Diseño y Análisis de Experimentos

Johan Steven Aparicio

Escuela de Estadística
Facultad de Ingeniería - Universidad del Valle
johan.aparicio@correounivalle.edu.co

Mayo, 2022

Tabla de Contenido

$$E[SCTrat] = E\left[\sum_{i=1}^{t}(\overline{Y}_{i.} - \overline{Y}_{..})^2 r_i\right] = \sum_{i=1}^{t}(\tau_i - \overline{\tau})^2 r_i + (t-1)\sigma^2$$

como,

$$CMTrat = \frac{SCTrat}{t-1}$$

entonces,

$$E[CMTrat] = \frac{E[SCTrat]}{t-1} = \frac{\sum_{i=1}^{t}(\tau_i - \overline{\tau})^2 r_i}{t-1} + \sigma^2$$

además, se tiene que:

$$E[CMError] = \frac{E[SCError]}{n-t} = \sigma^2$$

Demostrar

$$SCTrat = \sum_{i=1}^{t}(\overline{Y}_{i.} - \overline{Y}_{..})^2 r_i$$

$$\overline{Y}_{i.} = \sum_{j=1}^{r_i} \frac{Y_{ij}}{r_i} = \sum_{j=1}^{r_i} \frac{\mu + \tau_i + \varepsilon_{ij}}{r_i} = \mu + \tau_i + \overline{\varepsilon}_{i.}$$

$$\overline{Y}_{..} = \sum_{i=1}^{t}\sum_{j=1}^{n} \frac{Y_{ij}}{r_i} = \sum_{i=1}^{t}\sum_{j=1}^{r_i} \frac{\mu + \tau_i + \varepsilon_{ij}}{n} = \mu + \overline{\tau} + \overline{\varepsilon}_{..}$$

Por tanto,

$$E[SCTrat] = E\left[\sum_{i=1}^{t}(\overline{Y}_{i.} - \overline{Y}_{..})^2 r_i\right]$$

$$= E\left[\sum_{i=1}^{t}(\overline{Y}_{i.} - \overline{Y}_{..})^2 r_i\right] = E\left[\sum_{i=1}^{t}(\mu + \tau_i + \overline{\varepsilon}_{i.} - \mu - \overline{\tau} - \overline{\varepsilon}_{..})^2 r_i\right]$$

$$= E\left[\sum_{i=1}^{t}((\tau_i - \overline{\tau}) + (\overline{\varepsilon}_{i.} - \overline{\varepsilon}_{..}))^2 r_i\right]$$

$$= E\left[\sum_{i=1}^{t}((\tau_i - \overline{\tau})^2 r_i + (\overline{\varepsilon}_{i.} - \overline{\varepsilon}_{..})^2 r_i + 2(\tau_i - \overline{\tau})(\overline{\varepsilon}_{i.} - \overline{\varepsilon}_{..}) r_i)\right]$$

$$= E\left[\sum_{i=1}^{t}(\tau_i - \overline{\tau})^2 r_i\right] + E\left[\sum_{i=1}^{t}(\overline{\varepsilon}_{i.} - \overline{\varepsilon}_{..})^2 r_i\right] + 2E\left[\sum_{i=1}^{t}(\tau_i - \overline{\tau})(\overline{\varepsilon}_{i.} - \overline{\varepsilon}_{..}) r_i\right]$$

$$= \sum_{i=1}^{t}(\tau_i - \overline{\tau})^2 r_i + (t-1)\sigma^2$$

Tabla de análisis de varianza para el modelo con un solo factor

Fuente de Variabilidad	G.L	Suma de Cuadrados	Cuadrado Medio	F_0	E(Cuadrado Medio)
τ	$t-1$	$\sum_{i=1}^{t}\frac{y_{i.}^2}{r_i}-\frac{y_{..}^2}{n}$	$\frac{SC_\tau}{t-1}$	$\frac{CM_\tau}{CM_{Error}}$	$\frac{\sum_{i=1}^{t}(\tau_i-\bar{\tau})^2 r_i}{t-1}+\sigma^2$
Error	$n-t$	$\sum_{i=1}^{t}\sum_{j=1}^{r_i}y_{ij}^2-\sum_{i=1}^{t}\frac{y_{i.}^2}{r_i}$	$\frac{SC_{Error}}{n-t}$		σ^2
Total	$n-1$	$\sum_{i=1}^{t}\sum_{j=1}^{r_i}y_{ij}^2-\frac{y_{..}^2}{n}$			

Que se puede decir acerca de $E(CM)$ bajo H_0?

