21 de junio de 2022

Diseño de bloques completos aleatorizados

Módulo 3
—
Diseño de Experimentos

Agenda de hoy

1. Diseño de bloques completos aleatorizados
2. Factores de bloque
3. Fuentes de variabilidad
4. Modelo estadístico
5. Análisis de Varianza
6. Otros tipos de diseños experimentales

Diseño de bloques completos aleatorizados

Cuando se lleva a cabo un experimento en el que se comparan varios tratamientos y se estudia el efecto de un factor sobre la variable respuesta, se busca que las diferencias observadas entre los tratamientos se deban al factor que se está estudiando y no a otros factores que no se están teniendo en cuenta en el experimento. **Cuando esto no ocurre, es decir, no se logra controlar el efecto de otros factores, los resultados y conclusiones del experimento pueden verse afectados.**

Diseño de bloques completos aleatorizados

En el experimento donde se comparan los cuatro métodos de ensamble, si cada método es asignado a un operador diferente y uno de los operadores es más hábil que los demás, es posible que parezca que su método de ensamble sea el mejor. En este caso, la comparación que se está haciendo de los métodos no es la adecuada.

Es necesario anular el posible efecto del factor operador

Diseño de bloques completos aleatorizados

1

Utilizar el mismo operador en los 4 métodos de ensamble

↓

No es aconsejable porque es posible que el resultado del experimento no se mantenga al utilizar otros operadores

2

Cada operador trabaja durante el experimento con cada uno de los métodos de ensamble

↓

Es el más recomendable ya que permite tener resultados de la comparación de los métodos que son válidos para todos los operadores

BLOQUEO

Diseño de bloques completos aleatorizados

Bloque 1	**Bloque 2**	**Bloque 3**	**Bloque 4**
Operador 1	**Operador 2**	**Operador 3**	**Operador 4**

En cada bloque se prueban TODOS los tratamientos, es decir, todos los métodos de ensamble y la aleatorización se hace dentro de cada bloque

Diseño de bloques completos aleatorizados

Método de ensamble	Operador			
	1	2	3	4
A	6	9	7	8
B	7	10	11	8
C	10	16	11	14
D	10	13	11	9

Los 4 operadores ensamblan el producto con los 4 métodos de ensamble. El orden se define de forma aleatoria

Factores de bloque

Los factores de bloque son factores adicionales al factor de estudio que se incorporan a un experimento comparativo. **Estos factores se incluyen al experimento, no para analizar sus efectos sino para estudiar de manera adecuada el factor de estudio**

Factores de bloque

Es posible que existan otros factores de bloque que deban controlarse para comparar adecuadamente los tratamientos. Sin embargo, solo se deben incluir los factores que por conocimiento del proceso sabemos que pueden afectar considerablemente el resultado de la comparación

Fuentes de variabilidad

En un diseño de bloques completos aleatorizados se consideran tres fuentes de variabilidad:

El factor de estudio + **El factor de bloqueo** + **El error aleatorio**

Modelo estadístico

Cuando se utiliza un diseño de bloques completos aleatorizados, se supone que cada medición será el resultado del efecto del tratamiento al que pertenece, del efecto del bloque al que pertenece y de cierto error. En este caso, el modelo estadístico que describe los datos obtenidos está dado por:

$$Y_{ij} = \mu_i + \gamma_j + \varepsilon_{ij}$$

Modelo estadístico

$$i = 1,2,\dots,\mathrm{k}$$
$$j = 1,2,\dots,\mathrm{b}$$
$$Y_{ij} = \textit{medición correspondiente al tratamiento i y al bloque j}$$
$$\mu_i = \textit{respuesta media del tratamiento i}$$
$$\gamma_i = \textit{efecto debido al bloque j}$$
$$\varepsilon_{ij} = \textit{error asociado a la medicion } Y_{ij}$$

Modelo estadístico

La respuesta media de un tratamiento μ_i se puede expresar como:

$$\boxed{\mu_i = \mu + \tau_i}$$

$$\mu = \textit{media global de las poblaciones}$$
$$\tau_i = \textit{parametro que mide el efecto del tratamiento } i$$

Modelo estadístico

Si el análisis de varianza (ANOVA) resulta significativo, se puede estimar el modelo ajustado que esta dado por:

$$\hat{Y}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\gamma}_j$$

Modelo estadístico

Si el análisis de varianza (ANOVA) resulta significativo, se puede estimar el modelo ajustado que esta dado por:

$$\hat{Y}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\gamma}_j$$

$$\hat{Y}_{ij} = \bar{Y}_{..} + (\bar{Y}_{i.} - \bar{Y}_{..}) + (\bar{Y}_{.j} - \bar{Y}_{..})$$

$$\hat{Y}_{ij} = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..}$$

$$e_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}$$

La respuesta estimada para cada observación es se calcula de esta forma

Análisis de varianza (ANOVA)

