NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**INTERVALOS DE CONFIANZA DE MEDIAS EN LA EXPERIMENTACIÓN**	**PRÁCTICA NÚMERO**	**1**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

REQUERIMIENTOS PARA REALIZACION DE PRÁCTICAS EDUCATIVAS EN LABORATORIOS DE LA FIE

SOFTWARE REQUERIDO	
MINITAB	
OBSERVACIONES-COMENTARIOS	
NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

La presente práctica aborda la estadística inferencial como punto de partida para el diseño de experimentos, y es que al hacer inferencias sobre la media de una población y sus características, en realidad se está experimentando.

En esta práctica se determinará el intervalo esperado de la media de una población partiendo de dos casos, conociendo la desviación estándar y desconociéndola. Lo anterior utilizando el software Minitab.

2.- OBJETIVO (COMPETENCIA):

Analizar una población a través de sus datos y la estadística inferencial de la media, utilizando las herramientas estadísticas del software Minitab.

3.- TEORIA:

Intervalos De Confianza

La Estimación Puntual, consiste en encontrar estadísticos y sus propiedades para hacer la mejor aproximación numérica de un parámetro. Hemos visto que los estadísticos muéstrales más comunes como la media $\overline{X}$, la varianza S^2 sólo dan un valor numérico posible del parámetro correspondiente μ y σ^2 respectivamente, es decir $\overline{X}$ y S^2 son estimadores puntuales de μ y σ^2. Sin embargo, en la vida real, en el proceso de toma de decisiones, es deseable contar con un rango de posibles valores que pueden tomar estos parámetros, es decir un intervalo.

Dada la información de una muestra aleatoria, **La Estimación por Intervalos** de parámetros, consiste en encontrar estadísticos que representen los límites inferior y superior de los posibles valores que éstos pueden tomar con un nivel de probabilidad establecido antes de sacar la muestra. Para lograr calcular dichos límites, es necesario conocer la distribución de probabilidad de los estimadores o funciones de éstos y así establecer el nivel de probabilidad, la cual es convertida en el ámbito de confianza en el momento que los estadísticos que representan los límites son reemplazados con los valores muestrales.

3.1 Intervalo De Confianza Para μ, Cuando σ Es Conocida.

Considere una variable aleatoria X, con una muestra aleatoria de tamaño n, el valor de $\overline{X}$ (media muestral) se puede usar para estimar μ (que puede ser desconocido) por ejemplo con un nivel de confianza de 95% de la siguiente manera:

$$P(\mu - E \leq \overline{X} \leq \mu + E)$$

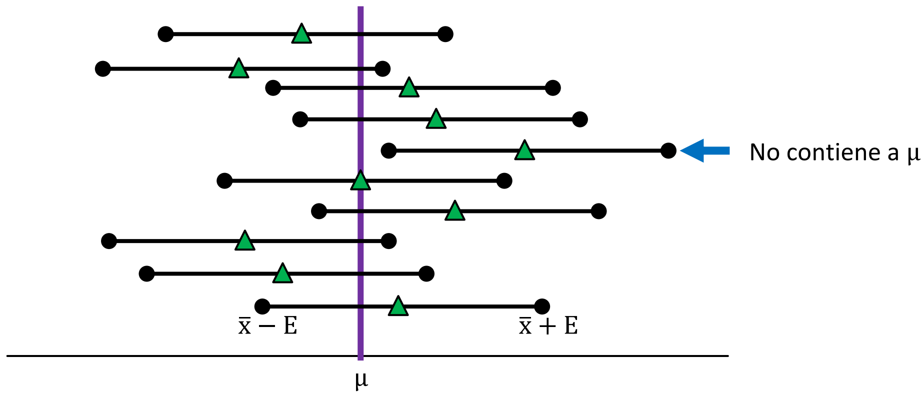
O su equivalente:

$$P(\overline{X} - E \leq \mu \leq \overline{X} + E)$$

Donde el margen de error o simplemente error E, es la diferencia entre la media poblacional y la media muestral, es decir: $E = \overline{X} - \mu$

A $(\overline{X} - E, \overline{X} + E)$ se le llama intervalo de confianza aleatorio de μ con nivel de confianza del 95 %.

En la siguiente figura se muestra que algunos intervalos $[\overline{x} - E, \overline{x} + E]$ contendrá a $\boldsymbol{\mu}$. Sin embargo habrá algunos intervalos que no contendrán a la verdadera $\boldsymbol{\mu}$, entonces, es conveniente especificar con una cierta confianza, cual intervalo si contiene a μ; este **nivel de confianza** ó **γ (gamma)** puede ser por ejemplo una probabilidad de 0.95, por lo tanto a la diferencia a uno (0.05) se le conoce como **nivel de error** o **α (alfa)**, es decir 1-α =γ ó 1-γ=α, por ejemplo 1-0.05 =0.95 ó 1-0.95=0.05 respectivamente.



Problema ejemplo1. Supongamos que X, es una variable aleatoria normal con media μ, cuyo valor desconocemos y desviación estándar σ=2. De una muestra aleatoria con remplazamiento de 25 valores de X, obtenemos una media muestral $\bar{\mathbf{x}}$ =10. Determinar el margen de error E para un intervalo de confianza del 95% para μ y hallar el correspondiente intervalo de confianza.

Partimos por definir la siguiente ecuación, para el intervalo de confianza aleatorio con un nivel de confianza del 95%.

$$P(\mu - E \leq \bar{X} \leq \mu + E) = 0.95$$

En este caso utilizamos la formula de transformación Z, pero como estamos trabajando con medias, dividiremos σ, es decir:

$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ o bien $Z = \frac{E}{\sigma/\sqrt{n}}$ donde: $E = \bar{X} - \mu$

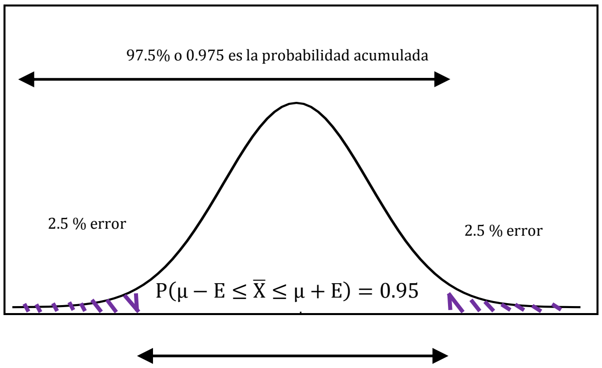
Ahora, el denominador para el problema sería $\sigma/\sqrt{n}$ = 2/5= 0.4

$$P(\mu - E \leq \bar{X} \leq \mu + E) = 0.95$$

Estandarizando se obtiene:

$P\left(-\frac{E}{0.4} \leq \frac{\bar{X}-\mu}{0.4} \leq \frac{E}{0.4}\right) = 0.95$ o bien $P\left(-\frac{E}{0.4} \leq Z \leq \frac{E}{0.4}\right) = 0.95$

Dado que estamos permitiendo un 5 % de error en la probabilidad, es conveniente especificar lo siguiente



Cuando buscamos el 95 % de confianza, implica un 5 % de nivel de error, el cual se reparte 2.5 % a la derecha e izquierda, de tal forma que el error queda a los extremos de la distribución. En estos casos buscamos en la tabla de distribución normal estándar acumulada el valor z correspondiente a una probabilidad de 0.975 y encontramos que z es 1.96, Es decir 1.96 es el valor crítico de z correspondiente a una probabilidad del 0.95, ya repartiendo un 2.5% a cada extremo.

Ahora bien, sustituimos en la formula

$Z = \frac{E}{\sigma/\sqrt{n}}$ donde $1.96 = \frac{E}{0.4}$ por lo tanto $E = (1.96)*(0.4) = 0.784$

$P(\overline{X} - E \leq \mu \leq \overline{X} + E) = 0.95$

$P(10 - 0.784 \leq \mu \leq 10 + 0.784) = 0.95$

$P(9.216 \leq \mu \leq 10.784) = 0.95$

Con esto se tiene un confianza del 95 % de que la media μ de X es algún valor del intervalo 9.216 hasta 10.784; lo que significa que la media de $\overline{x}$ toma los posibles valores $\overline{X}$. En otras palabras el 95 % de los intervalos $[\overline{x} - 0.784, \overline{x} + 0.784]$ contendrá a $\boldsymbol{\mu}$.

3.2 Intervalos De Confianza Para μ cuando σ Es Desconocida.

Cuando la desviación estándar σ de la variable aleatoria X es desconocida, se usan los valores de la desviación estándar muestral.

$$\mathbf{S} = \sqrt{\frac{\sum(\mathbf{x_i} - \overline{\mathbf{X}})^2}{\mathbf{n} - \mathbf{1}}}$$

Supongamos que la variable aleatoria X se distribuye normalmente y tiene como media μ. Sea $\overline{X}$ la media muestral correspondiente a muestras aleatorias de tamaño n, y sea S la correspondiente desviación estándar muestral, entonces se utiliza la variable aleatoria t en lugar de Z:

$$t = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

Con esta variable aleatoria podemos establecer un intervalo de confianza para μ cuando σ no se conoce, tomando su lugar S, y en el caso de Z entra t.

El margen de error se calcula de la siguiente manera:

$$E = \frac{t * S}{\sqrt{n}}$$

Problema Ejemplo 2. La media de una muestra aleatoria de 10 calificaciones de un examen es 75 (con escala de 0 a 100), y la desviación estándar muestral S=8.4; aceptando que el conjunto de todas las calificaciones están distribuidos aproximadamente como normal, hallar un intervalo de confianza del 95% para la calificación media.

Para buscar el valor crítico de t, debemos calcular los grados de libertad (v) mediante n-1 y el nivel de error α, el cual normalmente se divide entre 2 (α/2) para repartir el error en los dos extremos de la distribución.

Entonces: n=10 ∴ grados de libertad v = n-1 = 10-1=9

$t\alpha/2$, n-1 = t 0.025, 9

Buscamos en la tabla de la distribución t y encontramos que el valor critico de t es 2.262 con un α=0.025 y v=9

Calculamos el error E= t $*\frac{S}{\sqrt{n}}$ es decir 2.262 $*\frac{8.4}{\sqrt{10}}$ = 6.009

Calculamos el intervalo de confianza del 95%.

P[75-(2.262 $*\frac{8.4}{\sqrt{10}}$) ≤μ≤75+(2.262 $*\frac{8.4}{\sqrt{10}}$)]=0.95

P[75-6.009 ≤μ≤75+6.009)]=0.95

P[68.99≤μ≤81.009]=0.95

Es decir, con un 95% de confianza, se esperaría que el promedio real de las calificaciones estén entre 69 y 81.

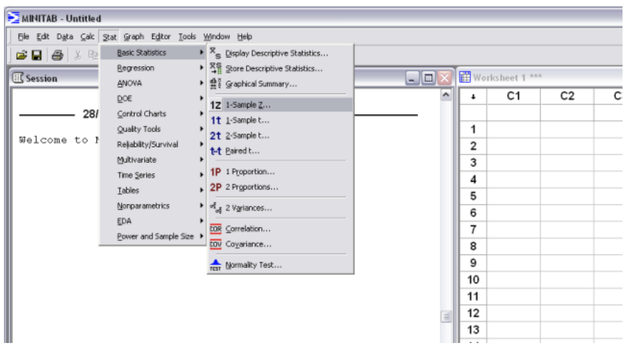
4.- DESCRIPCIÓN

A) PROCEDIMIENTO Y DURACION DE LA PRÁCTICA:

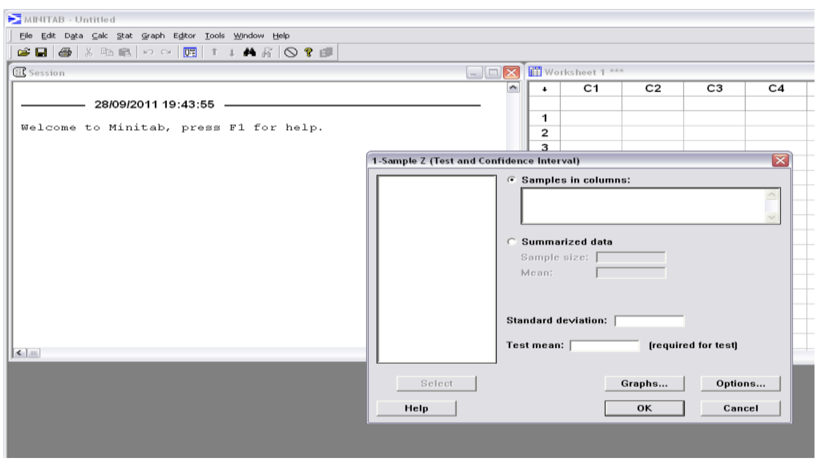
1.- Encienda su computadora y abra el programa Minitab.

2.- Siga la secuencia siguiente.

2.1 Intervalo de confianza de una media poblacional con desviación estándar conocida.



Aparece la siguiente ventana.



Datos de muestra presentes.

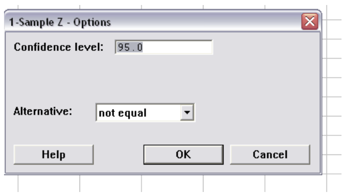
Si se tienen los datos, estos se deben de indicar en la opción de muestras en columnas (**samples in columns**).

Datos resumidos.

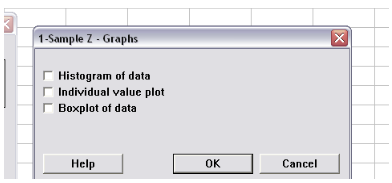
Si sólo se tiene los datos resumidos, es decir, el tamaño de muestra, promedio y desviación estándar, entonces se selecciona la opción de datos resumidos (**Summarized data**)

Ingrese el tamaño de la muestra, la media. Para probar que la media poblacional es igual a cierta variable aleatoria X, indique en **Test mean** el valor de X, con esto Minitab probará la hipótesis nula de que la media poblacional es igual a X.

Seleccione **options** e indique el nivel de confianza deseado. Acontinuación especifique la hipótesis alternativa.



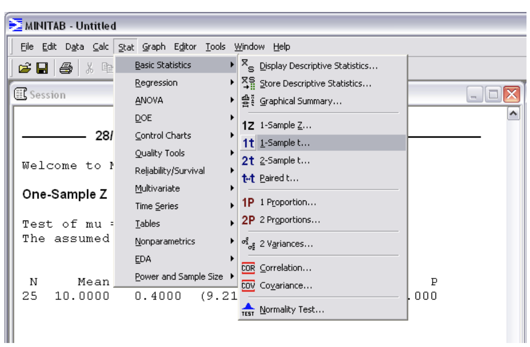
Para ver el comportamiento visual, seleccione **Graphs** e indique el tipo de gráfico que se desea obtener.



Nota: cuando no se tienen los datos, Minitab no puede graficar.

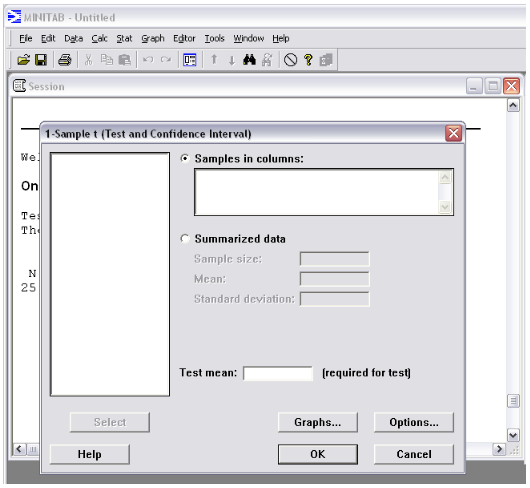
2.2 Intervalo de confianza de la media cuando se desconoce la desviación estándar poblacional.

Seleccione 1 sample t, como se muestra en la siguiente secuencia.

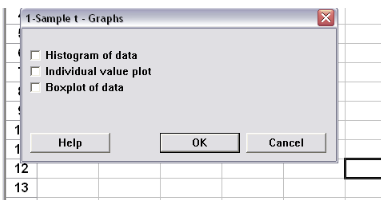


Si los datos están en diferentes columnas, seleccione la opción (**Samples incolumns**).

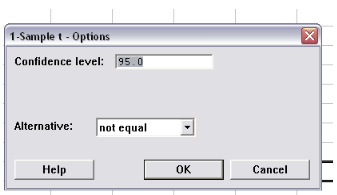
Si solo se tienen datos resumidos, es decir el tamaño de muestra, promedio y desviación estándar de la muestra, seleccione la opción (**Summarized data**). En la sección de **test mean**, indique la prueba de hipótesis de que la media es igual a cierto valor X.



En **graphs**, indique que gráfico desplegará Minitab. Si no se tienen datos, Minitab no podrá arrojar gráficos.



Seleccione **options** para especificar el nivel de confianza deseado, así como la prueba de hipótesis alternativa.



B) DESARROLLO

practique analizando los siguientes problemas

1. Se quiere hacer una estimación por intervalo del contenido neto promedio en gramos de bolsas de azúcar llenadas por una máquina automática. Se toma una muestra aleatoria de 50 bolsas resultando un peso promedio de =112 gramos, la desviación estándar se sabe que es igual a 25 gramos.
   a) Hallar el intervalo de confianza del 85% para la media µ del contenido neto de las bolsas de azúcar.
2. Se realizó un muestreo de 8 trozos de una misma tela para determinar su resistencia a la tensión (Libras/pulg^2). Los resultados fueron los siguientes: 24.4, 18.9, 12.8, 20.5, 19.1, 15.2, 21.7 y 14.6. Si se supone que la población de la que provienen estas resistencias a la tensión es aproximadamente normal, ¿Cuál será el intervalo de confianza del 98% para la media µ?