- Si H_0 es cierta, entonces $E[CMTrat] = \sigma^2$ ya que $\tau_1 = \tau_2 = \cdots = \tau_t = 0$; en tal caso $F \cong 1$.
- Si $F > 1$, indica que hay dispersión significativa entre los τ_i, posiblemente existen diferencias entre los efectos que producen los tratamientos en la variable respuesta.

Sea

$$V = \frac{SCTrat}{\sigma^2} \sim \chi^2_{(t-1,\lambda)}$$

Sea

$$V = \frac{SCTrat}{\sigma^2} \sim \chi^2_{(t-1,\lambda)} \quad , \quad E\left[\chi^2_{(t-1,\lambda)}\right] = (t-1) + \lambda$$

Sea

$$V = \frac{SCTrat}{\sigma^2} \sim \chi^2_{(t-1,\lambda)} \quad , \quad E\left[\chi^2_{(t-1,\lambda)}\right] = (t-1) + \lambda$$

$$E\left[V\right] = \frac{1}{\sigma^2} E\left[SCTrat\right] =$$

Sea

$$V = \frac{SCTrat}{\sigma^2} \sim \chi^2_{(t-1,\lambda)} \quad , \quad E\left[\chi^2_{(t-1,\lambda)}\right] = (t-1) + \lambda$$

$$E\left[V\right] = \frac{1}{\sigma^2} E\left[SCTrat\right] = (t-1) + \lambda$$

Sea

$$V = \frac{SCTrat}{\sigma^2} \sim \chi^2_{(t-1,\lambda)} \quad , \quad E\left[\chi^2_{(t-1,\lambda)}\right] = (t-1) + \lambda$$

$$E\left[V\right] = \frac{1}{\sigma^2} E\left[SCTrat\right] = (t-1) + \lambda$$

donde,

$$\lambda = \frac{E\left[SCTrat\right]}{\sigma^2} - (t-1)$$

Sea

$$V = \frac{SCTrat}{\sigma^2} \sim \chi^2_{(t-1,\lambda)} \quad , \quad E\left[\chi^2_{(t-1,\lambda)}\right] = (t-1) + \lambda$$

$$E\left[V\right] = \frac{1}{\sigma^2} E\left[SCTrat\right] = (t-1) + \lambda$$

donde,

$$\lambda = \frac{E\left[SCTrat\right]}{\sigma^2} - (t-1)$$

$$\lambda = \frac{\sum_{i=1}^{t} (\tau_i - \overline{\tau})^2 r_i + (t-1)\sigma^2}{\sigma^2} - (t-1)$$

Sea

$$V = \frac{SCTrat}{\sigma^2} \sim \chi^2_{(t-1,\lambda)} \quad , \quad E\left[\chi^2_{(t-1,\lambda)}\right] = (t-1) + \lambda$$

$$E\left[V\right] = \frac{1}{\sigma^2} E\left[SCTrat\right] = (t-1) + \lambda$$

donde,

$$\lambda = \frac{E\left[SCTrat\right]}{\sigma^2} - (t-1)$$

$$\lambda = \frac{\sum_{i=1}^{t} (\tau_i - \overline{\tau})^2 r_i + (t-1)\sigma^2}{\sigma^2} - (t-1)$$

$$\lambda = \frac{\sum_{i=1}^{t} (\tau_i - \overline{\tau})^2 r_i}{\sigma^2}$$

Finalmente,

$$\lambda = \frac{\sum_{i=1}^{t} (\tau_i - \overline{\tau})^2 r_i}{\sigma^2}$$

λ se conoce como parámetro de no centralidad y $\lambda = 0$ bajo H_0.

