

3. DISEÑO EN BLOQUES ALEATORIZADOS

En muchos experimentos además de que interesa investigar la influencia de un factor controlado sobre la variable de respuesta, como en la sección anterior, existe una fuente de variación adicional (bloque) que puede y debe ser sistematizada y controlada durante el experimento, con el propósito de disminuir el error experimental y obtener mejores conclusiones sobre el factor controlado.

3.1 Ejemplo Práctico de un Diseño por Bloques Aleatorizados

Se diseñó un experimento para estudiar el rendimiento de cuatro detergentes. Las siguientes lecturas de "blancura" se obtuvieron con un equipo especial diseñado para 12 cargas de lavado distribuidas en tres modelos de lavadoras:

Detergente	Lavad.1	Lavad.2	Lavad.3	Media
A	45	43	51	46.3
B	47	44	52	47.6
C	50	49	57	52
D	42	37	49	42.6
Media	46	43.2	52.2	47.16

Variable de respuesta: Blancura

Factor controlado: Tipo de detergente

Bloque: Tipo de lavadora

Modelo matemático o modelo estadístico:

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \quad i=1,2,\dots,a; \quad j=1,2,\dots,b$$

Donde y_{ij} es blancura; μ media global o media general; τ_i efecto del factor o efecto

del tipo de detergente; β_j efecto del bloque o efecto de la lavadora; \mathcal{E}_{ij} error aleatorio.

Hipótesis:

Ho: no hay efecto del tipo de detergente en la blancura de la ropa

Ha: si hay efecto del tipo de detergente en la blancura de la ropa

Es claro que lo que interesa investigar es si los cuatro detergentes son iguales o diferentes. Y si son diferentes ¿cuál es el mejor? Pero se tiene una variable extraña: el equipo de lavado, que puede influir en la blancura, y que por lo tanto es necesario bloquearla.

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
EFFECTOS PRINCIPALES					
A:LAVADORA	170.167	2	85.0833	65.17	0.0001
B:DETERGENTE	133.667	3	44.5556	34.13	0.0004
RESIDUOS	7.83333	6	1.30556		
TOTAL (CORREGIDO)	311.667	11			

3.1.1 Análisis de Varianza

En la tabla 3.1 se puede ver el Análisis de Varianza para el diseño en bloques,

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
EFFECTOS PRINCIPALES					
A:DETERGENTE	133.667	3	44.5556	34.13	0.0004
B:LAVADORA	170.167	2	85.0833	65.17	0.0001
RESIDUOS	7.83333	6	1.30556		
TOTAL (CORREGIDO)	311.667	11			

Tabla 3.1 Análisis de Varianza para blancura de la ropa

El valor de p para el detergente es de 0.0004, el cual es menor que $\alpha = 0.05$, nivel de confianza dado, se puede concluir que si existen diferencias significativas entre los tipos de detergentes.

Con el propósito de ver cual detergente es mejor, se realizó la prueba de rangos múltiples con el método LSD, que a continuación se muestra:

3.2 Método de la Mínima Diferencia Significativa para un diseño por bloques

Una vez rechazado la hipótesis nula, con base en una prueba F de análisis de varianza, se desea probar las diferencias para cada para de medias de los detergentes.

Suponiendo una hipótesis alterna bilateral, dos medias se consideran diferentes si

$$|\bar{Y}_i - \bar{Y}_j| > t_{\alpha/2, N-a} \sqrt{\frac{CM_E}{b}} .$$

La cantidad

$$LSD = t_{\alpha/2, N-a} \sqrt{\frac{2 CM_E}{b}}$$

se denomina mínima diferencia significativa.

Continuando con nuestro ejemplo de los detergentes, tenemos que

TIPO DE DETERGENTE	MEDIA
A	46.3
B	47.6
C	52
D	42.6

Analizando las tablas de la distribución t-student, para un nivel de confianza $\alpha=0.05$ y 6 grados de libertad el valor del estadístico t es:

$$t_{(\alpha/2, N-a)} = t_{(0.05/2, 6)} = t_{(0.025, 6)} = 2.446$$

En el análisis de varianza se puede ver que el valor de cuadrado medio del error es

$$CM_{Error} = 1.30556$$

y como $b=3$, el valor del LSD es

LSD=

$$t_{(\alpha/2, N-a)} \sqrt{\frac{2CM_{Error}}{b}} = 2.446 \sqrt{\frac{2(1.30556)}{3}} = 2.28$$

Comparando las medias de los detergentes **A** y el detergente **B**, tenemos que:

$$|46.3 - 47.6| = 1.3$$

Como $1.3 < 2.28$ por lo tanto el detergente **A** no es diferente al detergente **B**.

