FACULTAD MULTIDISCIPLINARIA DE OCCIDENTE

DEPARTAMENTO DE MATEMÁTICAS

TITULO:

DISEÑOS DE EXPERIMENTOS EN R

Fernando Ernesto Manzanares Morán

Docente:

Salvador Enrique Rodriguez Hernandez

Definición de un diseño experimental y conceptos básicos

El **diseño de un experimento** es la secuencia completa de los pasos que se deben tomar de antemano, para planear y asegurar la obtención de toda la información relevante y adecuada al problema bajo investigación, la cual será analizada estadísticamente para obtener conclusiones válidas y objetivas con respecto a los objetivos planteados.

Un **diseño experimental** es una prueba o serie de pruebas en las cuales existen cambios deliberados en las variables de entrada de un proceso o sistema, de tal manera que sea posible observar e identificar las causas de los cambios que se producen en la respuesta de salida.

Proposito de un diseño experimental

El propósito de cualquier Diseño Experimental, es proporcionar una cantidad máxima de información pertinente al problema que se está investigando. Y ajustar el diseño que sea lo mas simple y efectivo; para ahorrar dinero, tiempo, personal y material experimental que se va ha utilizar. Es de acotar, que la mayoría de los diseños estadísticos simples, no sólo son fáciles de analizar, sino también son eficientes en el sentido económico y en el estadístico. De lo anterior, se deduce que el diseño de un experimento es un proceso que explica tanto la metodología estadística como el análisis económico.

Conceptos básicos

Diseño: Consiste en planificar la forma de hacer el experimento, materiales y métodos a usar, etc.

Experimento: Conjunto de pruebas o ensayos cuyo objetivo es obtener información, que permita mejorar el producto o el proceso en estudio.

Tratamiento: - Es un conjunto particular de condiciones experimentales definidas por el investigador; son el conjunto de circunstancias creadas por el experimento, en respuesta a la hipótesis de investigación y son el centro de la misma.

Factor: Es un grupo específico de tratamientos. (Ejemplo, Temperatura, humedad, tipos de suelos, etc.).

Niveles del factor: Son diversas categorías de un factor. (Por ejemplo, los niveles de temperatura son 20°C, 30°C, etc.). Un factor Cuantitativo tiene niveles asociados con puntos ordenados en alguna escala de medición, como temperatura; mientras que los niveles de un factor cualitativo representan distintas categorías o clasificaciones, como tipo de suelo, que no se puede acomodar conforme a alguna magnitud.

Réplica: Son las repeticiones que se realizan del experimento básico.

Unidad experimental:

Es el material experimental unitario que recibe la aplicación de un tratamiento.

Es la entidad física o el sujeto expuesto al tratamiento independientemente de las otras unidades. La unidad experimental una vez expuesta al tratamiento constituye una sola réplica del tratamiento.

Es el objeto o espacio al cual se aplica el tratamiento y donde se mide y analiza la variable que se investiga.

Es el elemento que se está estudiando.

Unidad muestral: Es una fracción de la unidad experimental que se utiliza para medir el efecto de un tratamiento.

Error experimental: Es una medida de variación que existe entre dos o más unidades experimentales, que han recibido la aplicación de un mismo tratamiento de manera idéntica e independiente.

Factores controlables: Son aquellos parámetros o características del producto o proceso, para los cuales se prueban distintas variables o valores con el fin de estudiar cómo influyen sobre los resultados.

Factores incontrolables: Son aquellos parámetros o características del producto o proceso, que es imposible de controlar al momento de desarrollar el experimento.

Variabilidad natural: es la variación entre las unidades experimentales, que el experimentador no puede controlar ni eliminar.

Variable dependiente: es la variable que se desea examinar o estudiar en un experimento. (Variable Respuesta).

Hipótesis:

Es una suposición o conjetura que se plantea el investigador de una realidad desconocida.

Es el supuesto que se hace sobre el valor de un parámetro (constante que caracteriza a una población) el cual puede ser validado mediante una prueba estadística.