Finalmente,

$$\lambda = \frac{\sum_{i=1}^{t}(\tau_i - \overline{\tau})^2 r_i}{\sigma^2}$$

λ se conoce como parámetro de no centralidad y $\lambda = 0$ bajo H_0.

La potencia de la prueba se puede determinar a través de la estimación de λ. De igual forma se puede prefijar la potencia con el fin de determinar los r_i (tamaño del experimento).

En un DCA, el investigador está interesado en determinar el número de réplicas que le permitan al experimento detectar diferencias significativas entre los tratamientos, es decir, para determinar si hay o no evidencia para que con la prueba F del análisis de varianza se rechace o no la hipótesis nula.

La técnica ofrecida por la teoría estadística para decidir sobre el número de repeticiones necesarias en un experimento, es el cálculo de la potencia de las pruebas estadísticas de interés.

Para determinar el número de replicaciones por tratamiento (r), lo primero que se debe hacer es predeterminar a λ:

$$\lambda = \frac{\sum_{i=1}^{t} (\tau_i - \overline{\tau})^2 r_i}{\sigma^2}$$

El investigador debe tener claro, cuando estudia diferencia de efectos de tratamientos, qué magnitud es significativa en la práctica.

Errores y Potencia en las Pruebas de Hipótesis:

- Un error tipo I ocurre cuando H_0 is verdadera pero es rechazada.

Errores y Potencia en las Pruebas de Hipótesis:

- Un error tipo I ocurre cuando H_0 is verdadera pero es rechazada.
- Un error tipo II ocurre cuando H_0 es falsa pero es aceptada.

Errores y Potencia en las Pruebas de Hipótesis:

- Un error tipo I ocurre cuando H_0 is verdadera pero es rechazada.
- Un error tipo II ocurre cuando H_0 es falsa pero es aceptada.
- El nivel de significancia de una prueba es la probabilidad de cometer un error tipo I, es decir, la probabilidad de rechazar H_0 cuando esta es verdadera.

Errores y Potencia en las Pruebas de Hipótesis:

- Un error tipo I ocurre cuando H_0 is verdadera pero es rechazada.
- Un error tipo II ocurre cuando H_0 es falsa pero es aceptada.
- El nivel de significancia de una prueba es la probabilidad de cometer un error tipo I, es decir, la probabilidad de rechazar H_0 cuando esta es verdadera.
- La potencia de la prueba es la probabilidad de rechazar H_0 cuando H_a es verdadera:

Errores y Potencia en las Pruebas de Hipótesis:

- Un error tipo I ocurre cuando H_0 is verdadera pero es rechazada.
- Un error tipo II ocurre cuando H_0 es falsa pero es aceptada.
- El nivel de significancia de una prueba es la probabilidad de cometer un error tipo I, es decir, la probabilidad de rechazar H_0 cuando esta es verdadera.
- La potencia de la prueba es la probabilidad de rechazar H_0 cuando H_a es verdadera:

  $$\text{potencia} = 1 \text{ - P(cometer error tipo II}|H_0 \text{ es falsa)} = 1 - \beta$$

- Un buen test debe tener un nivel de prueba pequeño y una gran potencia.

- Un buen test debe tener un nivel de prueba pequeño y una gran potencia.

