Universidad Autónoma de Guerrero

Maestría en Matemáticas Aplicadas

Diseño Aleatorio Simple

Autor: Carlos David Carretillo Moctezuma

Facultad de Matemáticas

karre.97.moctezuma@hotmail.com

Contenido

Introducción

En una amplia variedad de disciplinas académicas, los investigadores realizan experimentos con el propósito de obtener aspectos de procesos o sistemas específicos. Desde una perspectiva literal, un experimento representa una evaluación o prueba. Desde un enfoque más formal, se puede definir como una serie de pruebas en las que se introducen cambios controlados en las variables de entrada de un proceso o sistema con el fin de observar y comprender las causas subyacentes de las posibles modificaciones observadas en las variables de salida (Motgomery C. D. 2005).

Conceptos básicos

1.¿Qué es el diseño de experimentos?

Consiste en manipular intencionalmente la variable independiente de un modelo para observar y medir sus efectos en la variable dependiente.

2. Cómo se debe realizar la alatorización en un DCA?

Se puede asignar números a los participantes, o a los tratamientos, y utilizar una tabla de números aleatorios para elegir los participantes y los grupos de tratamiento.

3. Explique los tres principios básicos del diseño de experimentos

El principio de aleatorización, bloqueo, la factorización del diseño.

4. ¿Qué es un factor?

Son las variables independientes que pueden influir en la variabilidad de la variable de interés.

5. ¿Qué es un tratamiento?

Es una combinación específica de los niveles de los factores en estudio. Son, por tanto, las condiciones experimentales que se desean comparar en el experimento.

6. ¿Qué es una unidad experimental?

Es la entidad o elemento sobre la cual se realizan las observaciones y las mediciones.

6. ¿Qué es un diseño de efectos fijos?

Es aquel en el que el investigador elije los niveles del factor que se va a estudiar.

7. ¿Qué es un diseño de efectos aleatorios?

Es aquel en donde el investigador elije los niveles del factor de interés de forma aleatoria.

8. ¿Qué es un diseño de efectos mixtos?

Es aquel en el que hay factores de efectos fijos y factores de efectos aleatorios.

Diseño Aleatorio Simple

Cuadro: Observaciones en un experimento completamente al azar

Tratamiento	Observaciones	Suma	Media
1	$y_{11}\ \ y_{12} \ldots y_{1n}$	$y_{1.} = \sum_{j=1}^{n} y_{1j}$	$\bar{y}_{1.} = \frac{1}{n}\sum_{j=1}^{n} y_{1j}$
2	$y_{21}\ \ y_{22} \ldots y_{2n}$	$y_{2.} = \sum_{j=1}^{n} y_{2j}$	$\bar{y}_{2.} = \frac{1}{n}\sum_{j=1}^{n} y_{2j}$
$\vdots$	$\vdots$	$\vdots$	
a	$y_{a1}\ \ y_{a2} \ldots y_{an}$	$y_{a.} = \sum_{j=1}^{n} y_{aj}$	$\bar{y}_{a.} = \frac{1}{n}\sum_{j=1}^{n} y_{aj}$
Total		$y_{..} = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}$	$\bar{y}_{..} = \frac{1}{an}\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}$

Notación:

Suma	$y_{..} = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}$	Media general	$\bar{y}_{..} = \frac{y_{..}}{an}$
Suma del tratamiento i	$y_{i.} = \sum_{j=1}^{n} y_{ij}$	Media del tratamiento i	$\bar{y}_{i.} = \frac{y_{i.}}{n}$

y_{ij}	Variable de respuesta	
μ	Media general	Parámetro
τ_i	Efecto que produce el tratamiento i	Parámetro
ϵ_{ij}	Error experimental	Parámetro
σ^2	Varianza de ϵ_{ij}	Parámetro

Modelo del diseño

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

Supuestos del modelo $\epsilon_{ij} \sim IIDN(0, \sigma^2)$

Independencia	$\epsilon_{11}, \ldots\ldots, \epsilon_{an}$	Son independientes
Distribución Normal	$\epsilon_{ij} \sim N(0, \sigma^2)$	
Homogeneidad de varianzas	$\sigma_1^2 = \sigma_2^2 = \ldots.. = \sigma_a^2$	

Estimación de los parámetros: $\mu, \tau_1, \tau_2, \ldots., \tau_a$ y σ^2 se estiman por Mínimos Cuadrados (Máxima Verosimilitud produce los mismos resultados).

