Diseño de experimentos

Isaac Cortés Olmos

Universidad de Atacama

25 de julio de 2025

Esquema

- 1 Diseño de experimentos
- ▶ Diseño de experimentos

Diseño de experimentos

Diseño de experimentos

Factor:

• Un factor es una variable que se manipula o controla en un experimento para observar su efecto en la variable de respuesta.

- Tipo de explosivo: Se quiere evaluar si diferentes tipos de explosivos (A, B, C) tienen un impacto en la fragmentación de la roca.
- Concentración de reactivo en flotación: Se busca determinar si variar la cantidad de un reactivo (por ejemplo, colector) afecta la recuperación de mineral.
- Tamaño de molienda: Estudiar cómo diferentes granulometrías de molienda (gruesa, media, fina) influyen en la eficiencia de un proceso de lixiviación.

Niveles:

 Los niveles son los diferentes valores o categorías específicas que toma un factor en el experimento.

- Tipo de explosivo:
 - ► Nivel 1: Explosivo A
 - ► Nivel 2: Explosivo B
 - ► Nivel 3: Explosivo C
- Concentración de reactivo en flotación:
 - ► Nivel 1: 50 g/ton
 - ► Nivel 2: 75 g/ton
 - ► Nivel 3: 100 g/ton
- Tamaño de molienda:
 - ▶ Nivel 1: Molienda gruesa (P80 de 200 micras)
 - ► Nivel 2: Molienda media (P80 de 150 micras)
 - ► Nivel 3: Molienda fina (P80 de 100 micras)

Tratamiento:

 Un tratamiento es una combinación específica de los niveles de todos los factores en un experimento. Es la condición experimental a la que se somete una unidad experimental.

- Experimento con dos factores: Tipo de explosivo (Niveles: A, B) y Patrón de perforación (Niveles: Cuadrado, Triangular).
 - ▶ Tratamiento 1: Explosivo A + Patrón Cuadrado
 - ► Tratamiento 2: Explosivo A + Patrón Triangular
 - ▶ Tratamiento 3: Explosivo B + Patrón Cuadrado
 - ► Tratamiento 4: Explosivo B + Patrón Triangular
- Experimento con un solo factor: (Concentración de reactivo):
 - ► Tratamiento 1: 50 g/ton
 - ► Tratamiento 2: 75 g/ton
 - ► Tratamiento 3: 100 g/ton

Diseño de experimentos

Unidad experimental:

• Es la entidad o el objeto al que se le aplica un tratamiento y sobre el cual se realiza la medición de la variable de respuesta. Es la unidad más pequeña sobre la que se realizan las observaciones o mediciones.

- En voladura: Un bloque de roca específico donde se aplica un determinado diseño de voladura (combinación de explosivo y patrón de perforación).
- En flotación: Una celda de flotación individual, o una muestra representativa de pulpa de mineral en un laboratorio, a la que se le aplica una combinación específica de reactivos y condiciones.
- En lixiviación: Una columna de lixiviación o una muestra de mineral en un vaso de precipitado a la que se le aplican ciertas condiciones de lixiviación (concentración de ácido, tamaño de partícula).

Diseño con un factor

Factor con dos niveles:

- El objetivo principal de un diseño experimental con un solo factor y dos niveles es determinar si existe una diferencia significativa en la variable de respuesta cuando el factor se cambia de un nivel a otro.
- En otras palabras, se busca establecer si el cambio en la condición del factor tiene un efecto medible y estadísticamente relevante en el resultado del proceso.

Diseño con un factor

- Objetivo: Evaluar si la adición del nuevo reactivo mejora significativamente la recuperación o la calidad del concentrado de mineral.
- Factor: Presencia/Ausencia de un nuevo reactivo.
- Niveles:
 - ▶ Nivel 1: Flotación sin el nuevo reactivo.
 - Nivel 2: Flotación con el nuevo reactivo (a una concentración fija).
- Variable de Respuesta: Recuperación de mineral valioso (ej. % de cobre recuperado) o ley del concentrado.

Contrastes de hipótesis:

• Cuando un factor tiene 2 niveles, el problema de comparación de los dos efectos se plantea como el contraste de dos hipótesis:

$$\begin{array}{lll} \mathsf{H}_0 & : & \mu_1 = \mu_2 \\ \mathsf{H}_1 & : & \mu_1 \neq \mu_2 \end{array}$$

en donde μ representa la media antes de cualquier tratamiento, suponiendo que $\mu_i = \mu + \delta_i$, donde δ_i es entonces lo que añade el tratamiento i a la μ . Note así que $\delta_i = \mu_i - \mu$, el llamado efecto del i-ésimo nivel.