En este tipo de diseño experimental, también nos interesa determinar si hay alguna diferencia entre las medias muestrales de cada uno de los tratamientos que se están comparando. Para lograrlo, utilizamos el método del **análisis de varianza con dos criterios de clasificación** ya que se controlan dos fuentes de variación: el factor de estudio y el factor de bloqueo.

Recordemos que con este método se hacen suposiciones de normalidad, varianza constante e independencia que se deben verificar.

Análisis de varianza (ANOVA)

La hipótesis fundamental a probar es:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_K$$

$$H_a: No\ todas\ las\ medias\ poblacionales\ son\ iguales$$

Análisis de varianza (ANOVA)

También es posible probar la siguiente hipótesis respecto al factor de bloqueo:

$$H_0: \gamma_1 = \gamma_2 = \gamma_3 = \cdots = \gamma_b = 0$$

$$H_a: \gamma_j \neq 0 \; para \; algun \; bloque \; j$$

Análisis de varianza (ANOVA)

Suma de cuadrados de tratamientos

$$SC_{TRAT} = \sum_{i=1}^{k} n_i(\bar{Y}_{i.} - \bar{Y}_{..})^2$$

Suma de cuadrados de bloques

$$SC_B = \sum_{j=1}^{b} \frac{Y_{.j}^2}{k} - \frac{Y_{..}^2}{N}$$

Suma de cuadrados del error

$$SC_E = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_{i.})^2$$

Suma de cuadrados total

$$SC_T = SC_{TRAT} + SC_B + SC_E$$

Análisis de varianza (ANOVA)

Usemos la técnica del análisis de varianza para probar la igualdad de los tiempos promedio de los métodos de ensamble. **Esta prueba nos permitirá determinar si existe un efecto del método de ensamble sobre el tiempo de ensamble y también si existe un efecto del operador sobre el tiempo de ensamble**

Análisis de varianza (ANOVA)

Prueba F

1 **Definir las hipótesis (para el factor de estudio)**

$$H_0: \mu_A = \mu_B = \mu_C = \mu_D$$

$$H_a: No\ todos\ los\ tiempos\ promedio\ de\ ensamble\ son\ iguales$$

Análisis de varianza (ANOVA)

Prueba F

1. **Definir las hipótesis (para el factor de bloqueo)**

$$H_0: \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$$

$$H_a: \gamma_j \neq 0 \; para \; algun \; bloque \; j$$

Análisis de varianza (ANOVA)

Prueba F

2 **Definir el estadístico de prueba (para el factor de estudio)**

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

Análisis de varianza (ANOVA)

Prueba F

2 **Definir el estadístico de prueba (para el factor de bloqueo)**

$$F = \frac{CM_B}{CM_E}$$

Análisis de varianza (ANOVA)

Prueba F

3 **Definir la regla de rechazo**

Valor P

Rechazo H_0 si :

$$valor\ p \leq \alpha$$

Análisis de varianza (ANOVA)

Si llevamos a cabo estos cálculos, esta es la ANOVA resultante:

	Suma de cuadrados	Grados de libertad	Cuadrados medios	F	Valor-p
Métodos	61.5	3	20.5	10.25	0.003
Operadores	28.5	3	9.5	4.75	0.030
Error	18.0	9	2.0		
Total	**108.0**	**15**			

Análisis de varianza (ANOVA)

Utilicemos la función **aov()** de R para llevar a cabo estas pruebas....

Análisis de varianza (ANOVA)

- **Conclusión (para el factor de estudio):** Como el valor-p es igual a 0.003 y es menor al nivel de significancia de 0.05, se rechaza la hipótesis nula y se concluye que al menos dos de los tiempos promedio de ensamble son diferentes.

- **Conclusión (para el factor de bloqueo):** Como el valor-p es igual a 0.03 y es menor al nivel de significancia de 0.05, se concluye que el factor de bloqueo afecta el tiempo de ensamble. Es decir, el tiempo promedio que tardan los operadores ensamblando el producto es significativamente diferente.

