segundo enfoque se utiliza la prueba *F*.

El estadístico de prueba es $F_0 = t_0^2 = \dfrac{\left(\sum_{i=1}^{t} c_i y_{i\cdot} \right)^2}{r\,CM_E \sum_{i=1}^{t} c_i^2}$.

El estadístico de prueba puede escribirse como $F_0 = \dfrac{CM_C}{CM_E}$, donde la suma de cuadrados de los contrastes con un solo grado de libertad es:

Diseño balanceado.

$$SC_C = \frac{\left(\sum_{i=1}^{t} c_i y_{i\cdot} \right)^2}{r \sum_{i=1}^{t} c_i^2}$$

Diseño no balanceado.

$$SC_C = \frac{\left(\sum_{i=1}^{t} c_i y_{i\cdot} \right)^2}{\sum_{i=1}^{t} r_i c_i^2}$$

Regla de decisión. Rechazar H_0 si $|F_0| > F_{\alpha,1,N-t}$ No rechazar H_0 si $|F_0| < F_{\alpha,1,N-t}$

Intervalo de confianza para un contraste.

El intervalo de confianza de $(1-\alpha)$% para el contraste $\sum_{i=1}^{t} c_i \mu_i$ es

$$\sum_{i=1}^{t} c_i y_{i\cdot} - t_{\alpha/2,N-t} \sqrt{\frac{CM_E}{r} \sum_{i=1}^{t} c_i^2} \le \sum_{i=1}^{t} c_i \mu_i \le \sum_{i=1}^{t} c_i y_{i\cdot} + t_{\alpha/2,N-t} \sqrt{\frac{CM_E}{r} \sum_{i=1}^{t} c_i^2}$$

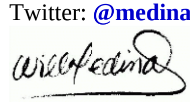
Contrastes ortogonales.

Dos contrastes con coeficientes $\{c_i\}$ y $\{d_i\}$ son ortogonales si $\sum_{i=1}^{t} c_i d_i = 0$ ó para un diseño no balanceado si $\sum_{i=1}^{t} r_i c_i d_i = 0$.

Para *t* tratamientos, el conjunto de $t-1$ contrastes ortogonales hace la partición de la suma de cuadrados debido a los tratamientos en $t-1$ componentes independientes con un solo grado de libertad. Por lo tanto, las pruebas que se realizan en los contrastes ortogonales son independientes.

En general, el método de contrastes (o de contrastes ortogonales) es útil para lo que se llama comparaciones preplaneadas. Es decir, los contrastes se especifican antes de llevar a cabo el experimento y de examinar los datos.

Autor: **MSc. Ing. Willians Medina**.
Teléfono / WhatsApp: **+58-424-9744352**
e-mail: **medinawj@gmail.com**
Twitter: **@medinawj**



El presente formulario está disponible en formato digital en la siguiente dirección:
https://www.tutoruniversitario.com/
Puerto La Cruz, abril de 2025.