

Diseño de Experimentos – Diseños de un Solo Factor Categórico

Resumen

La selección de *Diseños de un Solo Factor Categórico* sobre el menú de Crear un Diseño crea diseños experimentales para situaciones donde el interés primario se centra sobre la comparación de niveles de un simple factor categórico. El procedimiento puede generar corridas en cada nivel del factor. En adicción, uno o más factores de bloque pueden incluirse en el diseño. El procedimiento es capaz de generar

1. Diseño Completamente al Azar
2. Diseño por Bloques al Azar
3. Diseño por Bloques Balanceados Incompletos (BIB)
4. Diseño Cuadrado Latino
5. Diseño Cuadrado Grego-Latino
6. Diseño Cuadrado Hiper-Grego-Latino

Ejemplo

El ejemplo descrito en este documento fue tomado de Box, Hunter and Hunter (1978). Describen un experimento desarrollado para comparar el efecto de 7 tratamientos (A, B, C, D, E, F, y G) sobre el desgaste de calidad de un material particular. Desafortunadamente, la maquina usada para medir el desgaste solamente puede acomodar 4 muestras durante cualquier corrida. Se esperan diferencias potenciales entre las corridas de la maquina, deberán tratarse las corridas como una variable de bloque ordenada para reducir cualquier confusión potencial de las diferencias de corrida-a-corrida contra las diferencias entre los tratamientos.

Ejemplo StatFolio: doe singlecat.sgp

Creación del Diseño

Para crear un experimento en el cual la meta principal es comparar q niveles de un solo factor categórico, seleccione *Crear Diseño* del menú de diseño de experimentos y complete las cajas de diálogos que se describen abajo.

Caja de Dialogo #1 – Tipo de Diseño

La primera caja de dialogo despliega durante la especificación de la creación del diseño el tipo de diseño a ser creado:

Opciones de Creación de Diseños

Clase de Diseño

- De Cribado
- Superficie de Respuesta
- Mezcla
- Factorial Multinivel
- Arreglos Interno/Externo
- Un Solo Factor Categórico
- Multi-Factor Categóricos
- Componentes de Varianza (jerárquicos)

No. de Variables de Respuesta:

No. de Factores Experimentales:

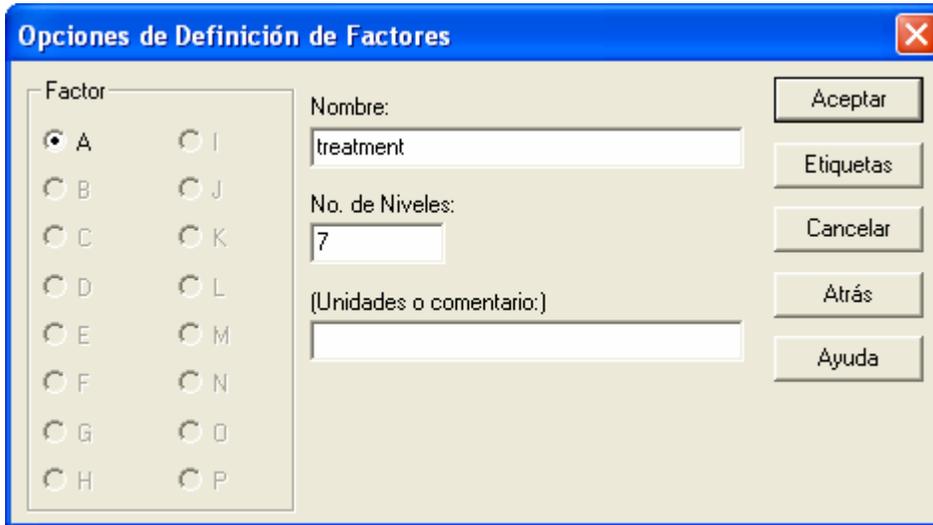
Comentario:

Aceptar
 Cancelar
 Ayuda

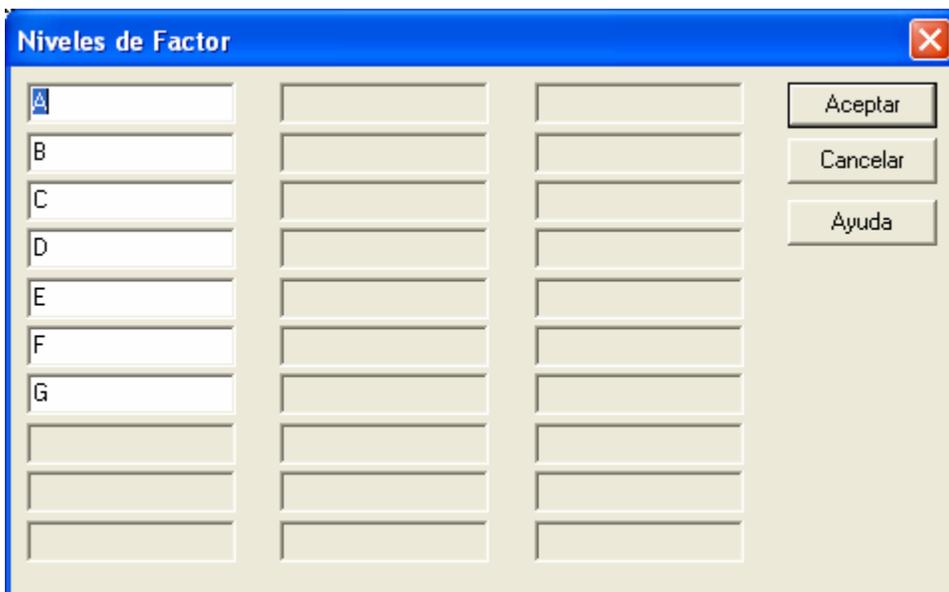
- **Clase de Diseño:** Tipo de diseño a ser creado.
- **No. Variables Respuestas:** El número de variables respuestas Y que deberán medirse durante cada corrida experimental. Este numero esta en un rango de 1 a 16.
- **Comentario:** Un comentario que aparecerá sobre las salidas de los procedimientos de los análisis.

Caja de Dialogo #2 – Factor Experimental

La segunda caja de dialogo requiere información acerca del factor experimental a ser estudiado:



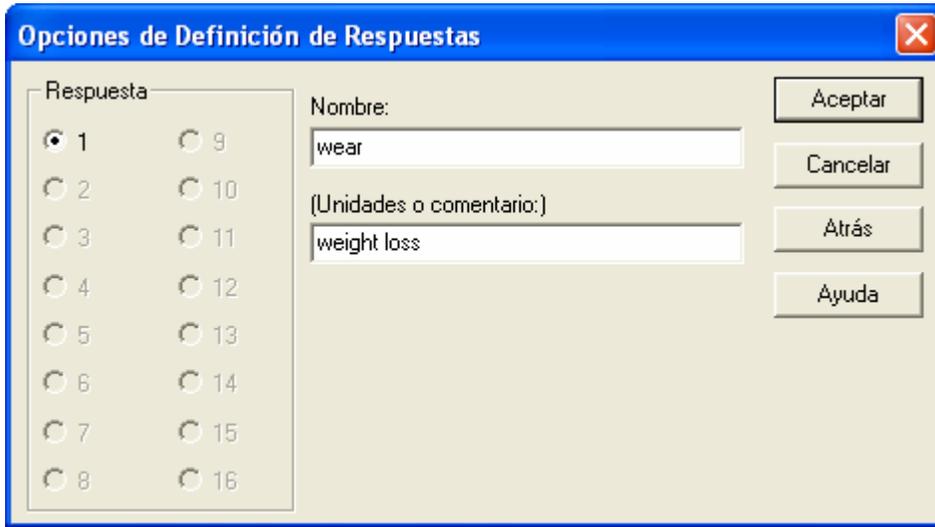
- **Nombre:** Ingrese un nombre para el factor que contenga hasta 32. Una columna puede crearse en la base de datos con el nombre indicado.
- **No. de Niveles:** Número de diferentes niveles del factor en el cuál los experimentos serán desarrollados.
- **Unidades o Comentario** – Una etiqueta opcional o un comentario hasta 64 caracteres que se incluyen sobre la hoja de trabajo experimental.
- **Botón Etiquetas:** Presione este botón para ingresar identificadores de cada nivel:



Si las etiquetas no son especificadas, los niveles del factor pueden numerarse del 1 hasta q.