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
Douglas C. Montgomery (2009). Design and Analysis of Experiments

6.- ANEXOS:

NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**INTERVALOS DE CONFIANZA DE PROPORCIONES EN LA EXPERIMENTACIÓN**	**PRÁCTICA NÚMERO**	**2**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

SOFTWARE REQUERIDO
MINITAB

OBSERVACIONES-COMENTARIOS

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

La presente práctica aborda la inferencia de proporciones de poblaciones a partir de datos muestrales.

En esta práctica se resolverá los intervalos esperados para una proporción utilizando el software Minitab.

2.- OBJETIVO (COMPETENCIA):

Determinar la proporción poblacional de una muestra como medida de estimación, utilizando las herramientas estadísticas del software Minitab.

3.- TEORIA:

3.1 Intervalo De Confianza Para Las Proporciones.

En ciertos fenómenos se halla una población que se divide en dos grupos. Los miembros de uno de esos grupos se llamarán exitosos, donde *p*= a la proporción desconocida de éxitos en una población. Para estos fenómenos, el intervalo de confianza para *p* tendrá la siguiente forma:

$$[\hat{p} - E, \hat{p} + E]$$

Donde $\hat{p}$ es la proporción de éxitos obtenidos en una muestra aleatoria y E es el margen de error.

$\hat{p}$ tiene como media $\mu_{\hat{p}} = p$ y desviacion estandar $\sigma_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

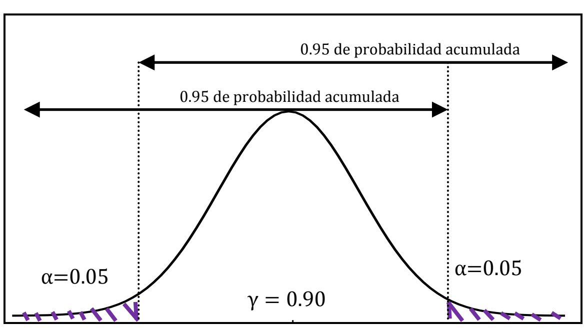
Y se aproxima a una distribución normal cuando n≥30.

$Z = \frac{\hat{p}-p}{\sigma_{\hat{p}}}$ o bien $Z = \frac{E}{\sigma_{\hat{p}}}$ por lo que el margen de error es $E = z * \sigma_{\hat{p}}$

Problema ejemplo 2. En una muestra aleatoria de 900 votantes, el 55% prefiere al candidato demócrata de presidente. Hallar el intervalo de confianza aproximado para la proporción de todos los votantes que prefieren al candidato demócrata con un nivel de confianza de a) 90% y b) 99%.

a) γ=90% y $\hat{p} = 0.55$

Buscamos en la tabla de distribución normal estándar acumulada el valor próximo de z=0.95 que corresponde a 1.65



-z=-1.65 z=1.65

Entonces

$Z = \frac{\hat{p}-p}{\sigma_{\hat{p}}} = \ E = z * \sigma_{\hat{p}}$

$\sigma_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ sustituyendo $\sigma_{\hat{p}} = \sqrt{\frac{0.55(1-0.55)}{900}} = 0.0166$

$E = z * \sigma_{\hat{p}}$ $E = 1.65 * 0.0166 = 0.0274$

El correspondiente intervalo de confianza del 90% es:

[0.55-0.0274, 0.55+0.0274] = 0.90

[0.523≤p≤ 0.577]=0.90

[0.523, 0.577]

4.- DESCRIPCIÓN

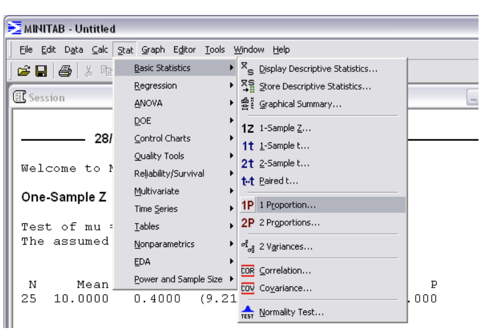
A) PROCEDIMIENTO Y DURACION DE LA PRÁCTICA:

1.- Encienda su computadora y abra el programa Minitab.

2.- Siga la secuencia siguiente.

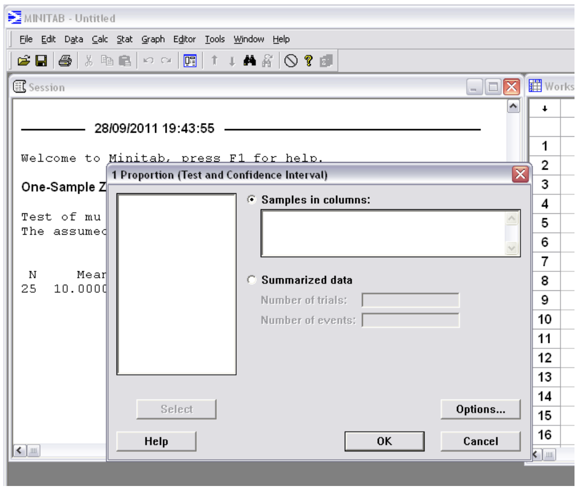
2.1 Intervalo De Confianza Para Proporciones

Seleccione la prueba siguiendo la siguiente secuencia

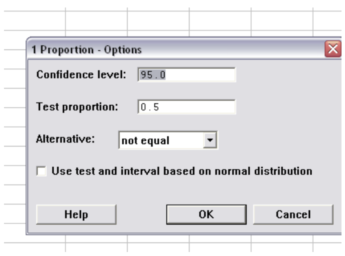


Si se tienen los datos para sacar la proporción, utilice **Samples in columns**.

Si solo se tienen datos resumidos, utilice la opción **summarized data**. En **Trials** deben de ir el número total de intentos y en **events**, el número total de eventos o de éxitos obtenidos



Seleccione el nivel de confianza en la casilla **Options**, posteriormente especifique la hipótesis nula de que la proporción de la población es igual a cierto valor X.



B) DESARROLLO

practique analizando los siguientes problemas

Suponga que una encuesta establece que [0.52,0.57] es el intervalo de confianza del 98% para la proporción de votantes de prefieren al candidato A.

a) ¿Qué porcentaje de la muestra prefieren al candidato A?
b) ¿Cuál es el margen de error?

NOMBRE DE LA MATERIA	DISEÑO DE EXPERIMENTOS	CLAVE	9020

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar el análisis de los roblemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
Douglas C. Montgomery (2009). Design and Analysis of Experiments

6.- ANEXOS:

NOMBRE DE LA PRÁCTICA	**INTERVALOS DE CONFIANZA PARA LA DIFERENCIA DE MEDIAS EN LA EXPERIMENTACIÓN**	**PRÁCTICA NÚMERO**	**3**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

SOFTWARE REQUERIDO	
MINITAB	
EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18
OBSERVACIONES-COMENTARIOS	
NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

La presente práctica aborda la estadística inferencial como punto de partida para el diseño de experimentos, y es que al hacer inferencias sobre la diferencia de medias de dos poblaciones y sus características, en realidad se está experimentando.

En esta práctica se determinará el intervalo esperado para la diferencia de dos medias de dos poblaciones, partiendo de dos casos, datos no pareados y datos pareados. Lo anterior utilizando el software Minitab.

2.- OBJETIVO (COMPETENCIA):

Comparar dos opciones o poblaciones mediante estadística inferencial para la correcta toma de decisiones utilizando las herramientas estadísticas del software Minitab.

3.- TEORIA:

3.1 Intervalo de confianza para las diferencias de medias con varianzas conocidas.

En ciertas ocasiones se requiere determinar si la media poblacional de un grupo de datos es estadísticamente igual a otro grupo, esto es que debido a la variación aleatoria, la $\boldsymbol{\mu_1}$ de un grupo 1, sea aproximadamente igual a la $\boldsymbol{\mu_2}$de un grupo 2, o en su caso determinar que estadísticamente no son iguales.

En estas situaciones se necesita determinar un intervalo de confianza para esta diferencia entre medias de 2 grupos, esto se logra con la siguiente fórmula:

$$P\left(\overline{X}_1 - \overline{X}_2 - \left[Z\alpha/_2 \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right] \leq \boldsymbol{\mu_1 - \mu_2} \leq \overline{X}_1 - \overline{X}_2 + \left[Z\alpha/_2 \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]\right) = \gamma$$

Problema Ejemplo. Un grupo de ingenieros realiza pruebas con dos aleaciones de aluminio para determinar cuál es la más resistente a la tensión en largueros de aluminio usados en la fabricación del ala de un avión de transporte comercial. A su vez se desea utilizar la más económica, pero dando prioridad a la más resistente.

Se realizaron pruebas de resistencia a las dos aleaciones diferentes y se conocen las desviaciones estándar por experiencias pasadas con el proceso de fabricación de largueros y el procedimiento de prueba. Los datos con los que se cuenta son los siguientes:

Aleación de aluminio	Tamaño de muestra	Media muestral de la resistencia (lb/pulgadas^2)	Desviación estándar (lb/pulgadas^2)	Precio por libra en dlls.
Tipo 1	$n_1 = 10$	$\overline{X}_1 = 87.6$	σ_1=1.0	\$5.0
Tipo 2	$n_2 = 12$	$\overline{X}_2 = 74.5$	σ_2=1.5	\$6.5

Si $\mu_{1\,y}\,\mu_2$ denotan las verdaderas medias de la resistencia a la tensión para los dos tipos de aleaciones, encontrar un intervalo de confianza del 90% para la diferencia de la resistencia media o promedio $\boldsymbol{\mu_{1-}\ \mu_2}$. Es decir se necesita probar si la diferencia que existe entre la media de la aleación tipo 1 y la del tipo 2, es tan grande que se puedan considerar diferentes, y con ello determinar cual tiene la mayor resistencia.

Se sustituye en la formula

$$P\left(87.6 - 74.5 - 1.65\sqrt{\frac{1.0}{10}+\frac{2.25}{12}} \leq \boldsymbol{\mu_{1-}\ \mu_2} \leq 87.6 - 74.5 + 1.65\sqrt{\frac{1.0}{10}+\frac{2.25}{12}}\right) = 0.90$$

$$P([87.6 - 74.5 - 1.65\,(0.536) \leq \boldsymbol{\mu_{1-}\ \mu_2} \leq 87.6 - 74.5 + 1.65\,(0.536)]) = 0.90$$

$$P([12.215 \leq \boldsymbol{\mu_{1-}\ \mu_2} \leq 13.984]) = 0.90$$

Es decir, el intervalo que se espera de la diferencias entre la $\mu_{1\ y}\ \mu_2$ de ambas aleaciones, es de 12.215 hasta 13.984. dicho en otra forma, siendo la aleación 1 la que tiene una mayor resistencia, se espera que ésta tenga una diferencia mayor de 12.215 y hasta 13.984 lb/pulgadas^2 con respecto a la tipo 2, por lo tanto El grupo de ingenieros deberían optar por utilizar el larguero de aluminio con la aleación tipo 1, ya que se espera que sea más resistente que la tipo 2 y además es la más económica.

Nota: Observe que este intervalo no incluye al cero, por lo que se puede decir que las medias son diferentes. De hecho puede afirmarse que se tiene una confianza del 90% de que la resistencia a la tensión de la aleación tipo 1, excede a la de tipo 2 entre 12.22 y 13.98 lb/pulgadas^2

3.2 Intervalo de confianza para la diferencia de medias con varianzas desconocidas.

Siguiendo con la diferencia de medias, en ciertas ocasiones se requiere determinar si la media poblacional de un grupo de datos es estadísticamente no diferente a otro grupo, esto es que debido a la variación aleatoria, la $\boldsymbol{\mu_1}$ de un grupo 1, sea aproximadamente igual a la $\boldsymbol{\mu_2}$de un grupo 2, o en su caso determinar que estadísticamente no son iguales. Sin embargo en muchas ocasiones no se conoce la verdadera varianza y ésta se debe estimar. En estas situaciones se necesita determinar un intervalo de confianza para esta diferencia entre medias de 2 grupos utilizando la variable aleatoria t en lugar de z, esto se logra con la siguiente fórmula:

$$P\begin{pmatrix} \overline{X}_1 - \overline{X}_2 - \left[\left(t_{\alpha/2,n_1+n_2-2}\right)(Sp)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right] \leq \boldsymbol{\mu_{1-}\ \mu_2} \leq \\ \overline{X}_1 - \overline{X}_2 + \left[\left(t_{\alpha/2,n_1+n_2-2}\right)(Sp)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right] \end{pmatrix} = \gamma$$

Donde:

$$Sp = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$$

Ejemplo 11. En un artículo de la revista Hazardous Waste and Hazardous Material, se reportaron los resultados obtenidos de un análisis del peso del calcio en el cemento estándar y del cemento dopado con plomo. Los niveles reducidos de calcio indicarían que el mecanismo de hidratación del cemento se bloquea y permitiría que el agua atacara varios lugares en la estructura del cemento. 10 muestras de cemento estándar tuvieron un peso promedio porcentual de calcio de $\overline{X}_1 = 90.0$ con una desviación estándar de S_1=5.0 y 15 muestras del cemento dopado con plomo tuvieron un peso promedio porcentual de calcio de $\overline{X}_2$= 87.0, con una desviación estándar muestral de S_2=4.0. suponer que el peso porcentual del calcio sigue una distribución normal. Encontrar un intervalo de confianza del 95% para la diferencia de medias $\mu_{1-}\ \mu_2$ para los dos tipos de cemento.

$\mathrm{Sp} = \sqrt{\frac{(9)(25)+(14)(16)}{23}}$ Sp=4.42

$\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = \sqrt{\frac{1}{10}+\frac{1}{15}}$ =0.408

Buscamos el valor crítico de t cuando:

$t_{\alpha/2}, n_1 - n_2 - 2$ es decir $t_{0.05/2}, 10 - 15 - 2$ encontrando t=2.069

Sustituimos en la formula:

$$P\begin{pmatrix} 90 - 87 - [(2.069)(4.42)(0.408)] \quad \leq \boldsymbol{\mu_{1-}\ \mu_2} \leq \\ 90 - 87 + [(2.069)(4.42)(0.408)] \end{pmatrix} = 0.95$$

$$P(-0.73 \quad \leq \boldsymbol{\mu_{1-}\ \mu_2} \leq 6.73) = 0.95$$

Nota: Observe que este intervalo **SI** incluye al cero, por lo que se puede decir que estadísticamente las medias no son diferentes. De hecho puede afirmarse que se tiene una confianza del 95% de que la concentración de calcio en el cemento estándar comparándolo con el cemento dopado con plomo, no son diferentes.

Lo anterior se puede ver de la siguiente manera:

El cemento estándar tiene en promedio más concentración de calcio $\overline{X}_1 = 90.0$ y el que contiene plomo $\overline{X}_2$= 87.0, al parecer el estándar es la mejor opción, sin embargo, recuerde que estamos tratando con medias muestrales, y nos interesa conocer el intervalo esperado que tendrán las medias poblacionales verdaderas y como el intervalo encontrado incluye al cero, esto significa que va a ver ocasiones en que la diferencia en concentración de calcio entre los dos cementos va a ser cero (no son diferentes). Inclusive al haber números negativos, hay la posibilidad que en muestras futuras, el cemento con plomo tenga mayor concentración de calcio. Por lo tanto estadísticamente no son diferentes.

4.- DESCRIPCIÓN

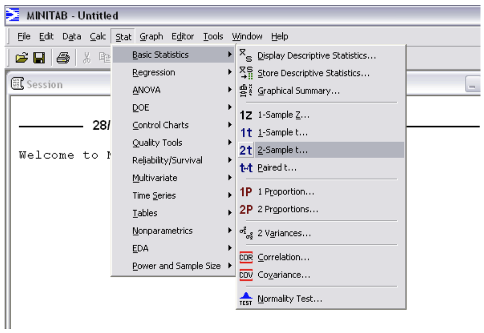
A) PROCEDIMIENTO Y DURACION DE LA PRÁCTICA:

1.- Encienda su computadora y abra el programa Minitab.

2.- Siga la secuencia siguiente.

2.1 Intervalo de confianza y prueba de hipótesis para la diferencia de dos medias.

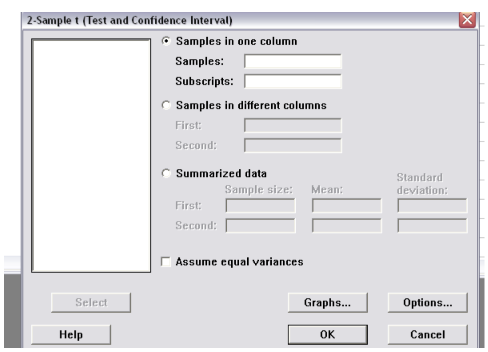
Seleccione la secuencia siguiente



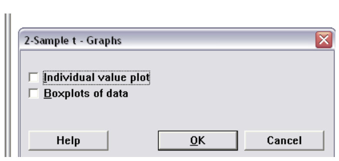
Si los datos están en la misma columna, seleccione la opción (**Samples in one column**).

Si los datos están en dos columnas, seleccione (**Samples in defferent columns**)

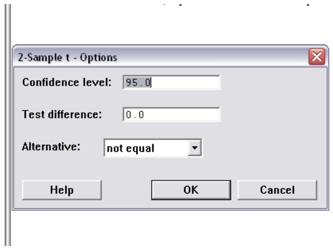
Si solo se tienen datos resumidos, es decir el tamaño de muestra, promedio y desviación estándar de cada muestra, seleccione la opción (**Summarized data**). Es recomendable seleccionar que se asume igualdad de varianzas y previamente haberlo verificado.



Al seleccionar **Graphs**, indique qué tipo de grafica desea obtener.



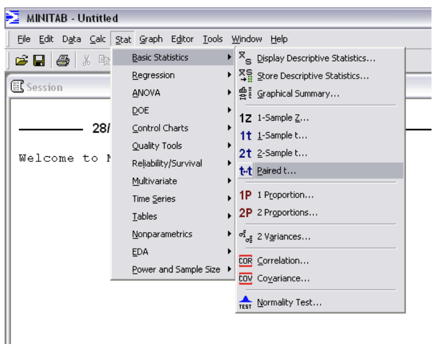
En **options**, especifique el nivel de confianza a utilizar, así como la diferencia que se desea probar y su respectiva hipótesis alternativa.