Análisis de varianza (ANOVA)

Si la prueba ANOVA para el efecto de bloques resulta significativa, implica que el factor de bloqueo tiene influencia sobre la variable respuesta y por lo tanto, deber ser tomado en cuenta para mejorar esta variable. En caso contrario, se concluye que la influencia del factor de bloqueo sobre la variable respuesta no es significativa y por lo tanto, no se debe controlar este factor en futuros experimentos.

Análisis de varianza (ANOVA)

- Es importante tener en cuenta que el objetivo final del experimento no es comparar a los operadores. Ellos se agregan en el estudio para lograr una comparación más justa y precisa de los métodos de ensamble.

- Con los resultados del experimento podemos escoger el método de ensamble más eficiente en términos de tiempo pero no escogemos un operador. Sin embargo, a partir de los resultados si es posible tomar decisiones como por ejemplo planear una sesión de entrenamiento a los operadores que obtuvieron tiempos de ensamble muy diferentes a los demás.

Comparaciones o pruebas múltiples

Cuando se rechaza la hipótesis de igualdad de los tratamientos con el análisis de varianza, es de interés saber cuáles de los tratamientos son diferentes y para esto llevamos a cabo pruebas múltiples para cada uno de los pares de tratamientos

Comparaciones o pruebas múltiples

Método LSD (Diferencia Mínima Significativa)

1 **Definir las hipótesis**

$$H_0: \mu_i = \mu_j$$

$$H_a: \mu_i \neq \mu_j$$

Comparaciones o pruebas múltiples

Método LSD (Diferencia Mínima Significativa)

2 **Definir el estadístico de prueba**

$$t = \frac{\bar{Y}_{i.} - \bar{Y}_{j.}}{\sqrt{CME\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}}$$

Comparaciones o pruebas múltiples

Método LSD (Diferencia Mínima Significativa)

3 **Definir la regla de rechazo**

Valor P

Rechazo H_0 si :

$$valor\ p \ \leq \ \alpha$$

Comparaciones o pruebas múltiples

Método LSD (Diferencia Mínima Significativa)

Otra forma de rechazar la hipótesis nula...

Rechazo H_0 si :

$$\left|\bar{Y}_{i.} - \bar{Y}_{j.}\right| \geq LSD$$

$$LSD = t_{\frac{\alpha}{2}, n-k} \sqrt{CME\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

Pruebas múltiples en nuestro experimento

$$H_0 : \mu_A = \mu_B \quad vs. \quad H_A : \mu_A \neq \mu_B$$
$$H_0 : \mu_A = \mu_C \quad vs. \quad H_A : \mu_A \neq \mu_C$$
$$H_0 : \mu_A = \mu_D \quad vs. \quad H_A : \mu_A \neq \mu_D$$
$$H_0 : \mu_B = \mu_C \quad vs. \quad H_A : \mu_B \neq \mu_C$$
$$H_0 : \mu_B = \mu_D \quad vs. \quad H_A : \mu_B \neq \mu_D$$
$$H_0 : \mu_C = \mu_D \quad vs. \quad H_A : \mu_C \neq \mu_D$$

Posibles pares de hipótesis

Pruebas múltiples en nuestro experimento

$$LSD = 2.26$$

Prueba	Diferencia muestras de tiempos promedio	Decisión
$H_0: \mu_A = \mu_B$	1.50	No significativa
$H_0: \mu_A = \mu_C$	5.25	Significativa
$H_0: \mu_A = \mu_D$	3.25	Significativa
$H_0: \mu_B = \mu_C$	3.75	Significativa
$H_0: \mu_B = \mu_D$	1.75	No significativa
$H_0: \mu_C = \mu_D$	2.00	No significativa

Pruebas múltiples en nuestro experimento

Utilicemos la función **LSD.test()** del paquete **agricolae** de R para llevar a cabo estas pruebas….

Pruebas múltiples en nuestro experimento

```
$groups
  tiempo groups
C  12.75      a
D  10.75     ab
B   9.00     bc
A   7.50      c
```

Los tratamientos que tienen diferentes caracteres en la columna grupos tienen medias significativamente diferentes

Otros tipos de diseños experimentales

- **Diseño en cuadro latino:** Diseño en el que se controlan dos factores de bloque y uno de tratamientos o de estudio y todos los factores tienen la misma cantidad de niveles **(Consultar en la página 109 del libro Análisis y Diseño de Experimentos).**

- **Diseño en cuadro grecolatino:** Diseño en el que se controlan tres factores de bloque y uno de tratamientos o de estudio y todos los factores tienen la misma cantidad de niveles **(Consultar en la página 115 del libro Análisis y Diseño de Experimentos).**