Diseños unifactoriales

En el análisis de los resultados de los experimentos se pueden observar diferentes aplicaciones de los Diseños Experimentales. Hay experimentos muy útiles en los cuales existe un sólo factor de interés; el cual se analiza por medio de la comparación de dos condiciones que intervienen en el Experimento (a menudo llamadas tratamientos o niveles del factor); a este tipo de experimentos se le denomina Experimentos de Comparación Simple, estudiados en los cursos de Estadística básica. El análisis de los datos de este tipo de Experimentos resulta ser sencillo, ya que se utilizan técnicas de la Inferencia Estadística, llamada Prueba de Hipótesis (o pruebas de significancia) que son las que ayudan al experimentador a comparar estas condiciones.

Si en el tipo de Diseño Experimental planteado anteriormente se requiere más de dos niveles del factor que se analiza, éstos son considerados como "Diseños Unifactoriales". Teniendo en cuenta que para el análisis de éstos se utiliza el Análisis de Varianza, ya que se requiere probar la igualdad de varias medias.

En los experimentos de los Diseños Unifactoriales, el número de observaciones recolectadas en cada tratamiento pueden ser iguales o diferentes. Cuando el número de observaciones sea diferente se dice que el Diseño está Desequilibrado o Desbalanceado; en caso contrario el Diseño está Equilibrado o Balanceado.

En este documento se hará uso del caso en el que el Diseño está Equilibrado o Balanceado y se utilizará un modelo de efectos fijos (en este tipo de Modelos se desea probar hipótesis en relación a las medias de los tratamientos y las conclusiones sólo se aplicarán a los niveles del factor considerados en el análisis, las conclusiones no pueden extenderse a tratamientos similares que no se consideraron.) y no se profundizara mas teóricamente ya que el fin de este documento es la aplicación de código del software estadistico R para la aplicación de dicho modelo.

Ejemplo 1.

El Ministerio de Educación esta interesado en implementar tres programas de estudio; con el objetivo de medir la habilidad de lectura en los alumnos. Para ello, se eligen alumnos del sexto grado de un Colegio de San Salvador, de los cuales fueron asignados al azar 27 alumnos, a cada uno de los tres grupos. Se utilizó un programa diferente en cada grupo, se llevó a cabo un examen al inicio y al final de la implementación de los programas, los valores obtenidos representan la diferencia que hay entre la nota del examen que se hizo al inicio y al final de la implementación del programa, obteniéndose los siguientes datos, en base 100:

Tratamiento	Datos								
Programa 1	20	18	18	23	22	17	15	13	21
Programa 2	15	20	13	12	16	17	21	15	13
Programa 3	12	15	18	20	18	17	10	24	16

Cuadro 1: Diferencias entre nota de inicio y nota final

Solución. Antes de realizar los cálculos matemáticos, se definirá la variable de estudio y las hipótesis que se desean probar.

Variable de estudio: Habilidad de Lectura.

Hipótesis:

$H_0 : \mu_1 = \mu_2 = \mu_3$ (No existe diferencia entre los grupos)

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$ (Existe diferencia entre los grupos)

El significado verbal de las hipótesis es:

H_0 : Con la implementación de los tres programas de estudio, no existe diferencia significativa en la habilidad de lectura entre los grupos de alumnos del sexto grado.