Finalmente de ok.

2.2 Intervalo de confianza y prueba de hipótesis para la diferencia de medias de conjuntos pareados.

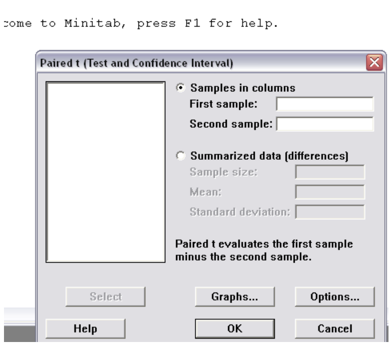
Ingrese con la siguiente secuencia en Minitab.



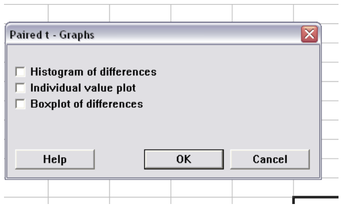
Si los datos están en la misma columna, seleccione la opción (**Samples in one column**).

Si los datos están en dos columnas, seleccione (**Samples in defferent columns**)

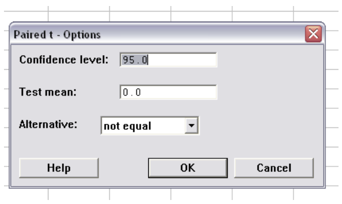
Nota: en esta prueba se deben tener los datos alineados en pares (apareados) de lo contrario Minitab no procederá con los cálculos.



Al seleccionar **Graphs**, indique qué tipo de grafica desea obtener.



En **options**, especifique el nivel de confianza a utilizar, así como la diferencia que se desea probar y su respectiva hipótesis alternativa.



Finalmente de ok para proceder con el cálculo.

B) DESARROLLO

practique analizando los siguientes problemas

1. Se usan dos máquinas para llenar botellas de plástico con un volumen neto de llenado de 16.0 onzas. Puede suponerse que el volumen neto de llenado es normal, con una desviación estándar de σ_2=0.020 y σ_2=0.025 onzas. Uno de los miembros del personal de ingeniería sospecha que ambas máquinas hacen el llenado con el mismo volumen neto medio, sea este volumen 16 onzas o no. Se toma una muestra aleatoria de 10 botellas de la producción de cada máquina, cuyos resultados se muestran enseguida:

   Máquina 1: 16.03, 16.04, 16.05, 16.05, 16.02, 16.01, 15.96, 15.98, 16.02 y 15.99

   Máquina 2: 16.02, 15.97, 15.96, 16.01, 15.99, 16.03, 16.04, 16.02, 16.01 y 16.00

   Encuentre un intervalo de confianza del 95% para la diferencia de las medias. Proponga una interpretación práctica para este intervalo.

2. Un investigador está determinando la utilidad de dos diferentes lenguajes de programación en cuanto a su facilidad de uso y la reducción del tiempo de programación. 12 expertos programadores, familiarizados con ambos lenguajes, trabajaron en realizar un determinado código en ambos lenguajes, tardando los siguientes minutos en hacerlo

Programador	1	2	3	4	5	6	7	8	9	10	11	12
Lenguaje 1	17	16	21	14	18	24	16	14	21	23	13	18
Lenguaje 2	18	14	19	11	23	21	10	13	19	24	15	20

   a) ¿Son iguales ambos lenguajes?
   b) ¿Cuál es el p value y qué significa?

c) ¿Son independientes los datos?, ¿por qué?

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

- Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
- Douglas C. Montgomery (2009). Design and Analysis of Experiments
- Douglas C. Montgomery y George C. Runger, (2005). Probabilidad y Estadística aplicadas a la Ingeniería. Editorial Limusa Wiley, segunda edición.
- Seymour Lipschutz y John Schiller, (2000). Introducción a la Probabilidad y Estadística. Editorial Mc Graw Hill.
- Humberto Gutiérrez y Román De La Vara, (2009). Control Estadístico de Calidad y Seis Sigma. Editorial Mc Graw Hill, segunda edición.

6.- ANEXOS:

REQUERIMIENTOS PARA REALIZACION DE PRÁCTICAS EDUCATIVAS EN LABORATORIOS DE LA FIE

NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**DISEÑO COMPLETAMENTE ALEATORIO**	**PRÁCTICA NÚMERO**	**4**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18
SOFTWARE REQUERIDO	
MINITAB	
OBSERVACIONES-COMENTARIOS	

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

La presente práctica aborda diseño completamente al azar, el cual es el más sencillo de los experimento al solo considerar un solo factor.

En esta práctica se analizará y comparará a dos o más poblaciones tomando en cuenta un sólo factor a dos o más niveles o tratamientos y la variación que estos producen. Lo anterior utilizando el software Minitab.

2.- OBJETIVO (COMPETENCIA):

Comparar la variación entre opciones de un experimento que contempla un solo factor a dos o más niveles, para la correcta toma de decisiones, desarrollando un Análisis de varianza y utilizando las herramientas estadísticas del software Minitab.

3.- TEORIA:

En muchas aplicaciones interesa comparar dos condiciones de un fenómeno, por ejemplo la resistencia de un alambre de cobre antes y después de un recubrimiento. En este caso pudiéramos describir este experimento como un experimento de un solo factor (resistencia del alambre de cobre) probado a dos niveles o tratamientos (antes y después del recubrimiento). Este tipo de experimento puede analizarse por medio de intervalos de confianza, prueba de hipótesis entre dos poblaciones sobre su igualdad de media o de varianza. A su vez estas pruebas estadísticas tienen una característica similar, todos consideran sólo dos poblaciones y un factor que se analiza. Sin embargo, al realizar investigación, por ejemplo la experimentación para la mejora de un proceso o producto, es común que se requiera comparar entre dos o más opciones de un mismo factor por ejemplo comparar entre: tres tipos de máquina a usar, cuatro tipos de material a emplear, dos proveedores a seleccionar, etc., en estos casos se tienen más de dos poblaciones y se trabaja de manera diferente.

El Análisis De Varianza

El nombre de análisis de varianza se deriva de la partición de la variabilidad total de un fenómeno en sus partes componentes. Suponga que se tienen ***a*** **tratamientos o niveles** diferentes **de un solo factor** que quieren compararse. La respuesta observada de cada uno de los ***a*** tratamientos es una variable aleatoria y se puede representar como en la siguiente tabla.

Tratamientos o Niveles	Observaciones				Total	Promedio
1	y_{11}	y_{12}	⋯	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	⋯	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
⋮	⋮	⋮	⋯	⋯	⋮	⋮

a	y_{a1}	y_{a2}	…	y_{an}	$y_{a.}$	$\bar{y}_{a.}$
					$y_{..}$	$\bar{y}_{..}$

Nota: algunos autores realizan el acomodo en vertical y denominan *k* tratamientos en lugar de *a* tratamientos.

	Tratamientos o Niveles			
	T1	T2	…	T_k
	Y_{11}	Y_{21}	…	Y_{k1}
	Y_{12}	Y_{22}	…	Y_{k2}
	⋮	⋮	…	…
	Y_{1n1}	Y_{2n2}	…	Y_{knk}
Sumatoria	$Y_{1.}$	$Y_{2.}$	…	$Y_{k.}$
promedio	$\bar{y}_{1.}$	$\bar{y}_{2.}$	…	$\bar{y}_{k.}$

Donde y_{ij} representa la *j*-ésima observación del factor en su nivel o tratamiento *i*. Así pues habrá *n* observaciones del tratamiento *i.*

Cuando se realiza un Anova de un solo factor, también se conoce comúnmente en ingles como ***one-way Anova***, ya que solo se investiga un solo factor. En estos casos se requiere que el experimento sea desarrollado en completo orden aleatorio, de tal manera que el ambiente en el que los tratamientos son aplicados (también llamados unidad experimental) sea lo más uniforme posible. Para hacer aleatorio el experimento, se puede apoyar en números aleatorios, por ejemplo de las calculadoras convencionales.

Supuestos de Anova.

Para las pruebas de hipótesis, se asume que los errores del modelo parten de una distribución normal y que son independientes, con media cero y varianza σ^2. Para ver la utilidad del Anova realizaremos un experimento con un solo factor

Diseño Completamente Al Azar (Dca)

Cuando se requiere analizar **un factor** a más de dos niveles o tratamientos, el método de Diseños Completamente al Azar, es uno de los más sencillos que se pueden realizar, ya que sólo se consideran dos fuentes de variabilidad: *los tratamientos y el error aleatorio*. Se llama así porque todas las corridas experimentales se realizan en orden aleatorio, de manera que los posibles efectos ambientales y temporales del experimento, se vayan repartiendo equitativamente entre los tratamientos.

Ejemplo

Un equipo de mejora investiga el efecto de cuatro métodos de ensamble A, B, C y D sobre el tiempo de ensamble en minutos. La estrategia de experimentación es aplicar cuatro veces los cuatro métodos de ensamble en orden completamente aleatorio obteniendo los siguientes resultados en minutos:

Método de ensamble

A	**B**	**C**	**D**
6	7	11	10
8	9	16	12
7	10	11	11
8	8	13	9

A partir de la tabla anterior, se realizan los cálculos siguientes

$$Y_{i.} = \text{Suma de las observaciones del tratamiento i, es decir} \qquad Y_{i.} = \sum_{j=1}^{n_i} \mathrm{Y}_{\mathrm{ij}}$$

$$\bar{Y}_{i.} = \text{Media de las observaciones del i-ésimo tratamiento, es decir} \qquad \bar{Y}_{i.} = \frac{\sum_{j=1}^{n_i} \mathrm{Y}_{\mathrm{ij}}}{n_i}$$

$$Y_{..} = \text{Suma total de las } \mathrm{N} = \mathrm{n}_1 + \mathrm{n}_2 + \ldots + \mathrm{n}_k \text{ mediciones, es decir} \qquad Y_{..} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} Y_{ij}$$

$\bar{Y}_{..}$ = Promedio de las $N = n_1 + n_2 + \ldots + n_k$ mediciones, es decir $\qquad \bar{Y}_{..} = \frac{Y_{..}}{N}$

$\hat{\tau}_i$ = Desviaciones respecto a la media global o efecto estimado delo tratamiento i, es decir

$$\hat{\tau}_i = \bar{Y}_{i.} - \bar{Y}_{..}$$

Para el ejercicio de los 4 métodos de ensamble, se probará la hipótesis de que todos los métodos son iguales contra de que son diferentes (al menos dos serían diferentes), es decir:

$$\boldsymbol{H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu}$$

$$\boldsymbol{H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq \mu}$$

Para esta prueba se debe desarrollar un análisis de varianza en la respectiva tabla Anova.

Tabla de ANOVA.

Fuente de variación	Suma de cuadrados (SC)	grados de libertad (gl)	Cuadrados medios (CM)	Estadístico F_0	Valor P
Tratamiento	$SC_{Trat} = \sum_{i=1}^{k} \frac{Y_{i.}^2}{n_1} - \frac{Y_{..}^2}{N}$	$k-1$	$CM_{Trat} = \frac{SC_{Trat}}{k-1}$	$\frac{CM_{Trat}}{CM_{Error}}$	$P(F > F_0)$
Error	$SC_{Error} = SC_T - SC_{Trat}$	$N-k$	$CM_{Error} = \frac{SC_{Error}}{N-k}$		
Total	$SC_T = \sum_{i=1}^{k} \sum_{j=1}^{n_i} Y_{ij}^2 - \frac{Y_{..}^2}{N}$	$N-1$			

En este caso se rechaza H_0 si $F_0 > F_{\alpha,\ k-1,\ N-k}$ o también se rechaza si el $valor\ p < \alpha$ donde el valor p es el área de la distribución $F_{\ k-1,\ N-k}$ a la derecha del estadístico F_0

4.- DESCRIPCIÓN

A) PROCEDIMIENTO Y DURACION DE LA PRÁCTICA:

1.- Encienda su computadora y abra el programa Excel y Minitab

2.- Desarrolle la tabla Anova anterior en Excel de los siguientes problemas.

3. Desarrolle el análisis de varianza utilizando Minitab de los siguientes problemas

B) DESARROLLO

practique analizando los siguientes problemas

1.- Se hace un estudio sobre la efectividad de 3 marcas de spray para matar moscos. Para ello cada producto se aplica sobre un grupo de 100 moscos, y se cuenta el número de moscos muertos expresado en porcentajes. Se hacen seis replicas y los resultados obtenidos se muestran a continuación.

	Numero de replica					
Marca de spray	1	2	3	4	5	6
1	72	65	67	75	62	73
2	55	59	68	70	53	50
3	64	74	61	58	51	69

A)Formule la hipótesis adecuada.

b) ¿Existe diferencia entre la efectividad promedio de los productos en spay?

c) ¿Hay algún spray mejor? Argumenta su respuesta

d) Da un intervalo al 95% de confianza para la efectividad promedio (porcentaje) de cada una de las marcas

e)Dibuje las graficas de medias y los diagramas de caja simultáneos, después interprételos

f)Verifique los supuestos de normalidad y de igual varianza entre las marcas

2.- En un centro de investigación se realiza un estudio para comparar varios tratamientos que al aplicarse previamente a los frijoles crudos, reducen su tiempo de cocción. Estos tratamientos son a base de bicarbonato de sodio (NaHCO3) y Cloruro de sodio o sal común (NaCl). El primer tratamiento es el llamado***control***, que consiste en no aplicar ningún tratamiento. El tratamiento T2 es el remojo de agua con bicarbonato de sodio, el T3 es remojar en agua con sal común y el T4 es remojar agua con una combinación de ambos ingredientes en proporciones iguales. La variable de respuesta es el tiempo de cocción en minutos. Los datos se muestran en la siguiente tabla.

Control	T2	T3	T4
213	76	57	84
214	85	67	82
204	74	55	85
208	78	64	92
212	82	61	87
200	75	63	79
207	85	63	90

a) ¿De que manera el experimentador debe aleatorizar los experimentos y el material experimental?
b) De ejemplos de qué otros factores que deben estar fijos durante las pruebas experimentales, para que no afecten los resultados y las conclusiones?
c) Formule y pruebe la hipótesis de que las medias de los tratamientos son iguales
d) Obtenga el diagrama de caja, después interprételo
e) ¿Hay algún tratamiento mejor? ¿Cuál es el tiempo de cocción esperado para el mejor tratamiento?
f) Algo importante a cuidar en un experimento es que no haya efectos colaterales no deseados, causados por el tratamiento ganador; en este caso, piense en los posibles efectos colaterales que podría causar el mejor tratamiento
g) ¿Se cumplen los supuestos del modelo? Verifique gráficamente
h) Pruebe la hipótesis de igualdad de varianzas entre tratamientos (que corresponde a un supuesto)

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
Douglas C. Montgomery (2009). Design and Analysis of Experiments

6.- ANEXOS:

NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**INTRODUCCION AL DISEÑO DE EXPERIMENTOS (DDE)**	**PRÁCTICA NÚMERO**	**5**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	CANTIDAD
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

REQUERIMIENTOS PARA REALIZACION DE PRÁCTICAS EDUCATIVAS EN LABORATORIOS DE LA FIE

SOFTWARE REQUERIDO	
MINITAB	
OBSERVACIONES-COMENTARIOS	
NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

El Diseño de Experimentos en La Industria

En la industria, es una práctica común hacer experimentos o pruebas con la intención de que al mover o hacer algunos cambios en los materiales, métodos o condiciones de operación de un proceso, se puedan detectar, resolver o minimizar los problemas de calidad. Por ejemplo, se prueban varias temperaturas en una máquina hasta encontrar la que da el mejor resultado, o se intenta un nuevo material con la intención de eliminar los problemas que tiene el material actual, o bien, se prueban diferentes velocidades para determinar la que minimiza la vibración excesiva del equipo.

Sin embargo, es común que estas pruebas o experimentos se hagan sobre la marcha, a prueba y error, apelando a la experiencia y a la intuición; en lugar de seguir un plan experimental adecuado que garantice una buena respuesta a las interrogantes planteadas. Algo similar puede decirse respecto al análisis de los datos experimentales, donde más que hacer un análisis riguroso de toda la información obtenida y que tome en cuenta la variación, se hace análisis informal “intuitivo”. Es tal el poder de la experimentación, que en ocasiones a pesar de que el experimento se hizo a prueba y error se logran mejoras. Sin embargo, en muchas situaciones no es suficiente aplicar experimentación a prueba y error, por lo que es mejor proceder siempre en una forma eficaz que garantice la obtención de las respuestas a las interrogantes planteadas, en un lapso corto de tiempo y utilizando pocos recursos.

El diseño estadístico de experimentos es precisamente la forma eficaz de hacer prueba en los procesos, ya que proporciona la técnica y la estrategia necesarias para llevar de manera eficaz los procesos a mejores condiciones de operación. En su parte medular, el *diseño de experimentos* consiste en determinar cuáles pruebas y cómo es que se deben realizar, para obtener datos que al analizarlos estadísticamente se obtengan conclusiones y decisiones que deriven en mejoras del desempeño del proceso.

2.- OBJETIVO (COMPETENCIA):

Investigar y comparar dos opciones mediante estadística inferencial, utilizando las herramientas estadísticas del software Minitab.

3.- TEORIA:

Experimento: Es un cambio en las condiciones de operación de un sistema o proceso, que se hace con el objetivo de medir el efecto del cambio sobre una o varias propiedades del producto. Dicho experimento permite aumentar el conocimiento acerca del sistema.

Diseño de experimentos: Consiste en planear un conjunto de pruebas experimentales, de tal manera que los datos generados puedan analizarse estadísticamente para obtener conclusiones válidas y objetivas acerca de un sistema o proceso.