$$\epsilon_{ij} = y_{ij} - \mu - \tau_i$$

$$SCE = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \mu - \tau_i)^2$$

Aplicamos la derivada a cada parámetro

$$\frac{\partial SCE}{\partial \mu} = \sum_{i=1}^{a}\sum_{j=1}^{n}\frac{\partial}{\partial \mu}(y_{ij} - \mu - \tau_i)^2$$

$$= -2\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \mu - \tau_i)$$

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \mu - \tau_i) = 0$$

$$\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij} - \sum_{i=1}^{a}\sum_{j=1}^{n} \mu - \sum_{i=1}^{a}\sum_{j=1}^{n} \tau_i = 0$$

$$any_{ij} - an\mu - n\sum_{i=1}^{a} \tau_i = 0$$

$$\sum_{i=1}^{a} \tau_i = 0$$

$$any_{ij} - an\mu = 0$$

$$y_{..} - an\mu = 0$$

$$y_{..} = an\mu$$

$$\frac{y_{..}}{an} = \mu$$

$$\Rightarrow \hat{\mu} = \overline{y}_{..}$$

Mismo proceso para τ_i

$$\frac{\partial SCE}{\partial \tau_1} = \sum_{j=1}^{n} \frac{\partial}{\partial \tau_1}(y_{1j} - \mu - \tau_1)^2$$

$$= -2\sum_{j=1}^{n}(y_{1j} - \mu - \tau_1)$$

$$\sum_{j=1}^{n}(y_{1j} - \mu - \tau_1) = 0$$

$$\sum_{j=1}^{n} y_{1j} - \sum_{j=1}^{n} \mu - \sum_{j=1}^{n} \tau_1 = 0$$

$$y_{1.} - n\mu - n\tau_1 = 0$$

$$y_{1.} = n\mu + n\tau_1$$

$$y_{1.} - n\overline{y}_{..} = n\tau_1$$

$$\tau_1 = \frac{y_{1.}}{n} - \frac{n\overline{y}_{..}}{n}$$

$$\Rightarrow \hat{\tau}_1 = \overline{y}_{1.} - \overline{y}_{..}$$

$$\vdots$$

$$\hat{\tau}_a = \overline{y}_{a.} - \overline{y}_{..}$$

Resumen de los estimadores bajo la restricción $\sum_{i=1}^{a} \tau_i = 0$

μ	Media general	$\hat{\mu} = \overline{y}_{..}$
τ_i	Efecto del tratamiento *i*	$\hat{\tau}_i = \overline{y}_{i.} - \overline{y}_{..}$
μ_i	Media del tratamiento *i*	$\hat{\mu}_i = \overline{y}_{i.}$
y_{ij}	Variable de respuesta	$\hat{y}_{ij} = \overline{y}_{i.}$
σ^2	Varianza de los errores	$\hat{\sigma}^2 = CME = \frac{SCE}{n-a}$

Prueba de Hipótesis

El efecto de cada nivel del factor en la VR es cero

Se denota:

Ho:	$\tau_1 = \tau_2 = \cdots = \tau_i = 0$
Ha:	Al menos un $\tau_i \neq 0$

Anova

Cuadro: Tabla de Análisis de Varianza (ANOVA)

FV	GL	SC	CM	Pr>F
Tratamientos	$a-1$	SC_{Trat}	$CM_{\text{Tratamientos}}$	α
Error	a(n-1)	SC_{Error}	CM_{Error}	
Total	an-1	SC_{Total}		

La suma de cuadrados total (SCT) corregida se descompone en dos SC en la de tratamientos (SCTrat) y en los errores (SCE):

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\overline{y}_{..})^2=\sum_{i=1}^{a}\sum_{j=1}^{n}\left[(y_{ij}-\overline{y}_{i.})+(\overline{y}_{i.}-\overline{y}_{..})\right]^2$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{n}\left[(y_{ij}-\overline{y}_{i.})^2 - 2(y_{ij}-\overline{y}_{i.})(\overline{y}_{i.}-\overline{y}_{..}) + (\overline{y}_{i.}-\overline{y}_{..})^2\right]$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\overline{y}_{i.})^2 - 2\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\overline{y}_{i.})(\overline{y}_{i.}-\overline{y}_{..}) + \sum_{i=1}^{a}\sum_{j=1}^{n}(\overline{y}_{i.}-\overline{y}_{..})^2$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\overline{y}_{i.})^2 + \sum_{i=1}^{a} n(\overline{y}_{i.}-\overline{y}_{..})^2$$

$$= n\sum_{i=1}^{a}(\overline{y}_{i.}-\overline{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\overline{y}_{i.})^2$$

$$SCT = SCTrat + SCE$$

Grados de libertad:

$$an - 1 = a - 1 + a(n - 1)$$

FV	Gl	SC	CM	F	Pr$>F$
Tratamientos	a-1	SCTrat	CMTrat=$\frac{SCTrat}{a-1}$	$\frac{CMTrat}{CME}$	α^*
Error	a(n-1)	SCE	CME=$\frac{SCE}{N-a}$		
Total	an-1	SCT			

$$SCTrat = \sum_{i=1}^{a} n(\overline{y}_{i.} - \overline{y}_{..})^2 = \sum_{i=1}^{a} \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{an}$$

$$SCE = \sum_{i=1}^{a}\sum_{i=1}^{n}(y_{ij} - \overline{y}_{i.})^2 = \sum_{i=1}^{a}\sum_{i=1}^{n} y_{ij}^2 - \sum_{i=1}^{a}\frac{y_{i.}^2}{n}$$

$$SCT = \sum_{i=1}^{a}\sum_{i=1}^{n}(y_{ij} - \overline{y}_{..})^2 = \sum_{i=1}^{a}\sum_{i=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{an}$$