• En el caso de 2 niveles, en lugar de considerar la comparación entre todos los contrastes $(\bar{y}_{i\cdot} - \bar{y}_{i'\cdot})$, $i \neq i'$, se puede pensar como más eficiente el analizar sólo las diferencias

$$\widehat{\delta}_i = (\overline{y}_i - \overline{y}_{\cdot \cdot}), i = 1, 2,$$

donde $\bar{y}_{..}$ es un estimador de μ

Contrastes de hipótesis:

• Definamos el efecto estimado del i-ésimo nivel del factor como

$$\widehat{\delta}_i = (\overline{y}_i - \overline{y}_{\cdot \cdot}), i = 1, 2.$$

- Si $\widehat{\delta}_i$ es pequeño se dirá que el efecto es bajo. Así si todas las $\widehat{\delta}_i$'s son cercanas a cero, se apoyaría a la hipótesis H_0 . en donde μ representa la media antes de cualquier tratamiento, suponiendo que $\mu_i = \mu + \delta_i$, donde δ_i es entonces lo que añade el tratamiento i a la μ . Note así que $\delta_i = \mu_i \mu$, el llamado efecto del i-ésimo nivel.
- En el caso de 2 niveles, en lugar de considerar la comparación entre todos los contrastes $(\bar{y}_{i\cdot} \bar{y}_{i'\cdot})$, $i \neq i'$, se puede pensar como más eficiente el analizar sólo las diferencias

$$\widehat{\delta}_i = (\overline{y}_{i.} - \overline{y}_{..}), i = 1, 2,$$

donde $\bar{y}_{..}$ es un estimador de μ

Supuestos:

ullet Si y_i representa a la respuesta ante el tratamiento i, se supone que

$$y_i = \mu_i + \varepsilon, i = 1, 2$$

donde ε representa el error experimental; respuestas medias ante cada tratamiento denotadas por μ_1, μ_2 potencialmente diferentes.

- Del término de error se supone que:
 - ightharpoonup arepsilon tiene media igual a cero.
 - La varianza de ε en cualquier tratamiento es constante, digamos igual a σ^2 (homogeneidad de varianzas).
- $m{e}$ como variable aleatoria es descrita adecuadamente por la función de densidad normal. Bajo los supuestos anteriores, ya con los datos, se tendrá entonces que

$$y_{ij} = \mu_i + \varepsilon_{i,j}, i = 1, 2; j = 1, ..., n_i.$$



Cuadro 1: Tabla de Análisis de Varianza (ANOVA)

Fuente de Variación	Grados de Libertad	Suma de Cuadrados	Cuadrados Medios	Estadístico F	Valor p
Tratamiento	2 - 1	$\sum_{i=1}^2 n_i \widehat{\boldsymbol{\delta}}_i^2$	$\frac{SC_{trat}}{2-1}$	CM _{trat} CM _{error}	p_{valor}
Error	N-2	$\sum_{i=1}^{2} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_{i.})^2$	$\frac{SC_{ermr}}{N-2}$		
Total	N-1	$\sum_{i=1}^{2} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_{})^2$			

Ejemplo:

Una minera quiere comparar si dos equipos (Equipo A y Equipo B) producen diferente rendimiento (toneladas/día). Se toman 6 mediciones por equipo.

Equipo A	Equipo B
102	110
105	113
98	108
100	112
104	111
101	109

Ejemplo:

Se evalúa si hay diferencia en la concentración de cobre (ppm) entre dos técnicas de lixiviación (Química y Biológica). Se toman 5 muestras por técnica.

Química	Biologica
320	300
310	295
315	298
318	302
317	301

Contrastes de hipótesis:

• Cuando un factor tiene *k* niveles, el problema de comparación de los *k* efectos se plantea como el contraste de dos hipótesis:

$$\mathsf{H}_0$$
 : $\mu_1 = \mu_2 = \ldots = \mu_k$

 H_1 : $\mu_i
eq \mu_{i'}$ para al menos un par $i \neq i'$

en donde μ representa la media antes de cualquier tratamiento, suponiendo que $\mu_i = \mu + \delta_i$, donde δ_i es entonces lo que añade el tratamiento i a la μ . Note así que $\delta_i = \mu_i - \mu$, el llamado efecto del i-ésimo nivel.

• En el caso de k niveles, en lugar de considerar la comparación entre todos los contrastes $(\bar{y}_{i\cdot} - \bar{y}_{i'\cdot})$, $i \neq i'$, se puede pensar como más eficiente el analizar sólo las diferencias

$$\widehat{\delta}_i = (\overline{y}_i - \overline{y}_{\cdot \cdot}), i = 1, 2, \dots, k$$

donde $\overline{y}_{..}$ es un estimador de μ .