Unidad experimental: Es la muestra de artículos que es necesario producir en una condición de operación del proceso para obtener, a partir de ellos, una medición o dato representativo de lo que allí ocurre. En cada diseño de experimentos es importante definir cuidadosamente la unidad experimental, ya que esta puede ser una pieza o un conjunto de piezas producidas, dependiendo del proceso que se estudia.

Variable de respuesta: Es la característica, variable de salida o propiedad del producto, cuyo valor interesa mejorar. Por lo general el valor de dicha característica determina algún aspecto de la calidad. La conjetura típica para utilizar el diseño de experimentos es que existe otra manera de operar el proceso en la cual el comportamiento de una o varias variables de respuesta sería mejor que el actual.

Factores controlables: Son variables de proceso o variables de entrada que se pueden fijar en un punto o en un nivel de operación. Algunos de ellos son los que usualmente se controlan durante la operación normal del proceso y se distinguen porque para cada uno de ellos existe la manera o el mecanismo para cambiar o manipular su nivel de operación. Esto último es lo que hace posible que se pueda experimentar con ellos.

Factores no controlables o de ruido: Son variables que no se pueden controlar durante la operación normal del proceso o que resulta impráctico o caro controlarlos. Por ejemplo, algunos factores que suelen ser no controlables son las variables ambientales, el ánimo de los operadores, la calidad del material que se recibe del proveedor y diversos usos que el cliente puede dar al producto. Un factor que ahora es no controlable puede convertirse en controlable, cuando se tenga el mecanismo o tecnología para ello.

Error aleatorio y error experimental: Siempre que se realiza un estudio experimental, parte de la variabilidad observada no se podrá explicar por los factores estudiados. Esto es, siempre habrá un remanente de variabilidad natural del proceso. Esta variabilidad constituye el llamado **error aleatorio**, que no es error en el sentido de equivocación, sino variabilidad no explicada. Por ejemplo será parte del error aleatorio, el pequeño efecto que tienen los factores que no se estudiaron, siempre y cuando se mantenga pequeño o despreciable, así como la variabilidad de las mediciones hechas bajo las mismas condiciones. Sin embargo el error aleatorio, también absorberá todos los errores (ahora sí en el sentido de equivocación) que el experimentador cometa durante los experimentos y si estos son graves, más que error aleatorio hablaremos de **error experimental**, y de predominar éste, la detección de cuáles de los factores estudiados tienen un efecto real sobre la respuesta será difícil, sino es que imposible.

Cuando se corre un diseño experimental es importante que la variabilidad de la respuesta observada se deba principalmente a los factores estudiados y en menor medida al error aleatorio, y además que este error sea efectivamente aleatorio. Cuando la variabilidad observada se debe a factores no estudiados o a error no aleatorio, no se podrá distinguir cual es el verdadero efecto que tienen los factores estudiados, con lo que el experimento resultaría inútil. De aquí la importancia de no dejar variar libremente ningún factor que pueda influir de manera significativa sobre el comportamiento de la respuesta.

Se debe ser muy cuidadoso en la planeación y análisis de un experimento. El punto de partida para una correcta planeación es aplicar los principios básicos del DDE: **aleatorización, repetición y bloqueo,** los cuales tienen que ver directamente con que los datos obtenidos sean útiles para responder las preguntas planteadas, es decir, la validez del análisis de los datos se apoya en estos principios.

Aleatorización: Consiste en hacer las corridas experimentales en orden aleatorio y con material seleccionado también aleatoriamente. Este principio aumenta la probabilidad de que el supuesto de independencia de los errores se cumpla, que es un requisito para la validez de las pruebas estadísticas que se realizan. También es una manera de asegurar que las pequeñas diferencias provocadas por materiales, equipo y todos los factores no controlados, se repartan de manera homogénea en todos los tratamientos. Por ejemplo, una evidencia de incumplimiento o violación de este principio se manifiesta cuando el resultado obtenido en una corrida del proceso está muy influenciado por la corrida inmediata anterior.

Repetición: Es correr más de una vez un tratamiento o combinación de factores dada. No confundir este principio con medir varias veces el mismo producto o artículos fabricados de una solo vez en cierta combinación de factores. Repetir es volver a correr el proceso, partir desde volver a fijar las condiciones de operación, para obtener un nuevo producto, hasta medir el resultado de esta otra corrida del proceso. **Las repeticiones permiten distinguir mejor que parte de la variabilidad total de los datos se debe al error aleatorio y cual a los factores.** Cuando no se hacen repeticiones no hay manera de estimar la variabilidad natural o error aleatorio y esto dificulta la construcción de estadísticos realistas en el análisis de los datos. El repetir aumenta la confiabilidad de las mediciones, ya que las repeticiones en el mismo tratamiento se esperan razonablemente parecidas, en particular cuando el proceso está en control estadístico.

Bloqueo: Es nulificar o tomar en cuenta en forma adecuada todos los factores que pueden afectar la respuesta observada. El nombre de este principio se deriva de experimentos agrícolas en los cuales se controla el efecto de parcela (bloque) al comparar varios tratamientos. Al bloquear se supone que el subconjunto de datos que se obtengan dentro de cada bloque, deben resultar más homogéneos que el conjunto total de los datos. Con el principio de bloqueo se logra mayor precisión en el experimento al eliminar la variabilidad no explicada que provocaría el error aleatorio. Por ejemplo si se quieren comparar cuatro máquinas del mismo tipo usando un operador, la comparación puede resultar injusta por favorecer a una de las máquinas (a la preferida por el operador), y entonces no hay garantía de que el resultado obtenido se mantenga con los otros operadores; en este caso se está utilizando un solo bloque (operador). Una mejor estrategia es experimentar con los cuatro operadores (cuatro bloques), donde cada uno de ellos prueba en orden aleatorio las cuatro máquinas; en este segundo caso la comparación de las máquinas resulta más justa y de mayor validez. Cada operador es un bloque porque se espera que las mediciones del mismo operador sean más parecidas entre sí que las mediciones de varios operadores.

4.- DESCRIPCIÓN

A) PROCEDIMIENTO Y DURACION DE LA PRÁCTICA:

1.- Encienda su computadora y abra el programa Minitab, espere instrucciones del profesor

2.-Escuche atentamente las explicaciones del maestro y desarrolle los ejercicios que le ida el profesor.

B) DESARROLLO

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en realizar una investigación de los elementos que comprenden la elaboración de un modelo básico de simulación, específicamente las locaciones, entidades, arribos y procesos

D) RESULTADOS:

Para esta práctica los resultados serán el conocimiento e identificación de los elementos básicos así como su creación dentro del ambiente Promodel

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

Charle R. Harrel, Biman K. Ghosh (2003,de julio 25) Simulation using PROMODEL.. Mc Graw Hill, E. U.

Eduardo Garcia Dunna, Heriberto García Reyes Heriberto, Simulación y Análisis con PROMODEL, España , Pearson Prentice Hall.

Paginas internet

Simulación promodel (manual español) (2006, 6 de diciembre) http://aeisc.wordpress.com

6.- ANEXOS:

NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**COMPARACIÓN DE PAREJAS DE MEDIAS DE TRATAMIENTOS**	**PRÁCTICA NÚMERO**	**6**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

REQUERIMIENTOS PARA REALIZACION DE PRÁCTICAS EDUCATIVAS EN LABORATORIOS DE LA FIE

SOFTWARE REQUERIDO
MINITAB y EXCEL
OBSERVACIONES-COMENTARIOS

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

En muchas soluciones prácticas, se desea comparar solo parejas de medias, para esto se pudiera determinar qué media difiere de las demás a través de pruebas de diferencias de medias de todos los pares de tratamientos.

En esta práctica se determinará cual tratamiento o nivel difiere de los demás por medio de comparaciones entre todas las medias utilizando el método de Fisher y Duncan. Lo anterior utilizando el software Minitab y Excel.

2.- OBJETIVO (COMPETENCIA):

Decidir entre tratamientos que sean diferentes por medio del análisis de sus medias y que cumplan con los requerimientos, mediante el uso del software Minitab y Excel.

3.- TEORIA:

Así estamos interesados en contrastar todos los pares de promedios de los ***a*** tratamientos con la consiguiente hipótesis nula

$$H_0: \mu_i = \mu_j \ para \ todas \ las \ i \neq j$$

Para esta prueba existen diferentes métodos, los cuales se presentan a continuación

3.1 MÉTODO LSD (Diferencia Mínima Significativa-Least Significant Difference)

Cuando se rechaza H_0 mediante el ANOVA, y se concluye que no hay igualdad entre las medias poblacionales de los tratamientos, pero no se tiene información específica sobre cuáles tratamientos son diferentes entre si, entonces es necesario probar la igualdad de todos los posibles pares de medias utilizando métodos de comparaciones múltiples o pruebas de rango múltiple. Una de esas pruebas es el método LSD y se utiliza para probar igualdad de todos los posibles pares de medias con la hipótesis:

$$H_0 : \mu_i = \mu_j$$

$$H_1 : \mu_i \neq \mu_j$$

Para *i ≠ j*. A su vez, para cada ***k*** tratamientos se tiene en total k(k-1)/2 pares de medias, por ejemplo si hay 4 tratamientos, es decir k=4, entonces existen 4x3/2= 6 posibles pares de medias.

El estadístico de prueba para cada una de las hipótesis, es la diferencia en valor absoluto entre sus medias muestrales

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

Se rechaza la hipótesis $H_0 = \mu_i - \mu_j$ si ocurre lo siguiente

$$|\bar{Y}_{i.} - \bar{Y}_{j.}| > t\alpha_{/2,N-k}\sqrt{CM_E\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} = LSD$$

Donde:

El valor $t\alpha_{/2,N-k}$ se lee en las tablas de la distribución T de student con N-k grados de libertad que corresponden al error.

El CM_E es el cuadrado medio del error y se obtiene de la tabla de Anova.

$n_{i\ y}\ n_j$ son el número de observaciones para los tratamientos *i y j.*

La cantidad LSD se llama Diferencia Mínima Significativa, ya que es la diferencia mínima que debe existir entre dos medias muestrales para considerar que los tratamientos correspondientes son significativamente diferentes. De esta manera **cada diferencia de medias muestrales en valor absoluto que sea mayor que el número LSD se declara significativo**.

Nota: si el diseño es balanceado, es decir, si $n_1 = n_2 = \cdots = n_k = n$ el LSD se reduce a

$$LSD = t\alpha_{/2,N-k}\sqrt{2CM_E/n}$$

En caso de rechazar H_0, se acepta la hipótesis alternativa H_1: $\mu_i \neq \mu_j$, es decir, que los tratamientos i y j son diferentes.

3.2 Método Duncan

En este método para comparación de medias, si las ▯ muestras son de igual tamaño, los ▯ promedios se acomodan en orden ascendente y el error estándar de los promedios se estima: $S_{\bar{Y}_i} = \sqrt{CM_{Error}/n}$

Si alguno de los tratamientos o todos los tratamientos tienen tamaños diferentes, se reemplaza ▯ por la media armónica que está dada por:

$$n_{AR} = \frac{k}{\sum_{i=1}^{k}\frac{1}{n_i}}$$

A su vez, los valores críticos se obtienen de la tabla de rangos significantes de Duncan así:

$r_{\alpha}(p,l),\ p = 2,3,\ldots,k$, donde $\boldsymbol{\alpha}$ es el nivel de significancia y ▯ son los grados de libertad para el error. Con estos ▯-▯▯valores se obtienen los rangos de significancia mínima dados por

$$R_p = r_\alpha(p,l)S_{\bar{Y}_{i.}};\ p = 2,3,\ldots,k,$$

Las diferencias observadas entre las medias muestrales se comparan con los rangos R_p así:

1. Se compara la diferencia entre la media más grande y la más pequeña con el rango R_k.
2. La diferencia entre la media más grande y la segunda más pequeña se compara con el rango R_{k-1}.
3. Estas comparaciones continúan hasta que la media mayor se haya comparado con todas las demás.
4. Enseguida se compara la diferencia entre la segunda media más grande y la media menor con el rango $R_{k-1.}$
5. Posteriormente la diferencia entre la segunda media más grande y la segunda más pequeña se compara con el valor de R_{k-2} y así sucesivamente hasta que se comparan los *k(k-1)/2* pares de medias posibles con el rango que les corresponda.

En las comparaciones donde la diferencia observada es mayor que el rango respectivo, se concluye que esas medias son significativamente diferentes. Si dos medias caen entre otras dos que no son muy diferentes, entonces esas dos medias poblacionales también se consideran estadísticamente iguales.

4.- DESCRIPCIÓN

A) PROCEDIMIENTO Y DURACION DE LA PRÁCTICA:

1.- Encienda su computadora y abra el programa Minitab.

2.- Resuelva los siguientes problemas utilizando Excel y Minitab

B) DESARROLLO

Practique analizando los siguientes problemas

4.1 Un equipo de mejora investiga el efecto de cuatro métodos de ensamble A, B, C y D sobre el tiempo de ensamble en minutos. La estrategia de experimentación es aplicar cuatro veces los cuatro métodos de ensamble en orden completamente aleatorio obteniendo los siguientes resultados en minutos:

Método de ensamble

A	**B**	**C**	**D**
6	7	11	10
8	9	16	12
7	10	11	11
8	8	13	9

4.2 Un fabricante de calzado desea mejorar la calidad de las suelas, las cuales se pueden hacer con uno de los cuatro tipos de cuero (A, B, C y D) disponibles en el mercado. Para ello prueba los cueros con una máquina que hace pasar los zapatos por una superficie. Como criterio de desgaste se usa la pérdida de peso después de un número fijo de ciclos, es decir se esperaría que las suelas de menor calidad se desgastaron más y por lo tanto pesan menos. Se prueban en orden aleatorio 24 zapatos, seis de cada tipo de cuero, cuyos datos se muestran en la siguiente tabla.

Tipo de cuero	**Observaciones**					
|A	264	260	258	241	262	255
B	208	220	216	200	213	206
C	220	263	219	225	230	228
D	217	226	215	227	220	222

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

- Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
- Douglas C. Montgomery (2009). Design and Analysis of Experiments
- Douglas C. Montgomery y George C. Runger, (2005). Probabilidad y Estadística aplicadas a la Ingeniería. Editorial Limusa Wiley, segunda edición.
- Seymour Lipschutz y John Schiller, (2000). Introducción a la Probabilidad y Estadística. Editorial Mc Graw Hill.
- Humberto Gutiérrez y Román De La Vara, (2009). Control Estadístico de Calidad y Seis Sigma. Editorial Mc Graw Hill, segunda edición.

6.- ANEXOS:

NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**DISEÑO DE CUADRO LATINO**	**PRÁCTICA NÚMERO**	**7**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

SOFTWARE REQUERIDO
MINITAB y EXCEL
OBSERVACIONES-COMENTARIOS

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

A continuación se analizan experimentos con tres factores simultáneos por medio del diseño de cuadro latino y el empleo del software Minitab y Excel.

2.- OBJETIVO (COMPETENCIA):

Comparar entre 3 tipos de factores y decidir cual es la mejor opción por medio del diseño latino y el empleo del software Minitab.

3.- TEORIA:

La siguiente es la tabla Anova que se realiza para obtener los componentes de variación.

Fuente de variación	**Suma de cuadrados (SC)**	**grados de libertad (gl)**	**Cuadrados medios (CM)**	**Estadístico F_0**	**Valor P**
Tratamiento	$SC_{Trat} = \sum_{j=1}^{k} \frac{Y_{.j.}^2}{k} - \frac{Y_{...}^2}{N}$	$k-1$	$CM_{Trat} = \frac{SC_{Trat}}{k-1}$	$\frac{CM_{Trat}}{CM_{Error}}$	$P(F > F_0)$
Renglones (B1)	$SC_{B1} = \sum_{i=1}^{k} \frac{Y_{i..}^2}{k} - \frac{Y_{...}^2}{N}$	$k-1$	$CM_{B1} = \frac{SC_{B1}}{k-1}$	$\frac{CM_{B1}}{CM_{Error}}$	$P(F > F_0)$
Columnas (B2)	$SC_{B2} = \sum_{l=1}^{k} \frac{Y_{..l}^2}{k} - \frac{Y_{...}^2}{N}$	$k-1$	$CM_{B2} = \frac{SC_{B2}}{k-1}$	$\frac{CM_{B2}}{CM_{Error}}$	$P(F > F_0)$
Error	$SC_{Error} = SC_T - SC_{Trat} - SC_{B1} - SC_{B2}$	$(k-2)(k-1)$	$CM_{Error} = \frac{SC_{Error}}{(k-2)(k-1)}$		
Total	$SC_T = \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{l=1}^{k} Y_{ijl}^2 - \frac{Y_{...}^2}{N}$	k^2-1			

4.- DESCRIPCIÓN

A) PROCEDIMIENTO Y DURACION DE LA PRÁCTICA:

1.- Encienda su computadora y abra el programa Minitab.

2.- Resuelva los siguientes problemas utilizando Excel y Minitab

Diseño de cuadro latino

Un experimentador estudia los efectos que tienen 5 formulaciones diferentes de la carga propulsora utilizada en los sistemas de expulsión de la tripulación de un avión basado en la rapidez de combustión. Cada formulación se hace con un lote de materia prima que sólo alcanza para probar cinco formulaciones. Además las formulaciones son preparadas por cinco operadores y puede haber diferencias sustanciales en las habilidades y experiencia de ellos. Por lo tanto, al parecer hay dos factores perturbadores que serán los lotes de materia prima y operadores.

El diseño apropiado para este problema consiste en probar cada formulación exactamente una vez con cada uno de los cinco operadores, con los cinco lotes, de tal manera que resultará en un cuadro de 5x5 llamado Cuadro latino. Algunos ejemplos de cuadro latino de 5x5 son los siguientes

A	B	C	D	E
B	A	E	C	D
C	D	A	E	B
D	E	B	A	C
E	C	D	B	A

B	C	A	D	E
C	A	D	E	B
A	D	E	B	C
D	E	B	C	A
E	B	C	A	D

Observe como en cualquier renglón o columna no se repite ninguna letra. Al primer cuadro que empieza con renglón y columna en el orden de abecedario se le llama cuadro latino estándar, sin embargo al utilizar cuadros de 5x5, se puede llegar a encontrar hasta 161280 formas, estos son sólo 2 de ellos.