Comparando las medias del detergente **A** y el detergente **C**, se tiene que:

$$|46.3 - 52| = 5.7$$

Como $5.7 > 2.28$ por lo tanto el detergente **A** es diferente al detergente **C**.

Comparando las medias del detergente **A** y el tipo de detergente **D**, se tiene que:

$$|46.3 - 42.6| = 3.7$$

Como $3.7 > 2.28$ por lo tanto hay diferencias entre el tipo detergente **A** y el detergente **D**.

Comparando las medias de los detergentes **B** y el detergente **C**, tenemos que:

$$|47.6 - 52| = 4.4$$

Como $4.4 > 2.28$ por lo tanto el detergentes **B** es diferente al detergente **C**.

Comparando las medias del detergente **B** y el detergente **D**, se tiene que:

$$|47.6 - 42.6| = 5$$

Como $5 > 2.28$ por lo tanto el detergente **B** es diferente al detergente **D**.

Comparando las medias del detergente **C** y el detergente **D**, se tiene que:

$$| 52-42.6 | = 9.4$$

Como $9.4 > 2.28$ por lo tanto el detergente **C** es diferente al detergente **D**.

De acuerdo a los resultados anteriores podemos concluir que el detergente **C** es el que da mayor blancura, ya que presenta el mayor promedio que los otros tres. Con esta prueba, podemos ver que el detergente que es mejor para la blancura es el **C**, ya que presenta mayor promedio y es diferente a los demás. En la tabla 3.2 se sintetizan estos resultados.

Método: 95.0 porcentaje LSD

DETERGENTE	Casos	Media LS	Sigma LS	Grupos Homogéneos
D	3	42.6667	0.659686	X
A	3	46.3333	0.659686	X
B	3	47.6667	0.659686	X
C	3	52.0	0.659686	X

Contraste	Sig.	Diferencia	+/- Límites
A - B		-1.33333	2.28282
A - C	*	-5.66667	2.28282
A - D	*	3.66667	2.28282
B - C	*	-4.33333	2.28282
B - D	*	5.0	2.28282
C - D	*	9.33333	2.28282

* indica una diferencia significativa.

Tabla 3.2 Prueba de comparación de medias para los tipos de detergentes

Una gráfica representativa de estos resultados es la **gráfica de medias**. En la figura 3.1 se puede ver la gráfica de medias de ejemplo del tipo de detergentes. Nótese que esta gráfica, se puede establecer que en promedio el detergente **C** es mejor, ya que presenta el mejor promedio de los 4 tipos de detergentes. Si el detergente **C** es el detergente recomendado se esperarían promedios entre 50.38 y 53.61 (ver tabla 3.3).

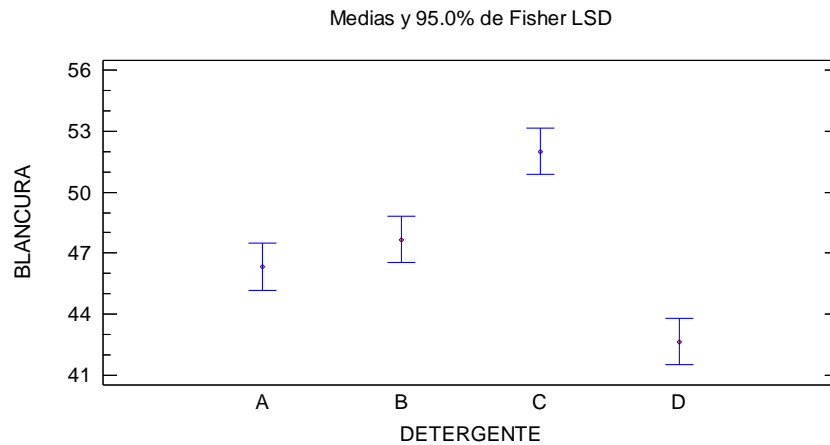


Figura 3.1 Graficas de Medias para detergentes

Tabla de Medias para BLANCURA con intervalos de confianza del 95.0%

			<i>Error</i>	<i>Límite</i>	<i>Límite</i>
<i>Nivel</i>	<i>Casos</i>	<i>Media</i>	<i>Est.</i>	<i>Inferior</i>	<i>Superior</i>
MEDIA GLOBAL	12	47.1667			
DETERGENTE					
A	3	46.3333	0.659686	44.7191	47.9475
B	3	47.6667	0.659686	46.0525	49.2809
C	3	52.0	0.659686	50.3858	53.6142
D	3	42.6667	0.659686	41.0525	44.2809