H_1 : Con la implementación de los tres programas de estudio, existe diferencia significativa en la habilidad de lectura entre los grupos de alumnos del sexto grado.

```
> library(xtable)
> #Leemos la hoja de datos
> datos<-read.csv("Ej1HL.csv", header = TRUE, sep = ";")
> #Analysis of Variance(aov) para un factor
> unifact<- aov(HL~Programa, data = datos)
> Resultado<-summary(unifact)
> #su equivalente
> Resultado2<- anova(unifact)
> #genarando el codigo latex para la tabla
> # de ANOVA
> Resultado3<- xtable(Resultado2)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Programa	2	36.22	18.11	1.44	0.2566
Residuals	24	301.78	12.57		

Cuadro 2: Anova de un factor

Conclusión

Si se observa la columna Pr(>F) ésta muestra el P-valor generado por el ANOVA y se interpreta de la siguiente manera, como el nivel de significancia de la prueba F con 2 grados de libertad en el numerador y 24 en el denominador es de del 5 por ciento y el P-valor es mayor a este, no hay sufciente evidencia estadistica para rechazar la hipostesis nula y puede concluirse que con la implementación de los tres programas de estudio, no existe diferencia significativa en la habilidad de lectura entre los grupos de alumnos del sexto grado.

Si al efectuar el Análisis de Varianza para un Modelo de Efectos Fijos, la hipótesis nula es rechazada. Se llega a la conclusión que existe diferencia entre las medias o que hay diferencia entre los tratamientos. En muchas situaciones en la industria, este resultado es de poco interés; ya que no se especifica exactamente cuales tratamientos son diferentes y el experimentador espera hallar diferencias, y está más interesado en investigar que tratamientos difieren entre si, o dicho de otra manera, en investigar contrastes entre los tratamientos.

Cuando se da esta situación puede ser útil realizar comparaciones adicionales entre grupos de medias de los tratamientos. Las comparaciones entre medias de tratamientos se realizan en términos de los totales de tratamientos $y_{i.}$ o de los promedios de tratamientos $\bar{y_{i.}}$. Los procedimientos para efectuar esta comparación se conocen como Métodos de Comparación Múltiple o Pruebas a Posteriori.

Esto quiere decir que si no existe diferencia significativa entre la media de los tratamientos no tiene sentido aplicar pruebas posteriori, sin embargo por efectos de la utilización del código en R, se aplicará la prueba de tukey ("TukeyHSD.$^{\text{en}}$ R) cuya hipótesis nula es: "No esxiste difencia significativa entre las comparaciones".

Estimacion de las medias de los tratamientos:

```
> library(ggplot2)
> library(gplots)
> boxplot(datos$HL~datos$Programa,
+         ylab = "Habilidad de lectura",
+         xlab = "Tratamientos",
+         col=rainbow(5, alpha = .3))
> plotmeans(datos$HL~datos$Programa,
+           ylab = "Habilidad de lectura",
+           xlab = "Tratamientos")
```

Evaluacion de las diferencias entre los tratamientos:

Prueba de Tukey:

```
> data.Prueba<-TukeyHSD(unifact, ordered = TRUE)
> data.Prueba.result<- data.frame(data.Prueba$Programa)
> Tabla<- xtable(data.Prueba.result)
```

	diff	lwr	upr	p.adj
Programa3-Programa2	0.89	-3.29	5.06	0.86
Programa1-Programa2	2.78	-1.40	6.95	0.24
Programa1-Programa3	1.89	-2.29	6.06	0.51

Cuadro 3: TukeyHSD

Puede observarse que el P-valor en la columna P.adj es superior al 5 porciento por lo que se concluye que no existe diferencia significativa entre las medias que se han comparado, el resultado es obvio dado que en el analisis de varianza se conluyo que no existia diferencia significativa entre las medias de los tratamientos.

Ejemplo 2.

Se supone que la cantidad de carbón usada en la producción de acero tiene un efecto en su resistencia a la tensión. En la tabla se muestran los valores de la resistencia a la tensión del acero para cada uno de los 4 diferentes porcentajes de carbón. Con estos datos efectúe el análisis apropiado e interprete sus resultados.

Variable de estudio: Resistencia a la tensión.

Hipótesis:

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ (No existe diferencia entre los grupos)

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$ (Existe diferencia entre los grupos)

El significado verbal de las hipótesis es:

H_0 : La cantidad de carbón usada en la producción de acero no tiene efecto significativo en la resistencia a la tensión

H_1 : La cantidad de carbón usada en la producción de acero tiene efecto significativo en la resistencia a la tensión.

% de Carbon	Observaciones			
0.10	23	28	28	30
0.20	31	29	36	38
0.30	36	40	42	44
0.40	48	45	40	40

Cuadro 4: Resistencia a la Tensión

```
> FCarbon<-c("0.10%","0.10%","0.10%",
+            "0.10%","0.20%","0.20%",
+            "0.20%","0.20%","0.30%",
+            "0.30%","0.30%","0.30%",
+            "0.40%","0.40%","0.40%",
+            "0.40%")
> Nivel<-factor(FCarbon)
> Datos<-c(23,28,28,30,31,29,36,38,36,40,42,44,48,45,40,40)
> Unifact<-aov(Datos~Nivel)
> Result<-summary(Unifact)
> Tabla <- xtable(Result)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Nivel	3	622.25	207.42	15.41	0.0002
Residuals	12	161.50	13.46		

Cuadro 5: Anova de un factor

Conclusión.