Al realizar la experimentación, se lograron los siguientes resultados con el siguiente orden.

Letras latinas= Formulas (Factores de interés o tratamiento)

Renglones = Lote de materia prima (Factor de bloque 1)

Columnas= Operadores (Factor de bloqueo 2)

Operadores

Lote de materia prima	1		2		3		4		5	
1	A	24	B	20	C	19	D	24	E	24
2	B	17	C	24	D	30	E	27	A	36
3	C	18	D	38	E	26	A	27	B	21
4	D	26	E	31	A	26	B	23	C	22
5	E	22	A	30	B	20	C	29	D	31

Con esta información genere la tabla de Anova correspondiente y pruebe la hipótesis de que las formulas producen la misma rapidez de combustión en milisegundos.

Posteriormente determine si aplica, ¿cuál formula es la más rápida?...utilice el criterio del método LSD que para un cuadro latino es el siguiente:

$$LSD = t_{\frac{\alpha}{2},(k-2)(k-1)}\sqrt{\frac{2CM_{Error}}{k}}$$

Diseño de cuadro latino en minitab

Considere el ejemplo de la carga propulsora, donde interactúan 3 factores (Formula A-B-C-D-E, Lote 1-2-3-4-5 y Operador1-2-3-4-5)

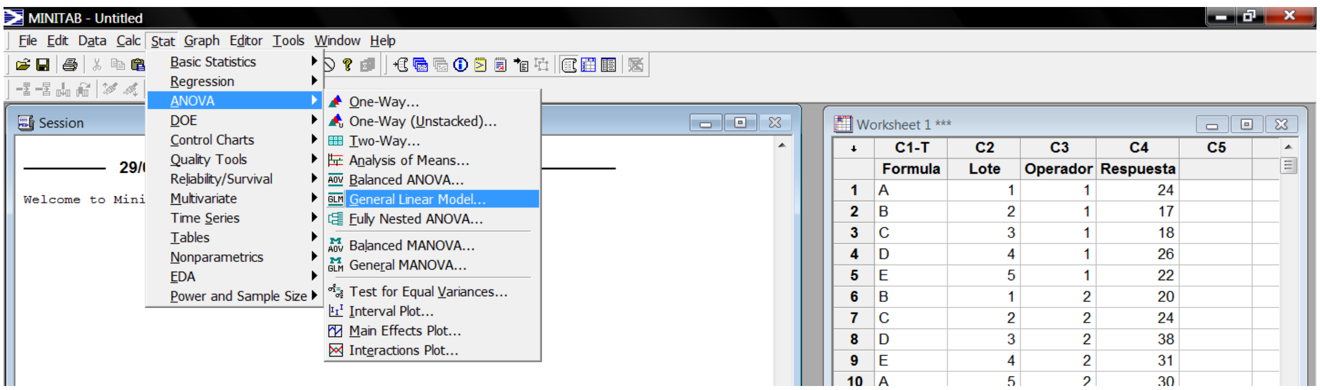
Observe que otra forma de representar este arreglo es en columnas como sigue:

Formula	**Lote**	**Operador**	**Respuesta**
A	1	1	24
B	2	1	17
C	3	1	18
D	4	1	26
E	5	1	22
B	1	2	20
C	2	2	24
D	3	2	38
E	4	2	31
A	5	2	30
C	1	3	19
D	2	3	30
E	3	3	26
A	4	3	26
B	5	3	20
D	1	4	24

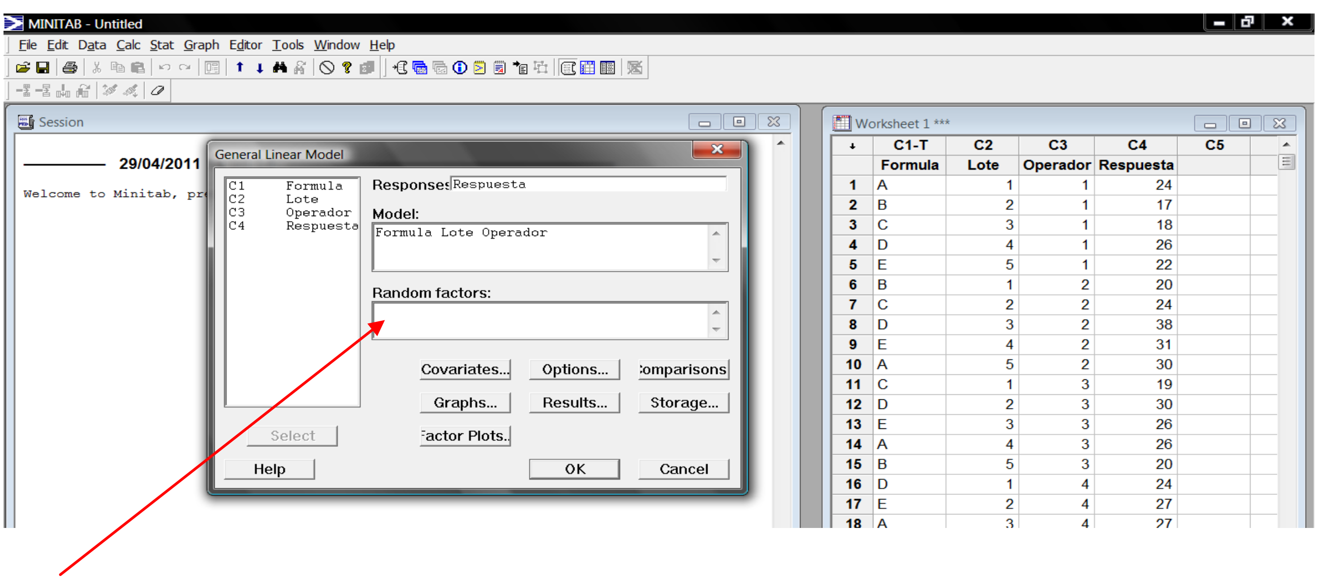
E	2	4	27
A	3	4	27
B	4	4	23
C	5	4	29
E	1	5	24
A	2	5	36
B	3	5	21
C	4	5	22
D	5	5	31

Observe que el operador 1, lote 1 y formula 1 tiene un valor de 24.

En minitab seleccione Stat / Anova / General Linear Model



Ingrese la columna de respuesta (los 25 datos o corridas), así como los tres factores dentro de modelo y dele ok.



Nota: Una variable es del tipo fijo si sólo se utilizan un número fijo de niveles ya sea por decisión del investigador o porque no se dispone de más. Por tanto una variable aleatoria (random) es la que se usan un número de niveles de una población más grande. *ejemplo* el factor operador se usó en 5 niveles (1,2,3,4,5) porque sólo había cinco operadores, por tanto es fijo. Si se hubieran seleccionado los cinco operadores de un grupo mayor (10 operadores disponibles) entonces sería una variable aleatoria (random) y se hubiera indicado en random factors.

Hoja de sesión

La hoja de sesión muestra primeramente los tres factores y el tipo que pertenecen (fixed= fijo) y 5 niveles para cada factor.

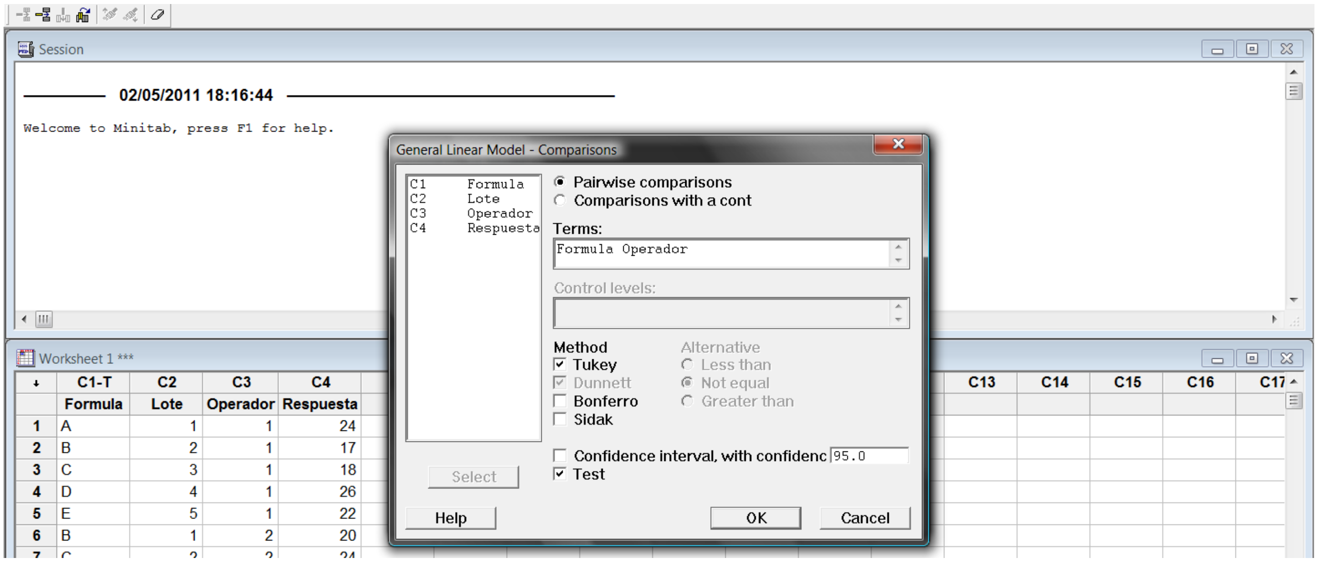
```
General Linear Model: Respuesta versus Formula, Lote, Operador
Factor    Type   Levels  Values
Formula   fixed       5  A, B, C, D, E
Lote      fixed       5  1, 2, 3, 4, 5
Operador  fixed       5  1, 2, 3, 4, 5


Analysis of Variance for Respuesta, using Adjusted SS for Tests

Source    DF  Seq SS  Adj SS  Adj MS     F      P
Formula    4  330.00  330.00   82.50  7.73  0.003
Lote       4   68.00   68.00   17.00  1.59  0.239
Operador   4  150.00  150.00   37.50  3.52  0.040
Error     12  128.00  128.00   10.67
Total     24  676.00


S = 3.26599   R-Sq = 81.07%   R-Sq(adj) = 62.13%


Unusual Observations for Respuesta

Obs  Respuesta      Fit  SE Fit  Residual  St Resid
 20    29.0000  24.0000  2.3551    5.0000      2.21 R
 22    36.0000  31.4000  2.3551    4.6000      2.03 R

R denotes an observation with a large standardized residual
```

Pvalue 0.003 es menor a α=0.05, por tanto se observa que las formulas producen tiempos de combustión diferentes, es decir la diferencia es significativa.

Operador también produce resultados diferentes (es significativo).

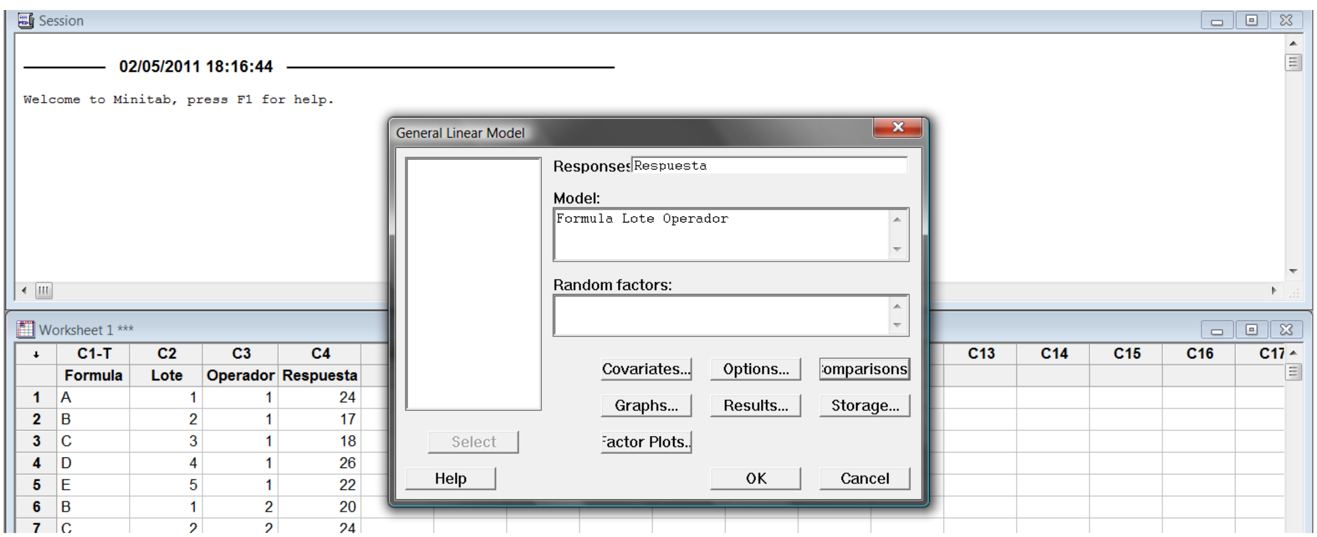
En cambio Lote no produce diferencia, ya que su pvalue es

Dado que ya sabemos que formula y operador si contribuyen significativamente a los tiempos de combustión, procederemos a determinar que formula y que operador son diferentes, es decir, realizaremos comparaciones de pares de medias, solo que en lugar del método LSD, utilizaremos el método de Tukey



Volvemos a correr el estudio mediante ***General linear model***, damos click en la casilla de ***comparison***, activamos ***pairwise comparisons-Tukey-test***, e ingresamos los factores *Formula y operador* en la celda ***Terms***. Con esto Minitab comparará exclusivamente estos dos factores en sus diferentes niveles (5 niveles) para poder determinar que factor y que nivel son diferentes, en base a los valores de pvalue.

Posteriormente corremos minitab dando click en ok



La hoja de sesión nos muestra primeramente la comparación del factor Formula en sus cinco niveles, primero comparando A con B, C, D y E. Recuerde que el promedio de A se resta a B, luego a C, D y E, por lo que Difference of Means (diferencia de promedios) es la diferencia esperada de B-A, C-A, D-A y E-A

```
Tukey Simultaneous Tests
Response Variable Respuesta
All Pairwise Comparisons among Levels of Formula
Formula = A  subtracted from:

          Difference       SE of            Adjusted
Formula     of Means  Difference  T-Value    P-Value
B             -8.400       2.066   -4.067     0.0111
C             -6.200       2.066   -3.002     0.0684
D              1.200       2.066    0.581     0.9754
E             -2.600       2.066   -1.259     0.7194
```

Pvalue 0.0111 menor que un α=0.05 indica que estadísticamente A y B son diferentes.

Por el contrario un Pvalue 0.0684 mayor que un α=0.05 indica que estadísticamente A y B son iguales… por lo tanto

```
Formula = B  subtracted from:

          Difference       SE of            Adjusted
Formula     of Means  Difference  T-Value    P-Value
C              2.200       2.066    1.065     0.8205
D              9.600       2.066    4.648     0.0042
E              5.800       2.066    2.808     0.0944
```

Comparando B con el resto de niveles

```
Formula = C  subtracted from:

          Difference       SE of            Adjusted
Formula     of Means  Difference  T-Value    P-Value
D              7.400       2.066    3.583     0.0254
E              3.600       2.066    1.743     0.4462
```

Comparando C con el resto de niveles

```
Formula = D  subtracted from:

          Difference       SE of            Adjusted
Formula     of Means  Difference  T-Value    P-Value
E             -3.800       2.066   -1.840     0.3967
```

Comparando D con E

D=E

Por lo tanto, a un nivel de significancia de 0.05 tenemos las siguientes relaciones

A≠B, A=C, A=D, A=E, B=C, B≠D, B=E, C≠D, C=E, D=E

Y dado los promedios de cada nivel A=28.6, B=20.2, C=22.4, D=29.8, E=26

Podemos decir que B tiene el menor tiempo de combustión promedio posible, pero que es igual a C y E, por lo tanto estadísticamente tanto B como C y E, darán un tiempo de combustión similar. Como siempre habrá que considerar más información para emitir una decisión final, como puede ser, precio, facilidad de fabricación, disponibilidad de la fórmula entre otros.

```
Tukey Simultaneous Tests
Response Variable Respuesta
All Pairwise Comparisons among Levels of Operador
```

```
Operador = 1  subtracted from:

          Difference       SE of            Adjusted
Operador    of Means  Difference  T-Value    P-Value
2              7.200       2.066    3.486     0.0300
3              2.800       2.066    1.356     0.6646
4              4.600       2.066    2.227     0.2342
5              5.400       2.066    2.614     0.1293


Operador = 2  subtracted from:

          Difference       SE of            Adjusted
Operador    of Means  Difference  T-Value    P-Value
3             -4.400       2.066   -2.130     0.2692
4             -2.600       2.066   -1.259     0.7194
5             -1.800       2.066   -0.871     0.9020


Operador = 3  subtracted from:

          Difference       SE of            Adjusted
Operador    of Means  Difference  T-Value    P-Value
4              1.800       2.066   0.8714     0.9020
5              2.600       2.066   1.2587     0.7194


Operador = 4  subtracted from:

          Difference       SE of            Adjusted
Operador    of Means  Difference  T-Value    P-Value
5             0.8000       2.066   0.3873     0.9946
```

En el caso del factor operador, tenemos según la hoja de sesión que a un nivel de significancia de 0.05 tenemos las siguientes relaciones:

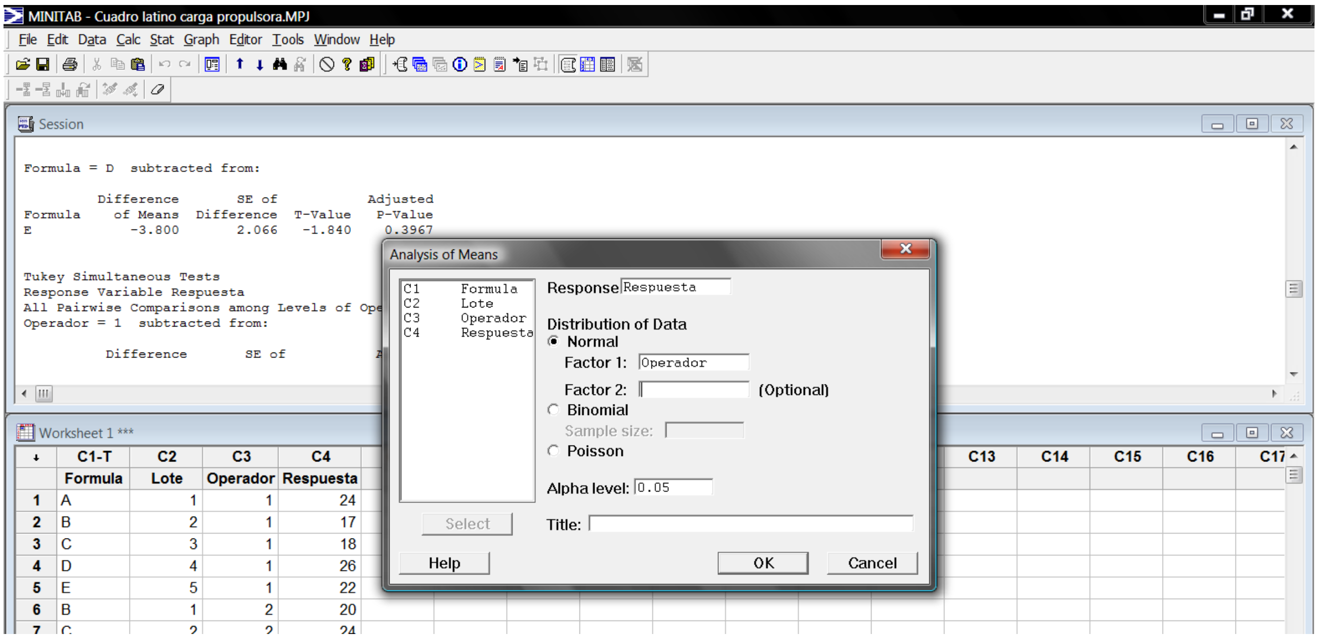
1≠2, 1=3, 1=4, 1=5, 2=3, 2=4, 2=5, 3=4, 3=5, 4=5

Es decir estadísticamente el operador 1 es diferente al menos de otro, y todas las demás relaciones en pares son estadísticamente iguales.

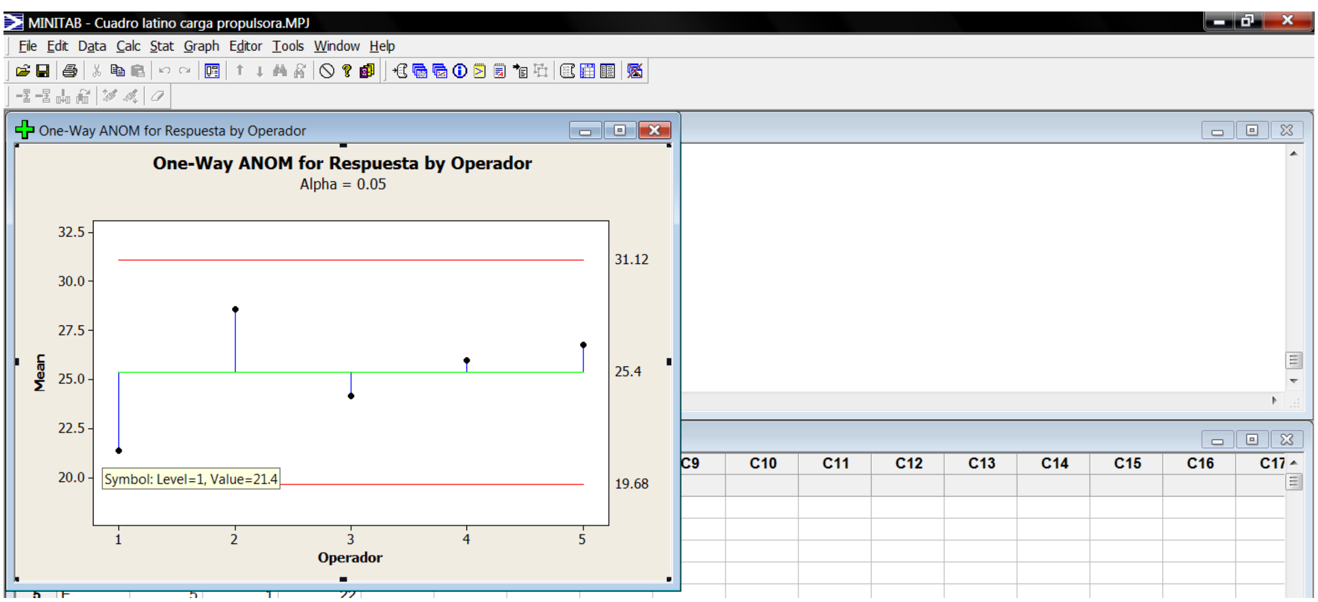
Y dado los promedios de operador en cada nivel: 1=21.4, 2=28.6, 3=24.2, 4=26 y 5=26.8

Podemos concluir que el operador 1 es el que más influye en el tiempo de combustión de manera significativa, ya que en promedio el tiempo es de 21.4 milisegundo cuando él prepara la formula. Sin embargo es estadísticamente igual al operador 3, 4 y 5, por lo que cualquiera de ellos podría dar un resultado similar. Otra consideración es homologar la forma que elaboran la formula, por lo que sería conveniente revisar como lo hacen y dar entrenamiento en la medida de lo posible.

Nota: para obtener los promedios de los niveles de cada factor, por ejemplo los promedios de los 5 niveles del factor operador, se puede utilizar la función analysis of Means, siguiendo la relación *Stat / ANOVA /Analysis of means* tal como se muestra en la figura.



Y el resultado es la siguiente gráfica, donde vemos el promedio del operador 1 simplemente colocando el cursor encima del punto.



B) DESARROLLO

practique analizando los problemas que proporcione el profesor

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

- Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
- Douglas C. Montgomery (2009). Design and Analysis of Experiments
- Douglas C. Montgomery y George C. Runger, (2005). Probabilidad y Estadística aplicadas a la Ingeniería. Editorial Limusa Wiley, segunda edición.
- Seymour Lipschutz y John Schiller, (2000). Introducción a la Probabilidad y Estadística. Editorial Mc Graw Hill.
- Humberto Gutiérrez y Román De La Vara, (2009). Control Estadístico de Calidad y Seis Sigma. Editorial Mc Graw Hill, segunda edición.

6.- ANEXOS:

NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**DISEÑO DE CUADRO GRECOLATINO**	**PRÁCTICA NÚMERO**	**8**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

SOFTWARE REQUERIDO
MINITAB y EXCEL
OBSERVACIONES-COMENTARIOS

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

A continuación se analizan experimentos con cuatro factores simultaneos por medio del diseño de cuadro Grecolatino y el empleo del software Minitab y Excel.

2.- OBJETIVO (COMPETENCIA):

Analizar experimentos por medio de la comparación entre 4 tipos de factores y decidir cual es la mejor opción utilizando el diseño Grecolatino y el empleo del software Minitab.

3.- TEORIA:

Al igual que el diseño de cuadro latino, el grecolatino considera un factor de interés (tratamiento), sólo que en lugar de 2 factores de bloqueo, aquí se puede utilizar 3 factores de bloqueo, para lo cual se organiza de la siguiente manera:

Letras latinas= Factores de interés (tratamiento), **Renglones** = Factor de bloque 1

Columnas= Factor de bloqueo 2, **Letras griegas**= Factor de bloqueo 3

El siguiente es un arreglo de cuadro grecolatino cuando interesa 4 niveles de tratamiento (4x4)

	Columna			
Renglon	1	2	3	4
1	Aα	Bβ	Cγ	Dδ
2	Bδ	Aγ	Dβ	Cα
3	Cβ	Dα	Aδ	Bγ
4	Dγ	Cδ	Bα	Aβ

Observe como cada letra griega (α,β,γ,δ) aparece una y sólo una vez con cada letra latina (A,B,C,D) en este arreglo. De esta manera se puede investigar cuatro factores, cada uno con k niveles en solo k^2 corridas. Para la tabla anterior de 4x4, k=4, por lo que habría k^2=16 corridas o datos.

A continuación se muestra la tabla Anova del diseño Grecolatino.

ANOVA DISEÑO EN CUADRO GRECOLATINO

Fuente de variación	**Suma de cuadrados (SC)**	**grados de libertad (gl)**	**Cuadrados medios (CM)**	**Estadístico F_0**	**Valor P**
Tratamiento	$SC_{Trat}=\sum_{j=1}^{k}\frac{Y_{.j..}^2}{k}-\frac{Y_{....}^2}{N}$	$k-1$	$CM_{Trat}=\frac{SC_{Trat}}{k-1}$	$\frac{CM_{Trat}}{CM_{Error}}$	$P(F>F_0)$
Renglones (B1)	$SC_{B1}=\sum_{i=1}^{k}\frac{Y_{i...}^2}{k}-\frac{Y_{....}^2}{N}$	$k-1$	$CM_{B1}=\frac{SC_{B1}}{k-1}$	$\frac{CM_{B1}}{CM_{Error}}$	$P(F>F_0)$
Columnas (B2)	$SC_{B2}=\sum_{m=1}^{k}\frac{Y_{...m}^2}{k}-\frac{Y_{....}^2}{N}$	$k-1$	$CM_{B2}=\frac{SC_{B2}}{k-1}$	$\frac{CM_{B2}}{CM_{Error}}$	$P(F>F_0)$
Letras griegas (B3)	$SC_{B3}=\sum_{l=1}^{k}\frac{Y_{..l.}^2}{k}-\frac{Y_{....}^2}{N}$	$k-1$	$CM_{B3}=\frac{SC_{B3}}{k-1}$	$\frac{CM_{B3}}{CM_{Error}}$	$P(F>F_0)$
Error	$SC_{Error}=SC_T-SC_{Trat}-SC_{B1}-SC_{B2}-SC_{B3}$	$(k-3)(k-1)$	$CM_{Error}=\frac{SC_{Error}}{(k-3)(k-1)}$		
Total	$SC_T=\sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{l=1}^{k}Y_{ijl}^2-\frac{Y_{...}^2}{N}$	k^2-1			

4.- DESCRIPCIÓN

A) PROCEDIMIENTO Y DURACION DE LA PRÁCTICA:

1.- Encienda su computadora y abra el programa Minitab.

2.- Resuelva los siguientes problemas utilizando Excel y Minitab

Diseño de cuadro Grecolatino

Suponga que en el experimento de la carga propulsora, existe un factor adicional que es los montajes de prueba. Sea que hay cinco montajes de prueba denotados por las letras griegas, se genera una tabla grecolatina de 5x5 como la que sigue y se procede a realizar la experimentación cuidando la aleatoriedad de la obtención de las corridas, dando como resultado los siguientes datos:

Operadores

Lote de materia prima	1		2		3		4		5	
1	Aα	24	Bγ	20	Cε	19	Dβ	24	Eδ	24
2	Bβ	17	Cδ	24	Dα	30	Eγ	27	Aε	36
3	Cγ	18	Dε	38	Eβ	26	Aδ	27	Bα	21
4	Dδ	26	Eα	31	Aγ	26	Bε	23	Cβ	22
5	Eε	22	Aβ	30	Bδ	20	Cα	29	Dγ	31

Donde las letras latinas contienen los datos del tratamiento (Formula) y las letras griegas contiene los montajes de prueba. Con esta información genera la tabla de Anova correspondiente y pruebe la hipótesis de que las formulas producen la misma rapidez de combustión en milisegundos.

Posteriormente determine si aplica, ¿cuál formula es la más rápida?...utilice el criterio del método LSD que para un cuadro grecolatino es el siguiente:

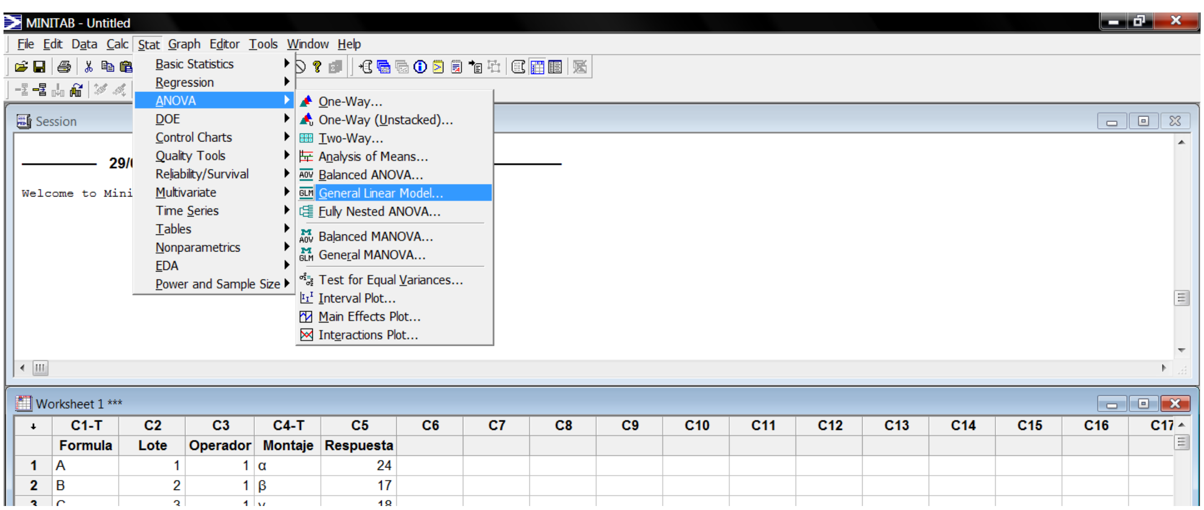
$$LSD = t_{\frac{\alpha}{2},(k-3)(k-1)}\sqrt{\frac{2CM_{Error}}{k}}$$

Observe que otra forma de representar este arreglo es en columnas como sigue:

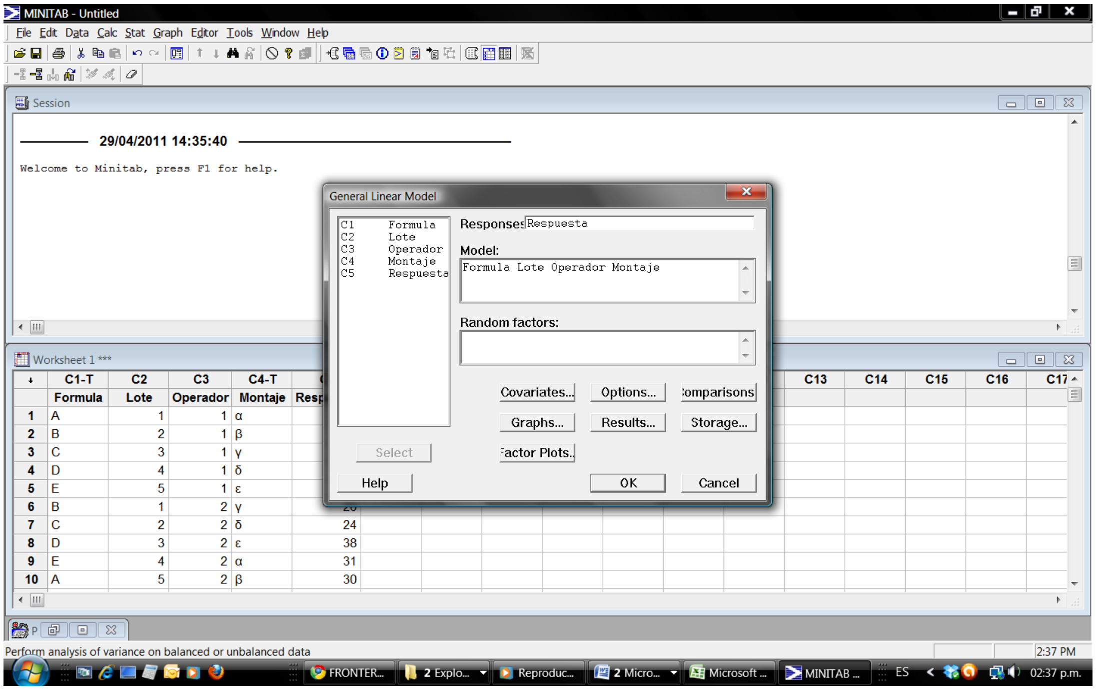
Formula	Lote	Operador	Montaje	Respuesta
A	1	1	α	24
B	2	1	β	17
C	3	1	ν	18
D	4	1	δ	26
E	5	1	ε	22
B	1	2	ν	20
C	2	2	δ	24
D	3	2	ε	38
E	4	2	α	31
A	5	2	β	30
C	1	3	ε	19
D	2	3	α	30
E	3	3	β	26
A	4	3	ν	26
B	5	3	δ	20
D	1	4	β	24
E	2	4	ν	27
A	3	4	δ	27
B	4	4	ε	23
C	5	4	α	29
E	1	5	δ	24
A	2	5	ε	36
B	3	5	α	21
C	4	5	β	22
D	5	5	ν	31

Observe que la formula A con operador 1, lote 1 y montaje α tiene un valor de 24.

En minitab seleccione Stat / Anova / General Linear Model



Ingrese la columna de respuesta (los 25 datos o corridas), así como los cuatro factores dentro de modelo y dele ok.