	H_a is rejected (H_0 is accepted)	H_0 is rejected (H_a is accepted)
H_0 is true	$\surd$	$\alpha = \mathrm{P}(\text{Type } I \text{ error})$
H_0 is false	$\beta = \mathrm{P}(\text{Type } II \text{ error})$	$\surd$

Recordando las hipótesis establecidas para el modelo de los efectos:

$$H_0 : \tau_1 = \cdots = \tau_t = 0 \quad \text{vs} \quad H_a : \text{No todos los } \tau_i \text{ son } 0.$$

Recordando las hipótesis establecidas para el modelo de los efectos:

$$H_0 : \tau_1 = \cdots = \tau_t = 0 \quad \text{vs} \quad H_a : \text{No todos los } \tau_i \text{ son } 0.$$

Rechazamos H_0 a un nivel α si el estadístico $F = \frac{CM_{trat}}{CM_{Error}}$ excede el valor crítico $F_{1-\alpha,t-1,n-t}$. Así que:

Recordando las hipótesis establecidas para el modelo de los efectos:

$$H_0 : \tau_1 = \cdots = \tau_t = 0 \quad \text{vs} \quad H_a : \text{No todos los } \tau_i \text{ son } 0.$$

Rechazamos H_0 a un nivel α si el estadístico $F = \frac{CM_{trat}}{CM_{Error}}$ excede el valor crítico $F_{1-\alpha,t-1,n-t}$. Así que:

$$\text{Potencia} = P(\text{Rechazar } H_0 | H_a \text{ es verdadera}) = P(F > F_{1-\alpha,t-1,n-t})$$

Recordando las hipótesis establecidas para el modelo de los efectos:

$$H_0 : \tau_1 = \cdots = \tau_t = 0 \quad \text{vs} \quad H_a : \text{No todos los } \tau_i \text{ son } 0.$$

Rechazamos H_0 a un nivel α si el estadístico $F = \frac{CM_{trat}}{CM_{Error}}$ excede el valor crítico $F_{1-\alpha,t-1,n-t}$. Así que:

$$\text{Potencia} = P(\text{Rechazar } H_0 | H_a \text{ es verdadera}) = P(F > F_{1-\alpha,t-1,n-t})$$

Para encontrar $P(F > F_{1-\alpha,t-1,n-t})$, debemos conocer la distribución de F.

- Cuál es la distribución de F bajo H_0? $F_{t-1,n-t}$
- Y bajo H_a?

Para $F = \frac{CM_{trat}}{CM_{Error}} = \frac{SC_{trat}/(t-1)}{SC_{Error}/(n-t)}$,

Para $F = \frac{CM_{trat}}{CM_{Error}} = \frac{SC_{trat}/(t-1)}{SC_{Error}/(n-t)}$,

- no importa bajo H_0 o H_a, siempre es cierto que:

Para $F = \frac{CM_{trat}}{CM_{Error}} = \frac{SC_{trat}/(t-1)}{SC_{Error}/(n-t)}$,

- no importa bajo H_0 o H_a, siempre es cierto que:

$$\frac{SC_{Error}}{\sigma^2} \sim \chi^2_{n-t}, \quad \text{y} \quad E(CM_{Error}) = \sigma^2$$

Para $F = \frac{CM_{trat}}{CM_{Error}} = \frac{SC_{trat}/(t-1)}{SC_{Error}/(n-t)}$,

- no importa bajo H_0 o H_a, siempre es cierto que:

$$\frac{SC_{Error}}{\sigma^2} \sim \chi^2_{n-t}, \quad \text{y} \quad E(CM_{Error}) = \sigma^2$$

- Para el numerador, SC_{trat}/σ^2, bajo H_0, tiene una distribución χ^2_{t-1}. Pero bajo H_a tiene una distribución chi-cuadrado no-central $\chi^2_{t-1,\lambda}$ con parámetro de no centralidad

Para $F = \frac{CM_{trat}}{CM_{Error}} = \frac{SC_{trat}/(t-1)}{SC_{Error}/(n-t)}$,

- no importa bajo H_0 o H_a, siempre es cierto que:

$$\frac{SC_{Error}}{\sigma^2} \sim \chi^2_{n-t}, \quad \text{y} \quad E(CM_{Error}) = \sigma^2$$