Se concluye que la cantidad de carbón usada en la producción de acero tiene efectos significativos en la resistencia a la tensión. Como se ha rechazado H_0, existe diferencia entre las medias de los tratamientos, pero no se especifica entre que medias de tratamientos existen las diferencias. Se podría estar interesado en querer saber entre que medias de tratamientos existe diferencia, para ello se utilizará el método de Tukey y LSD para contestar esta inquietud.

Estimacion de las medias de los tratamientos:

```
> library(ggplot2)
> library(gplots)
> boxplot(Datos~FCarbon,
+         ylab = "Resistencia a la tensión",
+         xlab = "Porcentaje de Carbón",
+         col=rainbow(5, alpha = .3))
> plotmeans(Datos~FCarbon,
+           ylab = "Resistencia a la tensión",
+           xlab = "Porcentaje de Carbón")
```

Evaluacion de las diferencias entre los tratamientos:

Prueba LSD:

```
> library(agricolae)
> data.Prueba1<- LSD.test(Unifact, "Nivel" )
> data.Prueba1.result1<- data.frame(data.Prueba1$statistics)
> data.Prueba1.result2<- data.frame(data.Prueba1$means)
> data.Prueba1.result3<- data.frame(data.Prueba1$groups)
> tabla1 <- xtable(data.Prueba1.result1)
```

```
> tabla2 <- xtable(data.Prueba1.result2)
> tabla3 <- xtable(data.Prueba1.result3)
```

Mean	CV	MSerror	LSD
36.12	10.16	13.46	5.65

Cuadro 6: Calculo del LSD

Porcentaje de carbón	Means	std	r	LCL	UCL	Min	Max
0.10 %	27.25	2.99	4	23.25	31.25	23.00	30.00
0.20 %	33.50	4.20	4	29.50	37.50	29.00	38.00
0.30 %	40.50	3.42	4	36.50	44.50	36.00	44.00
0.40 %	43.25	3.95	4	39.25	47.25	40.00	48.00

Cuadro 7: Estadisticos de los Tratamientos

	trt	means	M
1	0.40 %	43.25	a
2	0.30 %	40.50	a
3	0.20 %	33.50	b
4	0.10 %	27.25	c

Cuadro 8: Comparaciones

Las conclusiones principales se toman apartir de el cuadro 8, de manera que si se observa la columna ”M”se deben hacer las combinaciones de las letras y éstas diferiran en media significativamente si en las combinaciones las letras no se repiten, es decir, la media en la resistencia a la tensión según en el porcentaje de carbón para los grupos 1 y 2 (30 % y 40 %) no difiere significativamente, sin embargo para las demas combinaciones 1 y 3, 1 y 4, 2 y 3, 2 y 4, 3 y 4, si difieren significativamente en la media de la resistencia, Para reforzar esta conlclusión se realizara la prueba de tukey.

Prueba de Tukey:

```
> data.Prueba<-TukeyHSD(Unifact, ordered = TRUE)
> data.Prueba.result<- data.frame(data.Prueba$Nivel)
> Tabla<- xtable(data.Prueba.result)
```

Diferencias	diff	lwr	upr	p.adj
0.20 %-0.10 %	6.25	-1.45	13.95	0.13
0.30 %-0.10 %	13.25	5.55	20.95	0.00
0.40 %-0.10 %	16.00	8.30	23.70	0.00
0.30 %-0.20 %	7.00	-0.70	14.70	0.08
0.40 %-0.20 %	9.75	2.05	17.45	0.01
0.40 %-0.30 %	2.75	-4.95	10.45	0.72

Cuadro 9: TukeyHSD

Y efectivamente puede observarse que en la columna P.adj la comparación entre los grupos con 40 % y 30 % de carbón no difieren significativamente en media (igual al LSD) sin embargo en los grupos 20 % y 10 % y 30 % y 20 % de carbón dado que el P.adj es mayor a 0.05 no existe diferencia significativa en las medias de estas comparaciones y si existe diferencia en las demas combinaciones, en definitiva puede decirse que el la prueba de tukey es mas precisa que el la prueba LSD dado que esta reporto un grupo más que no difieren significativamente en media.