Hoja de sesión

La hoja de sesión muestra primeramente los cuatro factores y el tipo que pertenecen (fixed= fijo) y 5 niveles para cada factor.

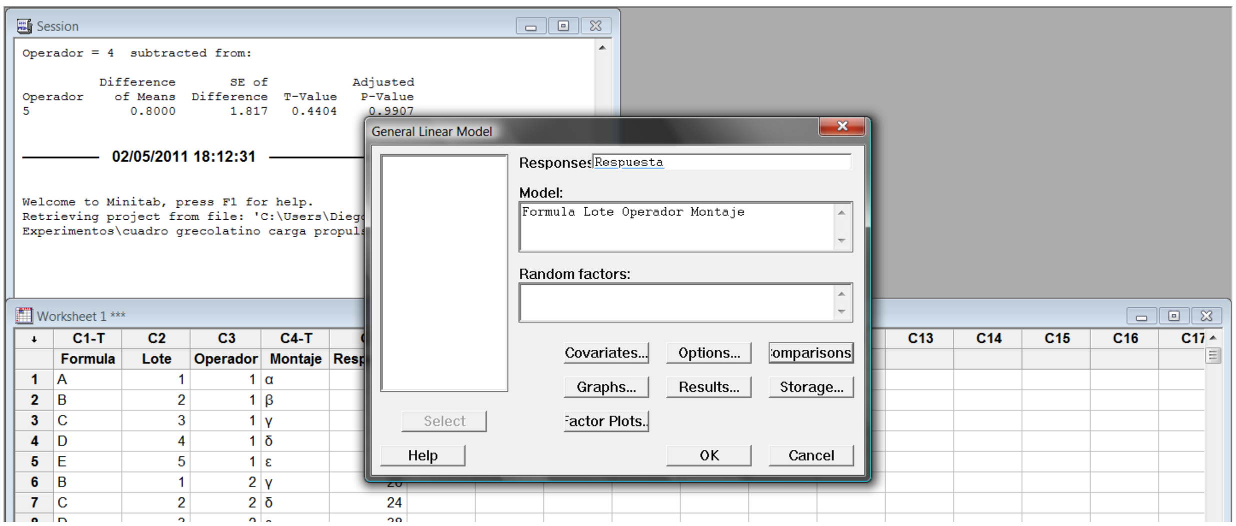
General Linear Model: Respuesta versus Formula, Lote, Operador, Montaje

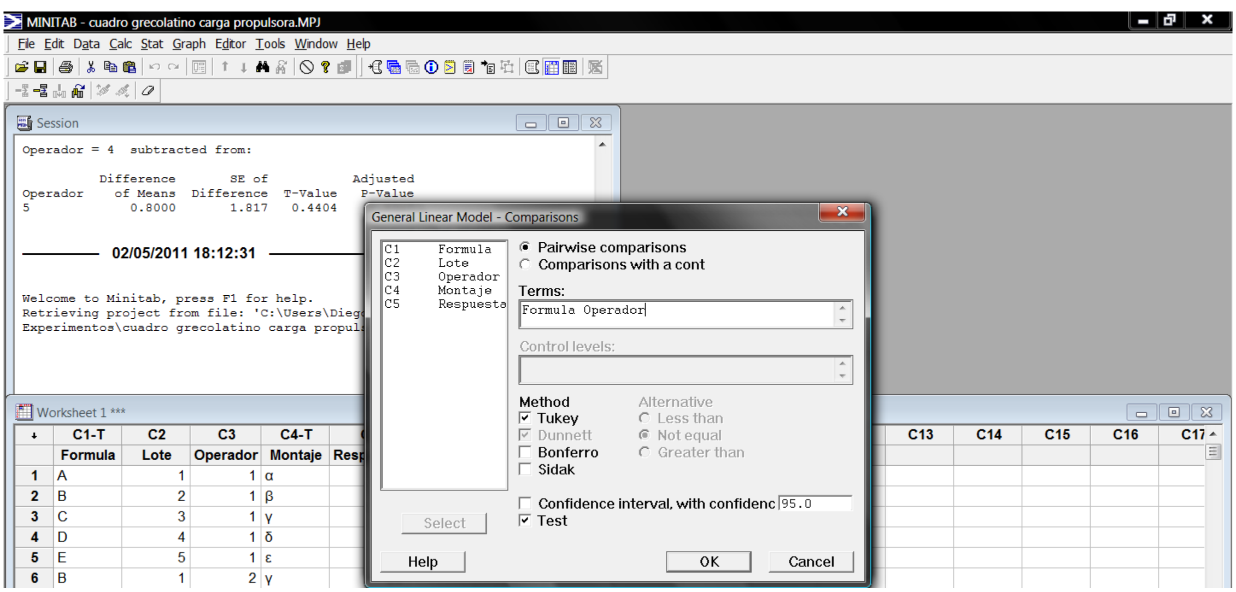
```
Factor    Type   Levels  Values
Formula   fixed       5  A, B, C, D, E
Lote      fixed       5  1, 2, 3, 4, 5
Operador  fixed       5  1, 2, 3, 4, 5
Montaje   fixed       5  α , β , γ , δ , ε


Analysis of Variance for Respuesta, using Adjusted SS for Tests

Source    DF   Seq SS   Adj SS  Adj MS      F      P
Formula    4  330.000  330.000  82.500  10.00  0.003
Lote       4   68.000   68.000  17.000   2.06  0.178
Operador   4  150.000  150.000  37.500   4.55  0.033
Montaje    4   62.000   62.000  15.500   1.88  0.208
Error      8   66.000   66.000   8.250
Total     24  676.000


S = 2.87228   R-Sq = 90.24%   R-Sq(adj) = 70.71%


Unusual Observations for Respuesta

Obs  Respuesta      Fit  SE Fit  Residual  St Resid
 20    29.0000  25.6000  2.3685    3.4000    2.09 R

R denotes an observation with a large standardized residual.
```

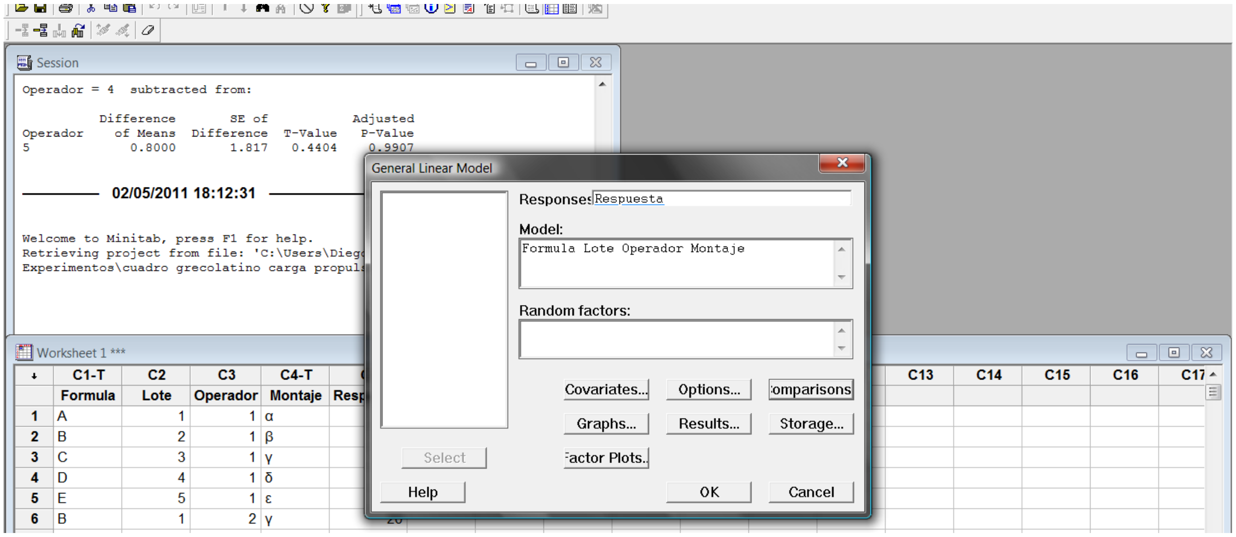
Pvalue 0.003 es menor a α=0.05, por tanto se observa que las **formulas** producen tiempos de combustión diferentes.

En este sentido **operador** también produce resultados diferentes. Sin embargo **lote** y **montaje** tienen p value mayor a 0.05, es decir no son significativos, por tanto al usar cualquier lote o montaje darían resultados parecidos en el tiempo

Dado que ya sabemos que formula y operador si contribuyen significativamente a los tiempos de combustión, procederemos a determinar que formula y que operador don diferentes, es decir realizaremos comparaciones de pares de medias, solo que en lugar del método LSD, utilizaremos el método de Tukey, por lo que volvemos a correr el estudio mediante *General linear model*, damos click en la casilla de *comparison*, activamos *pairwise comparisons-Tukey-test*, e ingresamos los factores *Formula y operador*. Con esto Minitab comparará exclusivamente estos dos factores en sus diferentes niveles (5 niveles) para poder determinar que factor y en que nivel son diferentes, esto empleando los valores de pvalue.







La hoja de sesión donde se especifica los resultados del factor *Formula* es la siguiente

```
Tukey Simultaneous Tests
Response Variable Respuesta
All Pairwise Comparisons among Levels of Formula
Formula = A  subtracted from:

           Difference        SE of           Adjusted
Formula      of Means   Difference  T-Value   P-Value
B              -8.400        1.817   -4.624    0.0108
C              -6.200        1.817   -3.413    0.0529
D               1.200        1.817    0.661    0.9597
E              -2.600        1.817   -1.431    0.6270


Formula = B  subtracted from:

           Difference        SE of           Adjusted
Formula      of Means   Difference  T-Value   P-Value
C               2.200        1.817    1.211    0.7463
D               9.600        1.817    5.285    0.0048
E               5.800        1.817    3.193    0.0715


Formula = C  subtracted from:

           Difference        SE of           Adjusted
Formula      of Means   Difference  T-Value   P-Value
D               7.400        1.817    4.074    0.0218
E               3.600        1.817    1.982    0.3526


Formula = D  subtracted from:

           Difference        SE of           Adjusted
Formula      of Means   Difference  T-Value   P-Value
E              -3.800        1.817   -2.092    0.3087
```

Por lo tanto, a un nivel de significancia de 0.05 tenemos las siguientes relaciones

A≠B, A=C, A=D, A=E, B=C, B≠D, B=E, C≠D, C=E, D=E

Y dado los promedios de cada nivel A=28.6, B=20.2, C=22.4, D=29.8, E=26

Podemos decir que se sigue manteniendo el resultado que en cuadro latino, donde B tiene el menor tiempo de combustión promedio posible, pero que es igual a C y E, por lo tanto estadísticamente tanto B como C y E, darán un tiempo de combustión similar. Como siempre habrá que considerar más información para emitir una decisión final, como puede ser, precio, facilidad de fabricación, disponibilidad de la fórmula entre otros.

La hoja de sesión donde se especifica los resultados del factor *Operador* es la siguiente

```
Tukey Simultaneous Tests
Response Variable Respuesta
All Pairwise Comparisons among Levels of Operador
Operador = 1  subtracted from:

          Difference       SE of            Adjusted
Operador    of Means  Difference  T-Value    P-Value
2              7.200       1.817    3.963     0.0253
3              2.800       1.817    1.541     0.5670
4              4.600       1.817    2.532     0.1754
5              5.400       1.817    2.973     0.0966


Operador = 2  subtracted from:

          Difference       SE of            Adjusted
Operador    of Means  Difference  T-Value    P-Value
3             -4.400       1.817   -2.422     0.2028
4             -2.600       1.817   -1.431     0.6270
5             -1.800       1.817   -0.991     0.8525


Operador = 3  subtracted from:

          Difference       SE of            Adjusted
Operador    of Means  Difference  T-Value    P-Value
4              1.800       1.817   0.9909     0.8525
5              2.600       1.817   1.4313     0.6270


Operador = 4  subtracted from:

          Difference       SE of            Adjusted
Operador    of Means  Difference  T-Value    P-Value
5             0.8000       1.817   0.4404     0.9907
```

En el caso del factor operador, tenemos según la hoja de sesión que a un nivel de significancia de 0.05 tenemos las siguientes relaciones:

1≠2, 1=3, 1=4, 1=5, 2=3, 2=4, 2=5, 3=4, 3=5, 4=5

Es decir estadísticamente el operador 1 es diferente al menos de otro, y todas las demás relaciones en pares son estadísticamente iguales.

Y dado los promedios de operador en cada nivel: 1=21.4, 2=28.6, 3=24.2, 4=26 y 5=26.8

*Podemos concluir que se mantiene la decisión hecha al utilizar 3 factores (cuadro latino) donde el operador 1 es el que más influye en el tiempo de combustión de manera significativa, ya que en promedio el tiempo es de 21.4 milisegundo cuando él prepara la formula. Sin embargo es estadísticamente igual al operador 3, 4 y 5, por lo que cualquiera de ellos podría dar un resultado similar. Otra consideración es homologar la forma que elaboran la formula, por lo que sería conveniente revisar como lo hacen y dar entrenamiento en la medida de lo posible***.**

B) DESARROLLO

practique analizando los problemas que proporcione el profesor

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

- Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
- Douglas C. Montgomery (2009). Design and Analysis of Experiments
- Douglas C. Montgomery y George C. Runger, (2005). Probabilidad y Estadística aplicadas a la Ingeniería. Editorial Limusa Wiley, segunda edición.
- Seymour Lipschutz y John Schiller, (2000). Introducción a la Probabilidad y Estadística. Editorial Mc Graw Hill.
- Humberto Gutiérrez y Román De La Vara, (2009). Control Estadístico de Calidad y Seis Sigma. Editorial Mc Graw Hill, segunda edición.

6.- ANEXOS:

NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**DISEÑO DE BLOQUES INCOMPLETOS BALANCEADOS**	**PRÁCTICA NÚMERO**	**9**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

SOFTWARE REQUERIDO
MINITAB y EXCEL
OBSERVACIONES-COMENTARIOS

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

En cierto experimentos en los que se utilizan diseño de bloques aleatorizados, quizá no sea posible correr todas las combinaciones de los tratamientos en cada bloque, esto debido a limitaciones del aparato experimental, de instalaciones, del material, tamaño físico del bloque o por cuestiones técnicas o económicas. Para estas situaciones es posible utilizar diseño de bloques aleatorizados en los que cada tratamiento no está presente en cada bloque. Estos diseños se conocen como *diseño de bloques incompletos aleatorizados*.

A continuación se analizan experimentos con diseño de bloques aleatorizados por medio del empleo del software Minitab y Excel.

2.- OBJETIVO (COMPETENCIA):

Analizar experimentos por medio de la comparación entre factores con diseño de bloques incompletos y decidir cual es la mejor opción utilizando el software Minitab.

3.- TEORIA:

Un ingeniero químico piensa que el tiempo de reacción de un proceso químico es una función del tipo de catalizador empleado. Se están investigando cuatro catalizadores, donde el procedimiento experimental consiste en seleccionar un lote de materia prima, cargar la planta piloto, aplicar cada catalizador en una corrida separada de la planta piloto y observar el tiempo de reacción. Debido a que las variaciones en los lotes de materia prima pueden afectar el desempeño de los catalizadores, el ingeniero decide usar los lotes de materia prima como bloques. Sin embargo, cada lote es apenas lo suficientemente grande para permitir que se prueben tres catalizadores. Por lo tanto, debe usarse un diseño de bloques incompletos aleatorizados, el cual se corre de manera aleatoria para cada bloque, quedando los siguientes resultados:

Tipo de catalizador		**Lote de materia prima (Bloque)**			
(Tratamiento)	**1**	**2**	**3**	**4**	**Yi.**
1	73	74	-	71	**218**
2	-	75	67	72	**214**
3	73	75	68	-	**216**
4	75	-	72	75	**222**
Y.j	**221**	**224**	**207**	**218**	**Y..= 870**

En este caso se toman los siguientes valores:

Número de tratamientos = a = **4**

Número de bloques = b = **4**

Número de tratamientos dentro de cada bloque = k = **3**

Número de veces que ocurre un tratamiento en el diseño = r = **3**

Número total de corridas/datos = N = **12**

Si a=b, se dice que es un diseño simétrico

Asimismo, el número de veces que cada par de tratamientos aparece en el mismo bloque es

Fuente de variación	Suma de cuadrados (SC)	grados de libertad (gl)	Cuadrados medios (CM)	Estadístico F_0	Valor P

$$\lambda = \frac{r(k-1)}{a-1}$$

Donde Q_i es el total ajustado del tratamiento i-esimo, el cual se calcula como:

$$Q_i = Y_{i.} - \frac{1}{k}\sum_{j=1}^{b} n_{ij} * Y_{.j}$$

$con\ n_{ij} = 1\ si\ el\ tratamiento\ aparece\ en\ el\ bloque\ j\ y\ 0\ si\ no.$

La siguiente es la tabla Anova

Tratamiento Ajustado	$SC_{TratAj} = \frac{k\sum_{i=1}^{a} Q_i^2}{\lambda a}$	$a-1$	$CM_{TratAj} = \frac{SC_{TratAj}}{a-1}$	$\frac{CM_{TratAj}}{CM_{Error}}$	$P(F >$
Bloques	$SC_B = \frac{1}{k}\sum_{j=1}^{b} Y_{.j}^2 - \frac{Y_{..}^2}{N}$	$b-1$	$CM_B = \frac{SC_B}{b-1}$	$\frac{CM_B}{CM_{Error}}$	$P(F >$
Error	$SC_{Error} = SC_T - SC_{TratAj} - SC_B$	$N-a-b+1$	$CM_{Error} = \frac{SC_{Error}}{N-a-b+1}$		
Total	$SC_T = \sum_{i=1}^{a}\sum_{j=1}^{b} Y_{ij}^2 - \frac{Y_{..}^2}{N}$	$N-1$			

$$\lambda = \frac{r(k-1)}{a-1}$$

$$Q_i = Y_{i.} - \frac{1}{k}\sum_{j=1}^{b} n_{ij} * Y_{.j} \quad con\ n_{ij} = 1\ si\ el\ trat\ aparece\ en\ el\ bloque\ j\ y\ 0\ si\ no$$

a= Número de tratamientos k= Número de tratamientos dentro de cada bloque

b= Número de bloques r= Número de veces que ocurre un tratamiento en el diseño

N= Número total de corridas/datos

4.- DESCRIPCIÓN

A) PROCEDIMIENTO Y DURACION DE LA PRÁCTICA:

1.- Encienda su computadora y abra el programa Minitab.