- Para el numerador, SC_{trat}/σ^2, bajo H_0, tiene una distribución χ^2_{t-1}. Pero bajo H_a tiene una distribución chi-cuadrado no-central $\chi^2_{t-1,\lambda}$ con parámetro de no centralidad

$$\lambda = \frac{\sum_{i=1}^{t} \tau_i^2 r_i}{\sigma^2}$$

$$E(SC_{trat}) = \begin{cases} (t-1)\sigma^2 & \text{bajo } H_0 \\ (t-1+\lambda)\sigma^2 & \text{bajo } H_a \end{cases}$$

Bajo H_a, se puede demostrar que

$$F = \frac{CM_{trat}}{CM_{Error}}$$

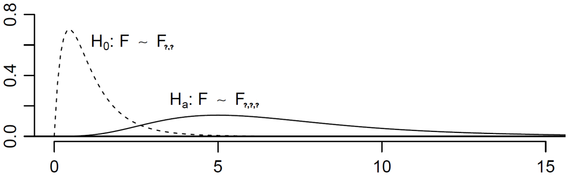
tiene distribución F no-central con $t-1$ y $n-t$ grados de libertad, con parámetro de no-centralidad λ, denotado como:

$$F \sim F_{t-1,n-t,\lambda} \quad \text{donde} \quad \lambda = \frac{\sum_{i=1}^{t} \tau_i^2 r_i}{\sigma^2}$$

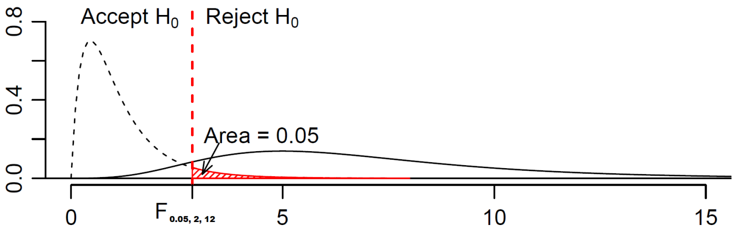
Ejercicio 1

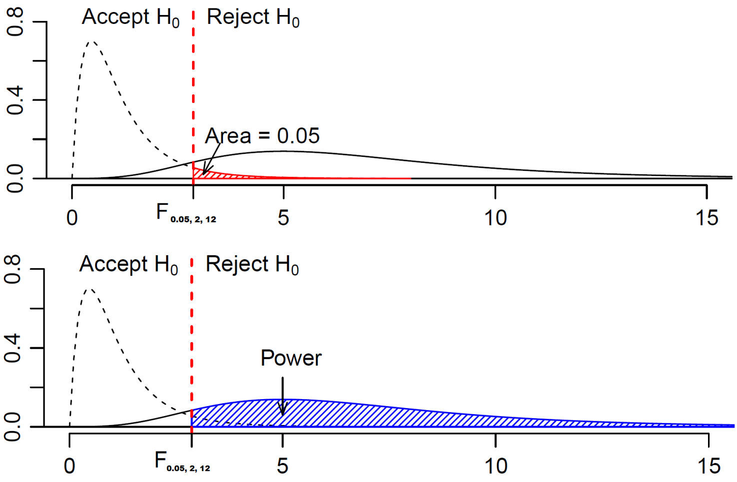
Realizar el cálculo de la potencia para el experimento de los fertilizantes.

$$F = \frac{CM_{trat}}{CM_{Error}} \sim \begin{cases} F_{t-1,n-t} & \text{bajo } H_0 \\ F_{t-1,n-t,\lambda} & \text{bajo } H_a \end{cases}$$



Distribución F no-central





Y el tamaño de muestra?