Contrastes de hipótesis:

• Definamos el efecto estimado del i-ésimo nivel del factor como

$$\widehat{\delta}_i = (\overline{y}_i - \overline{y}_{\cdot \cdot}), i = 1, 2, \dots, k$$

- Si $\widehat{\delta}_i$ es pequeño se dirá que el efecto es bajo. Así si todas las $\widehat{\delta}_i$'s son cercanas a cero, se apoyaría a la hipótesis H_0 . en donde μ representa la media antes de cualquier tratamiento, suponiendo que $\mu_i = \mu + \delta_i$, donde δ_i es entonces lo que añade el tratamiento i a la μ . Note así que $\delta_i = \mu_i \mu$, el llamado efecto del i-ésimo nivel.
- En el caso de k niveles, en lugar de considerar la comparación entre todos los contrastes $(\bar{y}_{i\cdot} \bar{y}_{i'\cdot})$, $i \neq i'$, se puede pensar como más eficiente el analizar sólo las diferencias

$$\widehat{\delta}_i = (\overline{y}_i - \overline{y}_i), i = 1, 2, \dots, k$$

donde $\bar{y}_{..}$ es un estimador de μ .

Supuestos:

ullet Si y_i representa a la respuesta ante el tratamiento i, se supone que

$$y_i = \mu_i + \varepsilon, i = 1, 2$$

donde ε representa el error experimental; respuestas medias ante cada tratamiento denotadas por μ_1, μ_2 potencialmente diferentes.

- Del término de error se supone que:
 - ightharpoonup arepsilon tiene media igual a cero.
 - La varianza de ε en cualquier tratamiento es constante, digamos igual a σ^2 (homogeneidad de varianzas).
- $m{e}$ como variable aleatoria es descrita adecuadamente por la función de densidad normal. Bajo los supuestos anteriores, ya con los datos, se tendrá entonces que

$$y_{ij} = \mu_i + \varepsilon_{i,j}, i = 1, 2; j = 1, \ldots, n_i.$$



Cuadro 2: Tabla de Análisis de Varianza (ANOVA)

Fuente de Variación	Grados de Libertad	Suma de Cuadrados	Cuadrados Medios	Estadístico F	Valor p
Tratamiento	k-1	$\sum_{i=1}^k n_i \widehat{\delta}_i^2$	$\frac{SC_{trat}}{k-1}$	CM _{trat} CM _{error}	Pvalor
Error	N-k	$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_{i.})^2$	$\frac{SC_{error}}{N-k}$		
Total	N-1	$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_{})^2$			

Ejemplo:

Se evalúa la eficiencia de 6 tipos de reactivos químicos sobre la recuperación de oro (%), usando 6 réplicas por grupo.

Tipo 1	Tipo 2	Tipo 3	Tipo 4	Tipo 5	Tipo 6
82.18	85.90	87.37	82.32	86.64	87.36
82.78	85.25	86.73	84.53	86.81	85.94
80.65	85.55	87.77	84.75	87.08	84.90
83.98	86.90	87.74	83.79	87.17	85.84
83.24	86.44	86.43	82.68	88.09	87.20
83.20	85.12	89.75	83.02	89.01	88.12

Ejemplo:

Se evalúa el efecto de 5 tiempos de lixiviación (12, 24, 36, 48, 60 horas) sobre la recuperación de oro (%). Se recolectan 5 réplicas por tiempo.

12 Hrs	24 Hrs	36 Hrs	48 Hrs	60 Hrs
70.23	75.95	78.59	77.89	78.45
70.43	75.50	76.53	78.14	78.45
69.73	74.97	77.25	79.80	78.26
72.66	75.28	77.80	79.57	78.20
70.62	73.12	76.43	78.28	77.14

Intervalos de Confianza

Diferencia de tratamientos:

• En general al construirse un intervalo de confianza para un parámetro θ , se afirma que con un nivel de confianza $100(1-\alpha)\%$ se espera que $\theta \in I(\widehat{\theta})$, donde $I(\widehat{\theta})$ representa un intervalo construido a partir de $\widehat{\theta}$, una estimación de θ , una estimación de θ ; por ejemplo, si $\theta = \mu_i - \mu_j$, un estimador respectivo es

$$\widehat{\boldsymbol{\theta}} = \widehat{\mu}_i - \widehat{\mu}_j = \overline{y}_{i\cdot} - \overline{y}_{j\cdot}; \tag{1}$$

ullet Típicamente el intervalo de confianza para heta será de la forma

$$\theta \in \widehat{\theta} \pm Q(\widehat{\theta}) \times (ES_{\widehat{\theta}})$$
(2)

donde $(ES_{\widehat{\theta}})$ representa el error estándar estimado de $\widehat{\theta}$, $Q(\widehat{\theta})$ representa un cuantil de orden α de la distribución muestral asociada a $\widehat{\theta}$, y dependiente del tipo de comparación por realizar.

Referencias I

- Devore, J. L. (2009). Probabilidad y estadística para ingeniería y ciencias. Cengage Learning Editores.
- Díaz Mata, A. (2013). Estadística aplicada a la Administración y Economía. McGraw-Hill.
- Ross, S. M. (2002). Probabilidad y estadística para ingeniería y ciencias. McGraw–Hill.
- Normalis Spiegel, M. R., & Stephens, L. J. (2009). Estadística. McGraw–Hill.