Diseños por Bloques

En todo Diseño de Experimento se desea que el Error Experimental sea lo más pequeño posible, ya que refleja tanto el Error aleatorio como la variabilidad entre las unidades experimentales o grupos de Unidades Experimentales; es decir, se trata de sustraer del Error Experimental la variabilidad producida por las unidades experimentales o grupos de Unidades Experimentales.

Para lograr lo anterior, las unidades experimentales se deben distribuir aleatoriamente en grupos o bloques, de tal manera que dentro de un bloque sean relativamente homogéneas y que el número de ellas dentro de un bloque sea igual al número de tratamientos por investigar. A esta forma de organizar y distribuir las unidades Experimentales se le denomina **Diseño por Bloques.**

El principio de bloqueo se basa en que las unidades experimentales dentro de cada bloque o grupo deben ser parecidas entre si (que exista homogeneidad dentro del bloque) y que los bloques debieran ser diferentes entre si (que exista heterogeneidad entre los bloques); y además que las unidades experimentales se deben distribuir aleatoriamente en los grupos o bloques de tal manera que dentro de un bloque sean relativamente homogéneas; lo que significa que el bloqueo o agrupamiento del material experimental debe hacerse de tal forma que las unidades experimentales dentro de un bloque sean tan homogéneas como sea posible y que las diferencias entre las unidades experimentales sean explicadas, en mayor proporción, por las diferencias entre los bloques. Esta técnica es una generalización de la prueba de comparación de medias de muestras apareadas, donde cada pareja es un bloque. Para la validez del análisis, se supone que no existe interacción entre los bloques y los tratamientos. En experimentos de campo, cuando no se tenga información segura sobre la uniformidad del terreno, siempre es preferible introducir bloques, que servirán para separar la variabilidad debido a dicha heterogeneidad.

Los Diseños que cumplen estas características y que son objeto de estudio, se clasifican en:

1. Diseños Aleatorizados por Bloques Completos.
2. Diseños Aleatorizados por Bloques Incompletos Balanceados.

En este documento se hara referencia unicamente a los Diseños Aleatorizados por Bloques Completos este Diseño se da cuando cada bloque tiene tantas unidades Experimentales como Tratamientos, y todos los tratamientos son asignados al azar dentro de cada bloque. En este tipo de Diseño la palabra ”Completo” significa que todos los tratamientos son probados en cada bloque, y los bloques forman una unidad experimental más homogénea con la cual se comparan los tratamientos, y ”Aleatorizado” porque dentro de cada bloque los tratamientos son asignados de forma aleatoria a las unidades Experimentales; es decir, que todas las Unidades Experimentales de un mismo bloque tienen la misma probabilidad de recibir cualquiera de los tratamientos. Esta estrategia de Diseño mejora efectivamente la precisión en las comparaciones al eliminar la variabilidad en los tratamientos.

Ejemplo 1.