Se tiene que,

$$\overline{Y}_{i.} = \mu + \tau_i + \overline{\varepsilon}_{i.}$$

Se tiene que,

$$\overline{Y}_{i.} = \mu + \tau_i + \overline{\varepsilon}_{i.}$$

- $E[\overline{Y}_{i.}] = \mu + \tau_i$
- $Var[\overline{Y}_{i.}] = \sigma^2/r_i$

Se tiene que,

$$\overline{Y}_{i.} = \mu + \tau_i + \overline{\varepsilon}_{i.}$$

- $E[\overline{Y}_{i.}] = \mu + \tau_i$
- $Var[\overline{Y}_{i.}] = \sigma^2/r_i$
- $\overline{Y}_{i.} \sim N(\mu + \tau_i, \sigma^2/r_i)$

Se tiene que,

$$\overline{Y}_{i.} = \mu + \tau_i + \overline{\varepsilon}_{i.}$$

- $E[\overline{Y}_{i.}] = \mu + \tau_i$
- $Var[\overline{Y}_{i.}] = \sigma^2/r_i$
- $\overline{Y}_{i.} \sim N(\mu + \tau_i, \sigma^2/r_i)$

y por otro lado,

$$\overline{Y}_{..} = \mu + \overline{\tau} + \overline{\varepsilon}_{..}$$

Se tiene que,

$$\overline{Y}_{i.} = \mu + \tau_i + \overline{\varepsilon}_{i.}$$

- $E[\overline{Y}_{i.}] = \mu + \tau_i$
- $Var[\overline{Y}_{i.}] = \sigma^2/r_i$
- $\overline{Y}_{i.} \sim N(\mu + \tau_i, \sigma^2/r_i)$

y por otro lado,

$$\overline{Y}_{..} = \mu + \overline{\tau} + \overline{\varepsilon}_{..}$$

- $E[\overline{Y}_{..}] = \mu + \overline{\tau}$
- $Var[\overline{Y}_{..}] = \sigma^2/n$

Se tiene que,

$$\overline{Y}_{i.} = \mu + \tau_i + \overline{\varepsilon}_{i.}$$

- $E[\overline{Y}_{i.}] = \mu + \tau_i$
- $Var[\overline{Y}_{i.}] = \sigma^2/r_i$
- $\overline{Y}_{i.} \sim N(\mu + \tau_i, \sigma^2/r_i)$

y por otro lado,

$$\overline{Y}_{..} = \mu + \overline{\tau} + \overline{\varepsilon}_{..}$$

- $E[\overline{Y}_{..}] = \mu + \overline{\tau}$
- $Var[\overline{Y}_{..}] = \sigma^2/n$
- $\overline{Y}_{..} \sim N(\mu + \overline{\tau}, \sigma^2/n)$

Dado lo anterior y que no conocemos σ^2

$$t = \frac{\overline{Y}_{i.} - \mu_i}{\hat{\sigma}^2/\sqrt{r_i}} \sim t_{n-t}$$

Dado lo anterior y que no conocemos σ^2

$$t = \frac{\overline{Y}_{i.} - \mu_i}{\hat{\sigma}^2/\sqrt{r_i}} \sim t_{n-t}$$

el intervalo de confianza para $\mu + \tau_i$ viene dado por:

$$\overline{Y}_{i.} \pm t_{\alpha/2,n-t}\hat{\sigma}^2/\sqrt{r_i}$$

Dado lo anterior y que no conocemos σ^2

$$t = \frac{\overline{Y}_{i.} - \mu_i}{\hat{\sigma}^2/\sqrt{r_i}} \sim t_{n-t}$$

el intervalo de confianza para $\mu + \tau_i$ viene dado por:

$$\overline{Y}_{i.} \pm t_{\alpha/2,n-t}\hat{\sigma}^2/\sqrt{r_i}$$

y para $\mu + \overline{\tau}$ se tiene:

$$\overline{Y}_{..} \pm t_{\alpha/2,n-t}\hat{\sigma}^2/\sqrt{n}$$

Si en la tabla ANOVA se rechaza H_0, lo que indica este rechazo es que existen diferencias significativas entre al menos dos tratamientos, entonces se debe proceder a analizar entre cuales tratamientos existen esas diferencias y si esas diferencias en la práctica son de importancia.

Si en la tabla ANOVA se rechaza H_0, lo que indica este rechazo es que existen diferencias significativas entre al menos dos tratamientos, entonces se debe proceder a analizar entre cuales tratamientos existen esas diferencias y si esas diferencias en la práctica son de importancia.