FV	Gl	SC	CM	F	Pr$>F$
Tratamientos	a-1	SCTrat	CMTrat$\frac{SCTrat}{a-1}$	$\frac{CMTrat}{CME}$	α^*
Error	N-a	SCE	CME=$\frac{SCE}{N-a}$		
Total	N-1	SCT			

$$SCTrat = \sum_{i=1}^{a} n_i(\overline{y}_{i.} - \overline{y}_{..})^2 = \sum_{i=1}^{a} \frac{y_{i.}^2}{n_i} - \frac{Y_{..}^2}{N}$$

$$SCE = \sum_{i=1}^{a}\sum_{i=1}^{n_i}(y_{ij} - \overline{y}_{i.})^2 = \sum_{i=1}^{a}\sum_{i=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{a}\frac{y_{i.}^2}{n_i}$$

$$SCT = \sum_{i=1}^{a}\sum_{i=1}^{n_i}(y_{ij} - \overline{y}_{..})^2 = \sum_{i=1}^{a}\sum_{i=1}^{n_i} y_{ij}^2 - \frac{y_{..}^2}{N} \quad N = \sum_{i=1}^{a} n_i$$

Cuadrados medios: SC dividida por los gl.

$$CMTrat = \frac{SCTrat}{a-1}, \; CME = \frac{SCE}{a(n-1)}$$

Esperanzas de Cuadrados Medios:

$$CME = E\left(\frac{SCE}{a(n-1)}\right) = \sigma^2$$

$$CMTrat = E\left(\frac{SCTrat}{a-1}\right) = \sigma^2 + \frac{n\sum_{i=1}^{a}\tau_i^2}{a-1}$$

H0: $\tau_i = 0 \qquad i = 1, 2, 3, ..., a$

Si H0: es cierta $\Rightarrow$ E(CME)= E(CMTrat)= σ^2

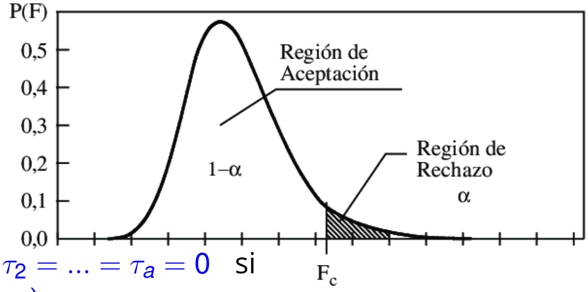
Si H0: es falsa $\Rightarrow$ E(CME) $<$ E(CMTrat)

$\frac{SCTrat}{\sigma^2}$ tiene distribución chi-cuadrada con a-1 gl.

$\frac{SCE}{\sigma^2}$ tiene distribución chi-cuadrada con N-a gl.

$\frac{SCTrat}{\sigma^2}$ y $\frac{SCE}{\sigma^2}$ son independientes

Sean U y V variables aleatorias independientes tales que $U \sim \chi^2_m$ y $V \sim \chi^2_n$, esto es U y V siguen una distribución chi-cuadrado con m y n grados de libertad respectivamente entonces la variable aleatoria $\dfrac{U/m}{V/n} \sim F_{m,n}$ donde $F_{m,n}$ denota la distribución F con m y n grados de libertad.



Se rechaza $H_0 : \tau_1 = \tau_2 = \ldots = \tau_a = 0$ si

1) $F_c > F_\alpha(a-1, N-a)$

2) $p - value < 0{,}05, (0{,}01)$. El P-valor es: $P(F(a-1, N-a) > F_c|\tau_2 = \ldots = \tau_a = 0;$

Estadístico de prueba $F_c = \dfrac{CMTrat}{CME}$

References I

Motgomery C. D. (2005). Design and analysis of experiments 6ta. Ed. John Wiley Sons, Inc. New York.
Motgomery C. D. (2005). Diseño and analisis of experimentos. 2a. Ed. John Wiley Sons, Inc. New York.
Motgomery C. D. (2010). Diseño y analisis of experimentos. Editorial Limusa. Mexico.
Peña D. (2002). Regresión y diseño de experimentos. Alianza Editorial. Madrid.