Se probaran 5 raciones respecto a sus diferencias en el engorde de novillos. Se dispone de 20 novillos para el experimento, que se distribuyen en 4 bloques (5 novillos por bloque) con base a sus pesos, al iniciar la prueba de engorde. Los 5 tratamientos (raciones) se asignaron al azar dentro de cada bloque. Los novillos más pesados se agruparon en un bloque, en otro se agruparon los 5 siguientes más pesados y así sucesivamente. Se obtuvieron los siguientes datos:

```
> library(xtable)
> #Leemos la hoja de datos
> datos<-read.csv("Ej1_B.csv", header = TRUE, sep = ";")
> #Analysis of Variance(aov) para un factor
> DBloques<- aov(datos$Observaciones ~ datos$Raciones + datos$Bloques)
> Resultado<-summary(DBloques)
> #su equivalente
> Resultado2<- anova(DBloques)
> #genarando el codigo latex para la tabla
> # de ANOVA
> Resultado3<- xtable(Resultado2)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Raciones	4	30.71	7.68	39.11	0.0000
Bloques(novillos)	3	0.46	0.15	0.78	0.5257
Residuals	12	2.36	0.20		

Cuadro 10: Análisis de Varianza

Conclusión

Las cinco raciones no son igualmente efectivas en el engorde de novillos o la cantidad de ración influye significativamente en el engorde de los novillos, otro punto importante que puede observarse es que los bloques (novillos) no son significativos, es decir no hay un efecto significativo de los bloques hacia el factor, o dicho de otra manera separar a los novillos en bloques no reduce significativamente la variabilidad, en palabras sencillas, dado que los bloques no son significativos, da igual utilizar un diseño unifactorial.

Como la hipotesis nula fue rechazada es necesario determinar que medias son significativamente diferentes; para observar esas diferencias entre las medias se utilizará la prueba LSD y tukey.

Estimacion de las medias de los tratamientos:

```
> library(ggplot2)
> library(gplots)
> boxplot(datos$Observaciones~datos$Raciones,
+         ylab = "Pesos",
+         xlab = "Raciones",
+         col=rainbow(5, alpha = .3))
> plotmeans(datos$Observaciones~datos$Raciones,
+           ylab = "Pesos",
+           xlab = "Raciones")
```

Prueba LSD:

```
> library(agricolae)
> data.Prueba1<- LSD.test(DBloques, "datos$Raciones")
> data.Prueba1.result1<- data.frame(data.Prueba1$statistics)
> data.Prueba1.result2<- data.frame(data.Prueba1$means)
> data.Prueba1.result3<- data.frame(data.Prueba1$groups)
> tabla1 <- xtable(data.Prueba1.result1)
> tabla2 <- xtable(data.Prueba1.result2)
> tabla3 <- xtable(data.Prueba1.result3)
```

Mean	CV	MSerror	LSD
2.21	20.10	0.20	0.68

Cuadro 11: Calculo del LSD

Raciones	Means	std	r	LCL	UCL	Min	Max
R1	1.50	0.58	4	1.02	1.98	0.90	2.30
R2	3.85	0.57	4	3.37	4.33	3.20	4.50
R3	0.70	0.23	4	0.22	1.18	0.50	0.90
R4	3.50	0.20	4	3.02	3.98	3.20	3.60
R5	1.48	0.43	4	0.99	1.96	0.90	1.80

Cuadro 12: Estadisticos

	trt	means	M
1	R2	3.85	a
2	R4	3.50	a
3	R1	1.50	b
4	R5	1.48	b
5	R3	0.70	c

Cuadro 13: Comparaciones

Conclusión

Puede observarse que las combinaciones 1 y 2, 3 y 4 no difieren significativamente en media y las combinaciones 1 y 3, 1 y 4, 1 y 5, 2 y 3, 2 y 4, 2 y 5, 3 y 5, 4 y 5, difieren significativamente en media, en el contexto las raciones que se le den al novillo si afectan en su peso.

Prueba de Tukey:

```
> data.Prueba<-TukeyHSD(DBloques, ordered = TRUE)
> data.Prueba.result<- data.frame(data.Prueba$`datos$Raciones`)
> Tabla<- xtable(data.Prueba.result)
```

		diff	lwr	upr	p.adj
1	R5-R3	0.78	-0.22	1.77	0.16
2	R1-R3	0.80	-0.20	1.80	0.14
3	R4-R3	2.80	1.80	3.80	0.00
4	R2-R3	3.15	2.15	4.15	0.00
5	R1-R5	0.02	-0.97	1.02	1.00
6	R4-R5	2.03	1.03	3.02	0.00
7	R2-R5	2.38	1.38	3.37	0.00
8	R4-R1	2.00	1.00	3.00	0.00
9	R2-R1	2.35	1.35	3.35	0.00
10	R2-R4	0.35	-0.65	1.35	0.79

Conclusión

Puede observarse que en las combinaciones 1,2,5,10 no existe diferencia significativa en la media del peso de los novillos.