Tabla 3.3 Tabla de Medias para los detergentes

3.3 SUPUESTOS DEL MODELO POR BLOQUES

Al igual que en el diseño unifactorial es necesario que en el diseño por bloques también se chequen los supuestos de los residuales, los cuales son:

1. **VARIANZA CONSTANTE**
2. **INDEPENDENCIA**
3. **NORMALIDAD**

Para ello, primeramente se calculan los residuales mediante la siguiente formula:

$$e_{ij} = y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..}$$

Donde y_{ij} es el valor obtenido en el experimento, \bar{y}_i es el promedio del i-ésimo detergente, \bar{y}_j es el j-ésimo bloque, $\bar{y}_{..}$ es el promedio general.

Los residuales del detergente **A**, se calculan así:

$$e_{11}=45-46.33-46+47.166=-0.16$$

$$e_{12}=43-46.33-43.25+47.166=0.583$$

$$e_{13}=51-46.33-52.2+47.166=-.364$$

Los Residuales del detergente **B** son

$$e_{21}=47-47.6-46+47.166=0.566$$

$$e_{22}=44-47.6-43.25+47.166=0.316$$

$$e_{23}=52-47.6-52.25+47.166=-0.634$$

Los residuales para el detergente **C**

$$e_{31}=50-52-46+47.166=-0.834$$

$$e_{32}=49-52-43.25+47.166=0.916$$

$$e_{33}=57-52-52.25+47.166=-0.084$$

Los residuales para el detergente **D**

$$e_{41}=42-42.6-46+47.166=0.566$$

$$e_{42}=37-42.6-43.25+47.166=-1.684$$

$$e_{43}=49-42.6-52.25+47.166=1.316$$

Los resultados anteriores se sintetizan en la tabla 3.4,

DETERGENTES			
A	B	C	D
-0.164	0.566	-0.834	0.566
0.586	0.316	0.916	-1.684
-0.364	-0.634	-0.084	1.316

Tabla 3.4 Residuales

3.3.1 SUPUESTO DE VARIANZA CONSTANTE

El supuesto de varianza constante se puede verificar en la figura 3.2, donde se puede ver que no hay ningún tipo de problema.

GRAFICA DE RESIDUALES VS LOS NIVELES DEL FACTOR

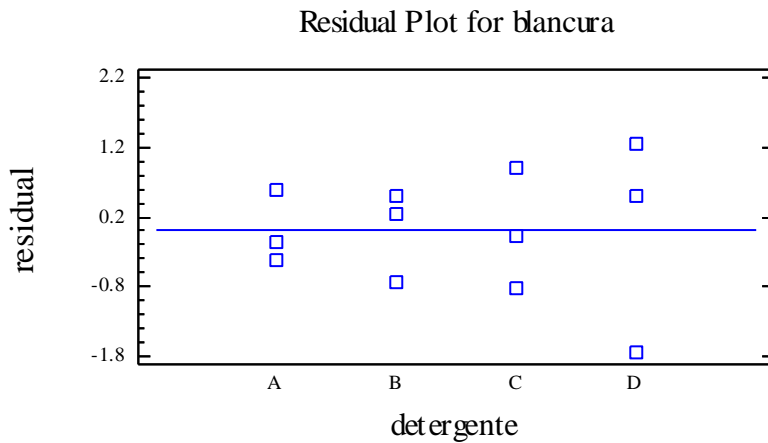


Figura 3.2 Verificación de la varianza constante

En la figura 3.2 se presentan gráficamente los niveles del factor contra los residuales. En esta gráfica no debe presentarse ningún tipo patrón de comportamiento. La variación de los puntos tanto de los niveles A, B, C y D no deben presentar diferencias. No se debe distinguir la presencia

de un embudo. En este caso no se presenta patrón inusual por lo que podemos concluir que si se cumple el supuesto de varianza constante.

3.3.2 SUPUESTO DE INDEPENDENCIA

Para checar el supuesto de independencia se requiere ordenar los residuales según el orden en que se corrió el experimento

BLOQUE	DETERGENTES			
	A	B	C	D
1	(2)-0.164	(1)0.566	(4)-0.834	(3)0.566
2	(4)0.586	(1)0.316	(2)0.916	(3)-1.684
3	(2)-0.364	(1)-0.634	(4)-0.084	(3)1.316

Por ejemplo, El detergente A se corrió en la segunda corrida dentro de la lavadora uno, en la corrida cuatro dentro de la lavadora dos y en la segunda corrida dentro de la lavadora tres, y así sucesivamente para los otros detergentes.