Dialogo #3 – Variables Respuesta

La tercera caja de dialogo requiere información acerca de cada una de las variables respuestas:

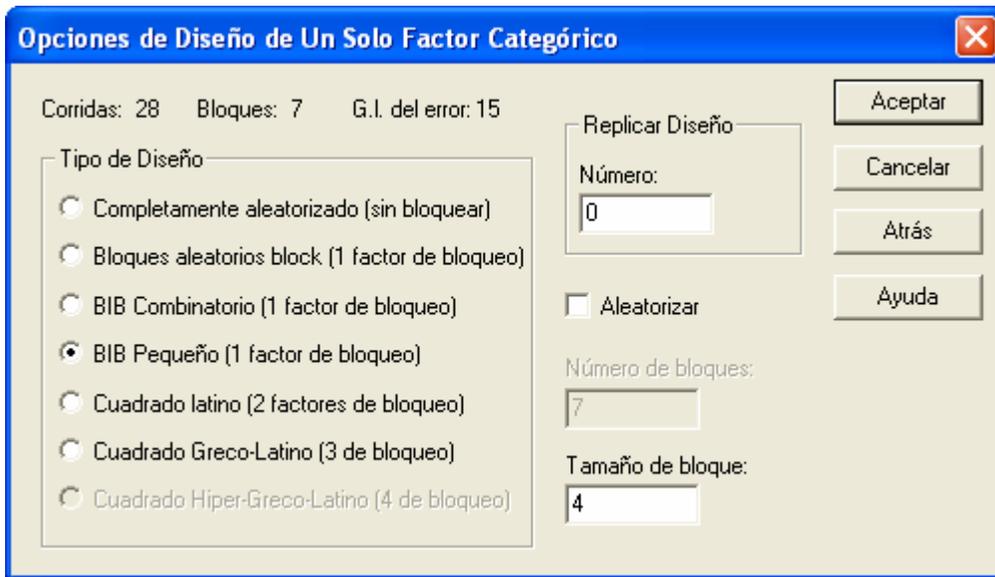


Clic sobre los números 1, 2, 3,..., una vez en el tiempo e ingrese la siguiente información para cada variable respuesta en el experimento:

- **Nombre** – Un nombre para cada respuesta conteniendo hasta 32 caracteres.
- **Unidades o Comentario** – Una etiqueta opcional o comentario hasta 64 caracteres que se incluyen sobre la hoja de trabajo experimental.

Dialogo #4 – Selección del Diseño

La cuarta caja de dialogo es usada para especificar el tipo de diseño a ser creado:



- **Tipo de Diseño:** Los siguientes tipos de diseños están disponibles, dependiendo sobre el numero de niveles del factor experimental:
 1. *Diseño Completamente al Azar* - Es un diseño en el cuál una muestra aleatoria de mediciones es tomada de cada nivel q , sin la tentativa de explicar los efectos de cualquier otro factor.