Algunos métodos de análisis son:

- Método de Contrastes Ortogonales
- Método de Tukey
- Método LSD
- Método de Scheffe
- Método de Rango Múltiple de Duncan
- Método de Bonferroni

Least Significant Difference (LSD)

$$H_0 : \mu_i = \mu_j \quad \text{vs} \quad H_a : \mu_i \neq \mu_j \quad \forall i \neq j$$

Estadístico de prueba:

$$|\bar{y}_{i\bullet} - \bar{y}_{j\bullet}|$$

$$LSD = t_{\alpha/2,n-t}\sqrt{CM_{Error}\left(\frac{1}{r_i} + \frac{1}{r_j}\right)} = t_{\alpha/2,n-t}\sqrt{\hat{\sigma}^2\left(\frac{1}{r_i} + \frac{1}{r_j}\right)}$$

Asi que, si μ_i y μ_j son significativamente diferentes al nivel α

$$|\bar{y}_{i\bullet} - \bar{y}_{j\bullet}| > LSD$$

Contrastes Ortogonales

Consiste en hacer una prueba F a cada uno de los $t-1$ contrastes contenidos en H_0

Si K_1 y K_2 son contrastes, entonces:

$$K_1 = \sum_{i=1}^{t} a_i \tau_i \quad , \quad \sum_{i=1}^{t} a_i = 0$$

Contrastes Ortogonales

Consiste en hacer una prueba F a cada uno de los $t-1$ contrastes contenidos en H_0

Si K_1 y K_2 son contrastes, entonces:

$$K_1 = \sum_{i=1}^{t} a_i \tau_i \quad , \quad \sum_{i=1}^{t} a_i = 0$$

$$K_2 = \sum_{i=1}^{t} c_i \tau_i \quad , \quad \sum_{i=1}^{t} a_i = 0$$

y son ortogonales sii $\sum a_i c_i = 0$

Recordemos que

$$\begin{aligned}
\hat{\mu} &= \bar{Y}_{..} \\
\hat{\tau}_1 &= \bar{Y}_{1.} - \bar{Y}_{..} \\
&\vdots \\
\hat{\tau}_t &= \bar{Y}_{t.} - \bar{Y}_{..}
\end{aligned}$$

Recordemos que

$$\begin{aligned}
\hat{\mu} &= \bar{Y}_{..} \\
\hat{\tau}_1 &= \bar{Y}_{1.} - \bar{Y}_{..} \\
&\vdots \\
\hat{\tau}_t &= \bar{Y}_{t.} - \bar{Y}_{..}
\end{aligned}$$

entonces,

$$\hat{K} = \sum_{i=1}^{t} a_i \hat{\tau}_i = \sum_{i=1}^{t} a_i (\bar{Y}_{i.} - \bar{Y}_{..}) =$$

Recordemos que

$$\begin{aligned}\hat{\mu} &= \bar{Y}_{..} \\ \hat{\tau}_1 &= \bar{Y}_{1.} - \bar{Y}_{..} \\ &\vdots \\ \hat{\tau}_t &= \bar{Y}_{t.} - \bar{Y}_{..}\end{aligned}$$

entonces,

$$\hat{K} = \sum_{i=1}^{t} a_i \hat{\tau}_i = \sum_{i=1}^{t} a_i(\bar{Y}_{i.} - \bar{Y}_{..}) = \sum_{i=1}^{t} a_i \bar{Y}_{i.}$$

Recordemos que

$$\begin{aligned}\hat{\mu} &= \bar{Y}_{..}\\ \hat{\tau}_1 &= \bar{Y}_{1.} - \bar{Y}_{..}\\ &\vdots\\ \hat{\tau}_t &= \bar{Y}_{t.} - \bar{Y}_{..}\end{aligned}$$

entonces,

$$\hat{K} = \sum_{i=1}^{t} a_i \hat{\tau}_i = \sum_{i=1}^{t} a_i(\bar{Y}_{i.} - \bar{Y}_{..}) = \sum_{i=1}^{t} a_i \bar{Y}_{i.}$$

donde,

$$E[\hat{K}] = \sum_{i=1}^{t} a_i \tau_i \quad \text{y} \quad V[\hat{K}] = \sum_{i=1}^{t} a_i^2 \frac{\sigma^2}{r_i}$$