RESIDUALES VS ORDEN DE CORRIDA (O SECUENCIA DE TIEMPO)

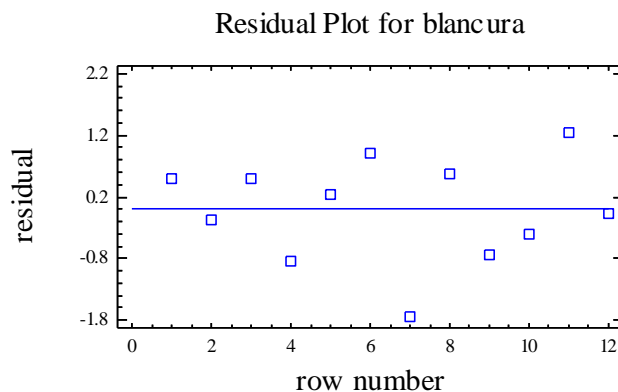


Figura 3.3 Grafica de verificación del supuesto de independencia de los residuos

En la figura 3.3 se presenta gráficamente el número de corrida contra el valor del residual. Es esta gráfica no se debe presentar ningún tipo de patrón de comportamiento, los puntos deben verse completamente dispersos. De acuerdo a esta gráfica se puede concluir que si se cumple el supuesto de independencia.

3.3.3 SUPUESTO DE NORMALIDAD

Un procedimiento útil consiste en construir una gráfica de probabilidad normal de los residuos. Una gráfica de este tipo es la representación de la distribución acumulada de los residuos sobre papel de probabilidad normal, en otras palabras, es papel para gráficas cuya escala de ordenadas es tal que la distribución normal acumulada sea una recta. Para construir una gráfica de probabilidad normal se hace el siguiente procedimiento:

1. Se ordenan los residuos en orden ascendente:

ORDEN K	RESIDUALES ORDENADOS
1	-1.75
2	-0.833333
3	-0.75
4	-0.416667
5	-0.166667
6	-0.083333
7	0.25
8	0.5
9	0.5
10	0.583333
11	0.916667
12	1.25

2. A cada residuo se le calcula su punto de probabilidad acumulada mediante la siguiente formula:

$$P_k = \frac{(k - 0.5)}{N}$$

ORDEN K	RESIDUALES ORDENADOS	$P_k = \frac{(k - 0.5)}{12}$
1	-1.75	0.041666667
2	-0.833333	0.125
3	-0.75	0.208333333
4	-0.416667	0.291666667
5	-0.166667	0.375
6	-0.0833333	0.458333333
7	0.25	0.541666667
8	0.5	0.625
9	0.5	0.708333333
10	0.583333	0.791666667
11	0.916667	0.875
12	1.25	0.958333333

3. Se grafican e_{ij} en el eje de las X's y los $P_k = \frac{(k - 0.5)}{12}$ en el eje de las Y's, como se muestra en la figura 3.4,

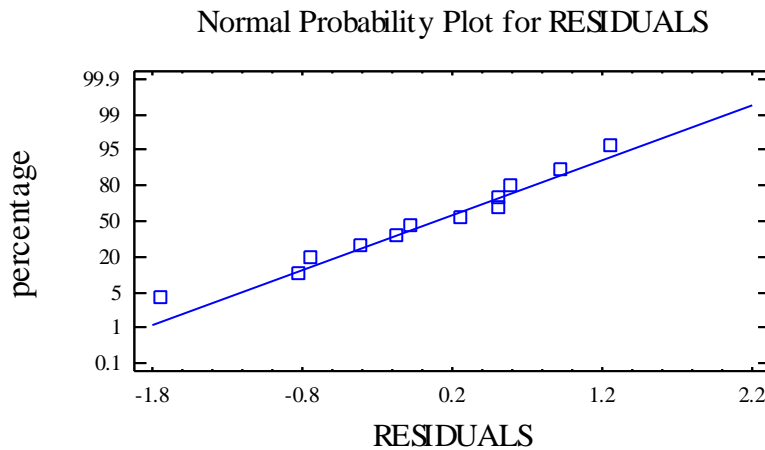


Figura 3.4 Grafica de probabilidad normal para residuos

En la gráfica de la figura 3.4 se puede ver que la mayoría de los puntos se ajustan a la línea recta, lo que significa que los residuales si cumplen el supuesto de normalidad.