2. *Diseño por Bloques al Azar* - Un diseño en el cual un número igual de observaciones es tomada de cada tratamiento a dos o más niveles de un factor de bloque o ruido. El efecto de Bloque será incluido en el modelo para reducir la magnitud del error experimental.
 3. *BIB Combinatorio* – Un diseño por bloques balanceados incompletos contiene una sola variable de bloque donde el numero de tratamientos en cada bloque es menor que q . Si k tratamientos deben correrse en cada bloque, el diseño requiere $\binom{q}{k}$ bloques, los cuales representan el número de alternativas al seleccionar k artículos fuera de q .
 4. *BIB Reducido* – Un diseño por bloques balanceados incompletos en el cual el numero de bloques es menor lo requerido por un *BIB Combinatorio* completo. Estos diseños están solamente disponibles para ciertas combinaciones del número de niveles del factor y el tamaño de bloque.
 5. *Cuadrado Latino* - Es un diseño en el cual los tratamientos están balanceados a través de dos factores de bloque, pero donde el número de corridas es menor que las requeridas por un diseño Multi-factorial.
 6. *Cuadrado Greco-Latino* – Es un diseño en el cual los tratamientos están balanceados a través de tres factores de bloques.
 7. *Cuadrado Hiper-Greco-Latino* - Es un diseño en el cuál los tratamientos están balanceados a través de cuatro factores de bloques.
- **Replica del Diseño** – El número de observaciones adicionales que son tomadas de cada nivel de tratamiento o combinación de bloque-tratamiento.
 - **Aleatorización** – Con o sin aleatorizar el orden de las corridas en el experimento.
 - **Número de Bloques** – Para definir en el diseño uno o más factores de bloque, el número de bloque.
 - **Tamaño de Bloque** – Para diseño BIB, el número de tratamientos que serán estudiados en cada bloque.

Basándose en el diseño seleccionado, la caja de dialogo calcula y despliega el numero total de corridas (prueba) a ser desarrolladas, el numero de bloques, y los grados de libertad que estarán disponibles para estimar el error experimental. Note que los grados de libertad son calculados asumiendo que los factores de bloque no interactúan con el factor principal experimental, el cuál es el supuesto más usual.

La caja de dialogo anterior requiere un diseño BIB capaz de estudiar cuatro tratamientos en un solo bloque. Como se indica, este diseño esta disponible en 7 bloques (corridas de la maquina). Un total de 28 pruebas pueden desarrollarse, sabiendo que cada uno de los 7 materiales deberá incluirse dentro de 4 bloques.

Atributos del Diseño

Puesto que la caja de dialogo tienen que ser completadas, la ventana de *Atributos del Diseño* será presentado:

Single Factor Categorical Design Attributes
 Clase de diseño: Factor Categóricio Individual
 Nombre del archivo: C:\DocData\weartest.sfx

Diseño Base
 Número de factores experimentales: 1
 Número de bloques: 7
 Número de respuestas: 1
 Número de corridas: 28
 Grados de libertad para el error: 15
 Aleatorizar: No

Factores	Niveles	Unidades
treatment	7	

Respuestas	Unidades
wear	weight loss

Este contiene un resumen del diseño generado. Para guardar el diseño, seleccione *Grabar Archivo de Diseño* del menú principal *Archivo*. El ejemplo fue grabado en el archivo *weartest.sfx*.

Hoja de Trabajo

Las corridas experimentales pueden ingresarse automáticamente en la base de datos. Esto también puede desplegarse en una Hoja de Trabajo de abajo:

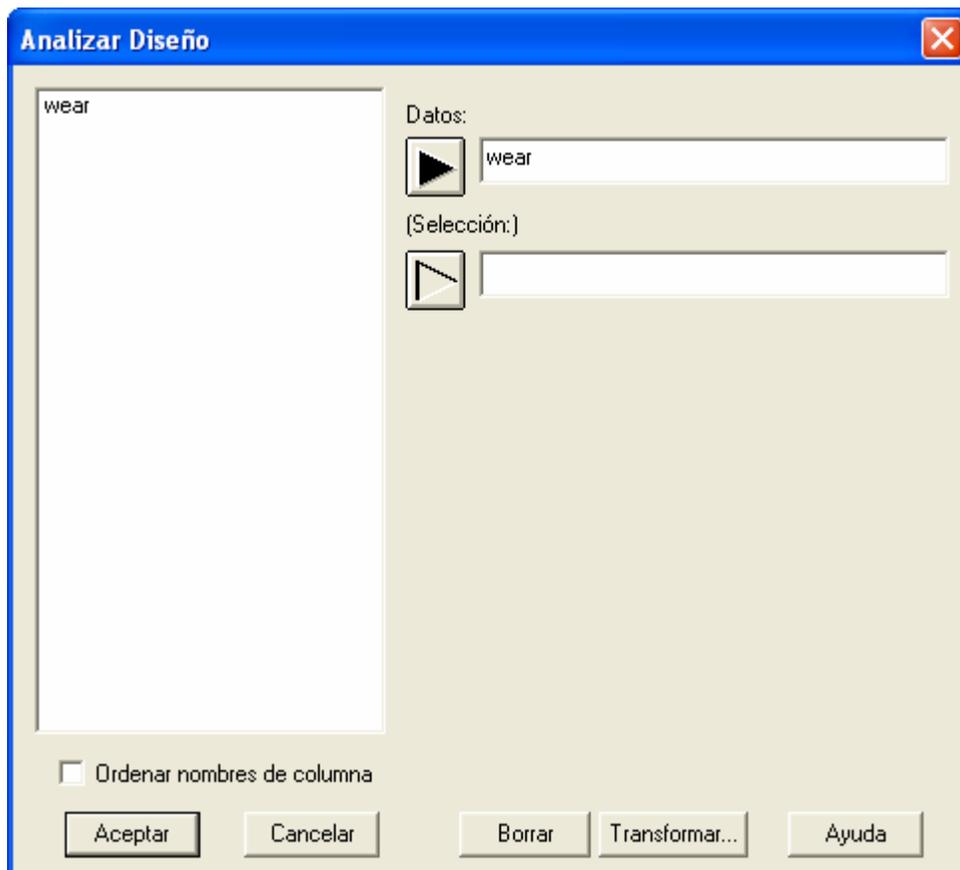
Hoja de trabajo para C:\DocData\weartest.sfx

corrida	bloque	treatment	wear
1	1	C	
2	1	D	
3	1	F	
4	1	G	
5	2	A	
6	2	B	
7	2	F	
8	2	G	
9	3	B	
10	3	D	
11	3	E	
12	3	G	
13	4	A	
14	4	C	
15	4	E	
16	4	G	
17	5	B	
18	5	C	
19	5	E	
20	5	F	
21	6	A	
22	6	D	
23	6	E	
24	6	F	
25	7	A	
26	7	B	
27	7	C	
28	7	D	

Examinado detalladamente la hoja de trabajo, podemos notar que cada tratamiento aparece 4 veces. En adición, cada par de tratamientos ocurre siempre en el mismo bloque una vez. Esto encuentra un diseño más balanceado así como sin posibles corridas para cada tratamiento en cada bloque.

Analizando los Datos

DEspués de que son ingresados los resultados de las corridas experimentales, selecciona Analizando Datos del menú DDE. Una caja de dialogo puede presentarse requiriendo la columna que contiene la respuesta a ser analizada:



- **Datos:** Columna que contiene los valores de la variable respuesta a ser analizada.
- **Selección:** Selección de un subconjunto de los datos.

Cuando OK es presionado, el programa invoca alguno de los dos procedimientos:

1. El procedimiento ANOVA Uni-factor es para un diseño completamente al azar sin variables de bloque.
2. El procedimiento ANOVA *Multi-factor* es para los diseños que contienen una o más variables de bloques.

Para detalles completos de estos análisis se encuentran contenidos en la documentación correspondiente.

Por interés particular en el ejemplo actual tenemos diversas tablas y gráficas:

Tabla de ANOVA

Esta tabla es utilizada para juzgar si no existe diferencia estadística significativa entre los niveles del factor experimental:

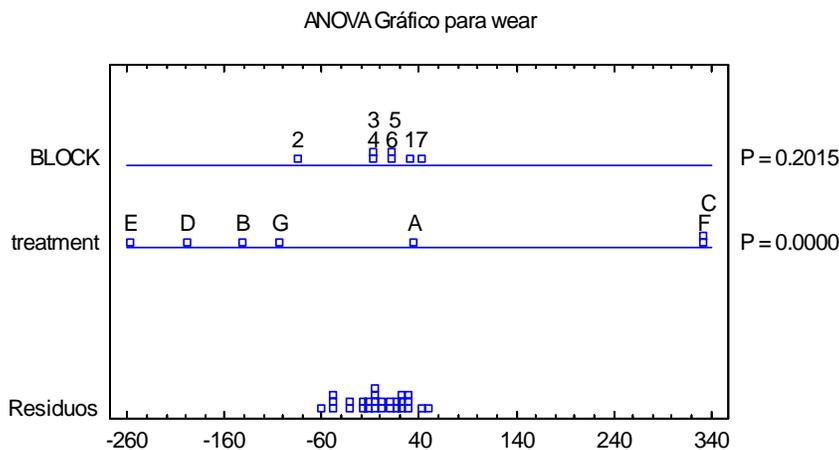
Análisis de Varianza para wear - Tipo III Suma de Cuadrados					
<i>Fuente</i>	<i>Suma de Cuadrados</i>	<i>Gl</i>	<i>Cuadrado Medio</i>	<i>Razón-F</i>	<i>Valor-P</i>
EFFECTOS PRINCIPALES					
A:treatment	506799.	6	84466.4	57.40	0.0000
B:BLOCK	14570.1	6	2428.35	1.65	0.2015
RESIDUOS	22071.4	15	1471.43		
TOTAL (CORREGIDO)	626265.	27			

Todas las razones-F se basan en el cuadrado medio del error residual

Un Valor-P pequeño del factor A (menor que 0.05 si esta en operación el nivel de significancia del 5%) indica que existen diferencias significativas entre los tratamientos. En el ejemplo actual, las diferencias son altamente significativas. Es de interés secundario el Valor-P para el BLOQUE. Puesto que el valor-P en la tabla anterior es mayor que 0.05, el efecto del bloque no es estadísticamente significativo, concluyendo que no existen diferencias grandes entre las corridas de la prueba del desgaste para la maquina.

Gráfico del ANOVA

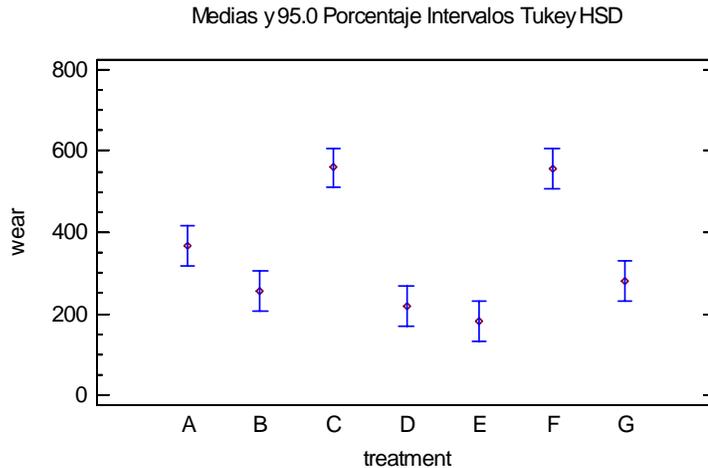
Un nuevo método para ilustrar las diferencias entre bloques y tratamientos, hecho por Hunter (2005), se presenta abajo:



El gráfico muestra las desviaciones escaladas de las medias con respecto de la gran media, junto con los residuales del modelo. El escalamiento significa que, si un factor no tiene efecto, la variación observada por cada factor puede ser comparable con los residuales. . Nótese que la variación sobre los bloques es aceptable más que la variación observada de los residuales, con la posible excepción del bloque #2. También notamos que el tratamiento E es el desgaste mínimo, algunos otros tratamientos son relativamente importantes.

Gráfica de Efectos Principales

El *Gráfico de Efectos Principales* puede ser utilizado para determinar cuales tratamientos tienen diferencias significativas con respecto de otros. Debido al número grande de tratamientos, el gráfico de abajo muestra las medias de los tratamientos con intervalos de Tukey HSD, los cuales permiten al experimentador comparar todos los pares de tratamientos con una tasa de error experimental del 5%:



El tratamiento E muestra el desgaste mínimo con respecto al promedio. Sin embargo, puesto que los intervalos se traslapan con los tratamientos B, D y G, no podemos declarar que no son estadísticamente mejor que cualquiera de los otros 3 tratamientos.