Entonces,

$$\hat{K} = \sum_{i=1}^{t} a_i \bar{Y}_{i.} \sim N\left(\sum_{i=1}^{t} a_i \tau_i, \quad \sigma^2 \sum_{i=1}^{t} a_i^2 / r_i\right)$$

Estadarizando se tiene que:

$$Z_k = \frac{\sum a_i \bar{Y}_{i.} - \sum a_i \tau_i}{(\sigma^2 \sum a_i^2 / r_i)^{1/2}} \sim N(0,1)$$

Entonces,

$$\hat{K} = \sum_{i=1}^{t} a_i \bar{Y}_{i.} \sim N\left(\sum_{i=1}^{t} a_i \tau_i, \quad \sigma^2 \sum_{i=1}^{t} a_i^2 / r_i\right)$$

Estadarizando se tiene que:

$$Z_k = \frac{\sum a_i \bar{Y}_{i.} - \sum a_i \tau_i}{(\sigma^2 \sum a_i^2 / r_i)^{1/2}} \sim N(0,1)$$

Sea,

$$W = Z_k^2 = \frac{(\sum a_i \bar{Y}_{i.} - \sum a_i \tau_i)^2}{\sigma^2 \sum a_i^2 / r_i} \sim \chi^2_{(1,\lambda)}$$

Entonces,

$$\hat{K} = \sum_{i=1}^{t} a_i \bar{Y}_{i.} \sim N\left(\sum_{i=1}^{t} a_i \tau_i, \quad \sigma^2 \sum_{i=1}^{t} a_i^2 / r_i\right)$$

Estadarizando se tiene que:

$$Z_k = \frac{\sum a_i \bar{Y}_{i.} - \sum a_i \tau_i}{(\sigma^2 \sum a_i^2 / r_i)^{1/2}} \sim N(0,1)$$

Sea,

$$W = Z_k^2 = \frac{(\sum a_i \bar{Y}_{i.} - \sum a_i \tau_i)^2}{\sigma^2 \sum a_i^2 / r_i} \sim \chi^2_{(1,\lambda)}$$

Si $\sum a_i \tau_i \neq 0$, entonces $\lambda \neq 0$ y se tiene una chi-cuadrado no central.

Recordemos que:

$$V = \frac{SC_{Error}}{\sigma^2} = \frac{\hat{\sigma}^2(n-t)}{\sigma^2} \sim \chi^2_{(n-t)}$$

Así que,

$$F = \frac{W/1}{V/(n-t)} = \frac{(\sum a_i \bar{Y}_{i.} - \sum a_i \tau_i)^2}{\hat{\sigma}^2 \sum a_i^2/r_i} \sim F_{(1,n-t,\lambda)}$$

Y bajo H_0 se tiene que:

$$F = \frac{(\sum a_i \bar{Y}_{i.})^2}{\hat{\sigma}^2 \sum a_i^2/r_i} = \frac{1}{\hat{\sigma}^2} SC_{K_j}$$

Finalmente, la SC_{trat} se puede descomponer en $t-1$ sumas de cuadrados correspondientes a $t-1$ contrastes ortogonales, con cada una de las cuales se puede hacer una prueba F independiente de las demás.

$$\sum_{j=1}^{t-1} SC_{K_j} = SC_{trat}$$

Para el ejemplo de los fertilizantes, se desea probar las siguientes hipótesis:

$$H_0 : \tau_1 - \tau_2 = 0 \quad \text{y} \quad H_0 : \tau_1 + \tau_2 - 2\tau_3 = 0$$

Realice los contrastes y decida acerca de las hipótesis planteadas.