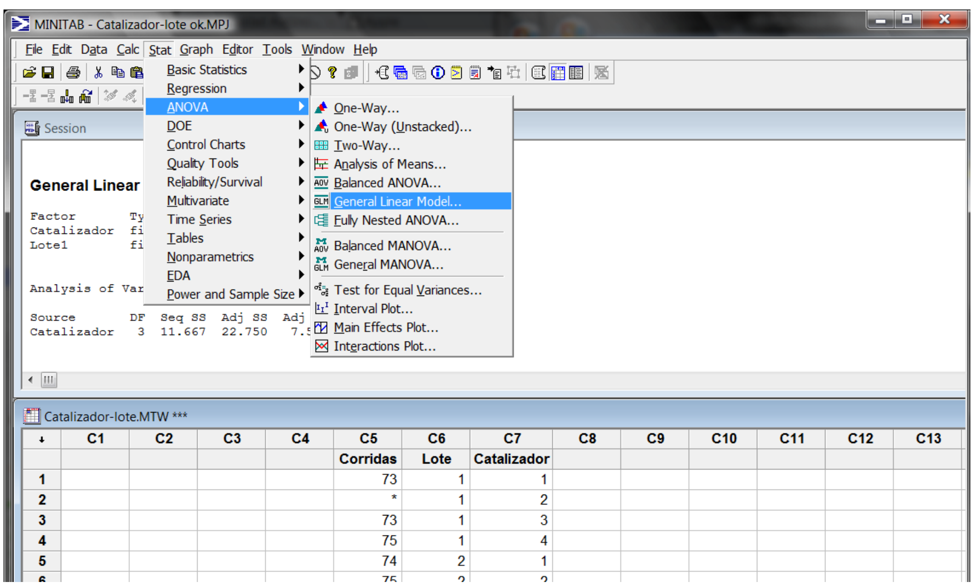
2.- Resuelva los siguientes problemas utilizando Excel y Minitab

Diseño de Bloques Incompletos Balanceados

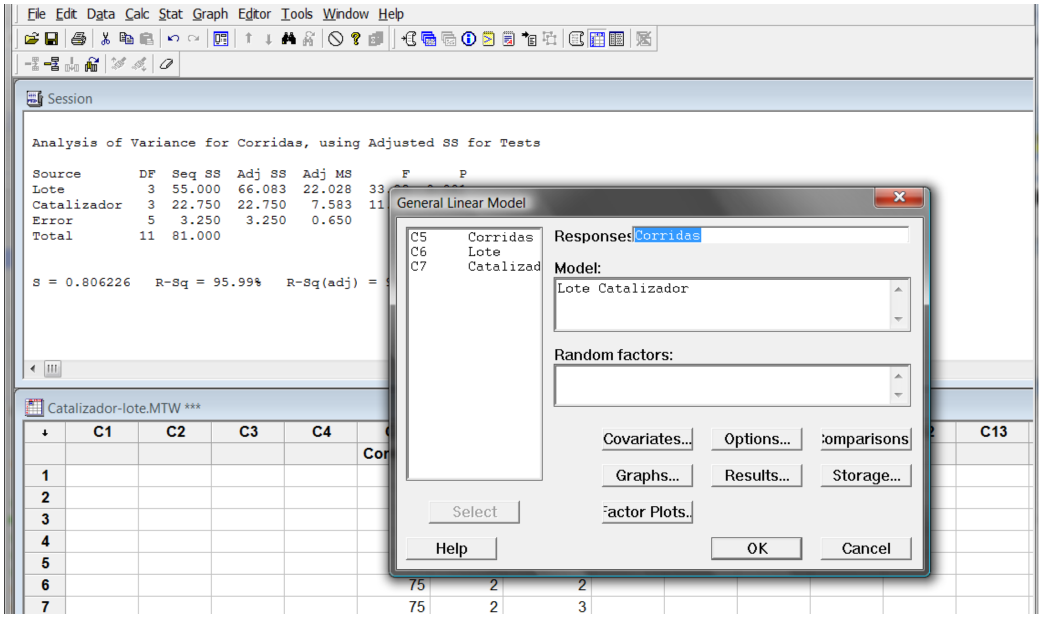
Observe que otra forma de acomodar los datos del ejercicio 4, es como sigue:

Corridas	Lote	Catalizador
73	1	1
*	1	2
73	1	3
75	1	4
74	2	1
75	2	2
75	2	3
*	2	4
*	3	1
67	3	2
68	3	3
72	3	4
71	4	1
72	4	2
*	4	3
75	4	4

De esta manera se pueden ingresar los datos en Minitab utilizando *Stat / Anova / General Linear Model*



Posteriormente se ingresa en la *respuesta*, la columna donde están los resultados (en este ejemplo se llama corridas) y los factores se ingresan en *Modelo* y se da click en *ok*.



La hoja se sesión quedaría como sigue:

Compárela con la obtenida manualmente en Excel.

General Linear Model: Corridas versus Lote, Catalizador

Factor	Type	Levels	Values
Lote	fixed	4	1, 2, 3, 4
Catalizador	fixed	4	1, 2, 3, 4

Analysis of Variance for Corridas, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Lote	3	55.000	66.083	22.028	33.89	0.001
Catalizador	3	22.750	22.750	7.583	11.67	0.011
Error	5	3.250	3.250	0.650		
Total	11	81.000				

S = 0.806226   R-Sq = 95.99%   R-Sq(adj) = 91.17%

Pvalue =0.001 es menor a α=0.05, por tanto se observa que los **Lotes** son significativos, es decir, dependiendo cual lote se usa, influye en el tiempo del proceso químico.

En este sentido, **el catalizador** a emplear también produce resultados diferentes, en otras palabras, es significativo, por lo que se rechaza la H0 de que los catalizadores producen

Dado que nos interesa saber que lote y que catalizador son los mejores, procedemos a hacer una comparación entre los cuatro niveles de los dos factores, y lo hacemos mediante Tukey…volvemos a dar Stat / Anova / *General Linear Model* e ingresamos las mismas columnas, posteriormente damos click en la casilla de *comparison*, activamos *pairwise comparisons-Tukey-test*, e ingresamos los factores *Catalizador y lote*. Con esto Minitab comparará exclusivamente estos dos factores en sus diferentes niveles (4 niveles) para poder determinar que factor y en que nivel son diferentes, esto empleando los valores de pvalue.

```
Tukey Simultaneous Tests
Response Variable Corridas
All Pairwise Comparisons among Levels of Lote
Lote = 1  subtracted from:

      Difference       SE of             Adjusted
Lote    of Means  Difference  T-Value     P-Value
2          2.125      0.6982    3.043      0.0970
3         -4.750      0.6982   -6.803      0.0040
4         -0.875      0.6982   -1.253      0.6247


Lote = 2  subtracted from:

      Difference       SE of             Adjusted
Lote    of Means  Difference  T-Value     P-Value
3         -6.875      0.6982   -9.847      0.0007
4         -3.000      0.6982   -4.297      0.0281


Lote = 3  subtracted from:

      Difference       SE of             Adjusted
Lote    of Means  Difference  T-Value     P-Value
4          3.875      0.6982    5.550      0.0098
```

Cuando se analiza el factor lote y se resta la media de nivel 1 (lote 1) a los demás niveles, por ejemplo cuando se resta a (lote 2), Pvalue =0.0970 es mayor a α=0.05, por tanto se observa que los **Lotes 1 y 2 NO SON SIGNIFICATIVOS**, es decir, son estadísticamente iguales.

Sin embargo cuando lote 1 se resta a lote 3, Pvalue =0.0040 es menor a α=0.05, entonces **lotes 1 y 3 SON SIGNIFICATIVOS**, es decir,

```
Tukey Simultaneous Tests
Response Variable Corridas
All Pairwise Comparisons among Levels of Catalizador
Catalizador = 1  subtracted from:

             Difference       SE of             Adjusted
Catalizador    of Means  Difference  T-Value     P-Value
2                0.2500      0.6982   0.3581      0.9825
3                0.6250      0.6982   0.8951      0.8085
4                3.6250      0.6982   5.1918      0.0130


Catalizador = 2  subtracted from:

             Difference       SE of             Adjusted
Catalizador    of Means  Difference  T-Value     P-Value
3                0.3750      0.6982   0.5371      0.9462
4                3.3750      0.6982   4.8338      0.0175


Catalizador = 3  subtracted from:

             Difference       SE of             Adjusted
Catalizador    of Means  Difference  T-Value     P-Value
4                 3.000      0.6982    4.297      0.0281
```

Por lo que al terminar de analizar las comparaciones en los niveles de LOTE tenemos: 1=2, 1≠3, 1=4, 2≠3, 2=4 y 3≠4

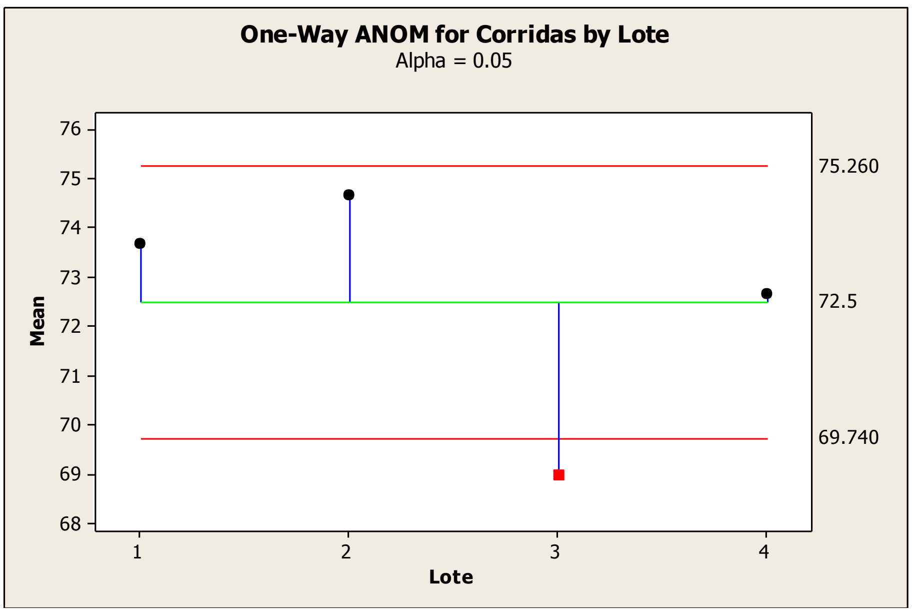
Por lo tanto, si los promedios de lotes son: 1=73.6, 2=74.6, 3=69 4=72.6

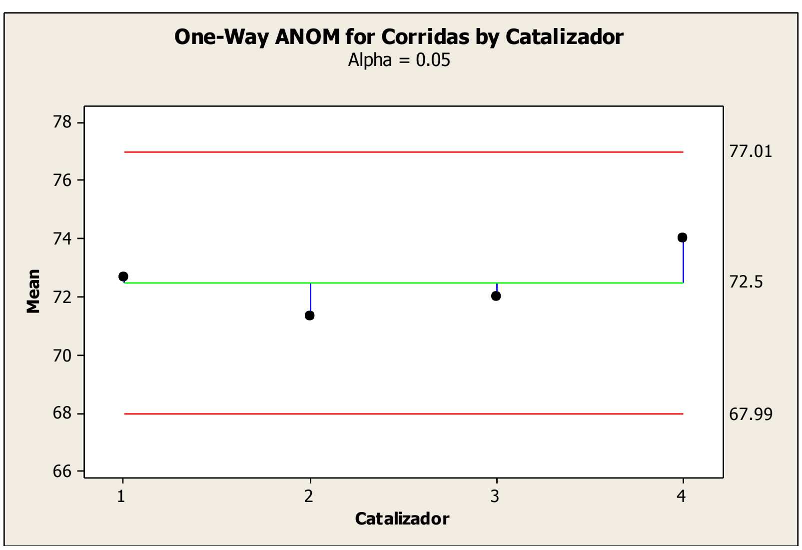
Entonces, si nos interesa los menores tiempos, **lote 3** da en promedio 69 segundos y es significativamente diferente de lote 1, 2 y 4.

Asimismo al analizar CATALIZADORES tenemos: 1=2, 1=3, 1≠4, 2=3, 2≠4 y 3≠4

Y si los promedios de Catalizadores en sus cuatro niveles son: 1=72.6, 2=71.3, 3=72 y 4=74
En catalizadores, el 2 da un tiempo de 71.3 seg. Pero es estadísticamente igual a 1 y 3, entonces, **tanto catalizador 1, 2 o 3 darán resultados similares**, no así el catalizador 4.
Finalmente se tiene evidencia que los menores tiempos serán al utilizar el lote 3 y los catalizadores 1,2 o 3. Sólo faltará revisar otras consideraciones como el precio, disponibilidad de cada catalizador entre otras cosas, para proceder con la decisión final.

Recuerde que mediante *Stat / Anova / Analysis of means* se obtienen los promedios de los cuatro niveles de lote y de catalizador, teniendo los siguientes resultados.





B) DESARROLLO

practique analizando los problemas que proporcione el profesor

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

- Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
- Douglas C. Montgomery (2009). Design and Analysis of Experiments
- Douglas C. Montgomery y George C. Runger, (2005). Probabilidad y Estadística aplicadas a la Ingeniería. Editorial Limusa Wiley, segunda edición.
- Seymour Lipschutz y John Schiller, (2000). Introducción a la Probabilidad y Estadística. Editorial Mc Graw Hill.
- Humberto Gutiérrez y Román De La Vara, (2009). Control Estadístico de Calidad y Seis Sigma. Editorial Mc Graw Hill, segunda edición.

6.- ANEXOS:

NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**DISEÑO FACTORIALES 2^k**	**PRÁCTICA NÚMERO**	**10**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

SOFTWARE REQUERIDO
MINITAB y EXCEL
OBSERVACIONES-COMENTARIOS

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:
Los diseños vistos hasta esta unidad, sólo contemplan un factor de tratamientos y el resto factores de bloque que tienen una importancia secundaria en la investigación.

En esta unidad se analizará los diseños factoriales, los cuales tienen por objetivo, estudiar el efecto de varios factores sobre una o varias respuestas, cuando todos los factores son de interés (no sólo uno). Y es que en muchas ocasiones se desea determinar en un experimento, una combinación de niveles de los factores, en la que el desempeño del proceso sea mejor.

Los factores pueden ser cuantitativos (temperatura, presión, velocidad, etc.) o cualitativos (tipo de máquina, operador, lote, etc.). Sin embargo, es necesario elegir al menos dos niveles de prueba de los factores,

A continuación se analizan experimentos con diseños factoriales por medio del empleo del software Minitab y Excel.

2.- OBJETIVO (COMPETENCIA):
Analizar experimentos por medio deL estudio del efecto de varios factores sobre una o varias respuestas, cuando todos los factores son de interés (no sólo uno) y decidir cual es la mejor opción utilizando el software Minitab.

3.- TEORIA:
Diseño factorial 2^k

El diseño factorial completo 2^2 (2 factores con dos niveles cada uno) es uno de los diseños de mayor impacto en la industria y la investigación, debido a su eficacia y versatilidad, y es que **es muy útil cuando el número de factores a estudiar está entre dos y cinco.** Es decir entre 4 y 32 tratamientos, cantidades que en la práctica son muy cómodas.

Nota Si el número de factores es mayor a 5, se debería usar un factorial fraccionado 2^{k-p}

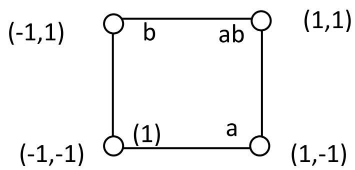
En el diseño 2^2, cada replica consiste de 2 x 2= 4 tratamientos que se puede denotar de diferentes maneras como en la siguiente tabla, donde A y B son dos factores.

	Factores		Factores		Factores		Factores		Factores		Factores		Notación de Yates
Tratamiento	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**	
1	bajo	bajo	A_1	B_1	A^-	B^-	-	-	0	0	-1	-1	(1)
2	alto	bajo	A_2	B_1	A^+	B^-	+	-	1	0	1	-1	a
3	bajo	alto	A_1	B_2	A^-	B^+	-	+	0	1	-1	1	b

4	alto	alto	A_2	B_2	A^+	B^+	+	+	1	1	1	1	ab

Representación geométrica.

El diseño factorial 2^2 se representa de manera geométrica por los vértices del cuadro que sigue. El área limitada por este cuadro se conoce como **región experimental** y, en principio, las conclusiones que se obtengan del experimento, sólo tienen validez sobre esta región.



4.- DESCRIPCIÓN

A) PROCEDIMIENTO Y DURACION DE LA PRÁCTICA:

1.- Encienda su computadora y abra el programa Minitab.

2.- Resuelva los siguientes problemas utilizando Excel y Minitab

Ejemplo 1 Supongamos que en un proceso de fermentación del tequila, interesa determinar si el rendimiento es influenciado por dos factores, los cuales se desean probar: el factor levadura, a dos niveles (L1 y L2) y el factor temperatura también a dos niveles (22°C y 30°C).

Si sólo interesaran estos dos factores, para estudiar la forma en que influyen sobre la variable de respuesta (rendimiento), entonces un **diseño factorial completo** requeriría correr aleatoriamente todas las posibles combinaciones que pueden formarse con los niveles de los factores (L1 y 22°C, L1 y 30°C, L2 y 22°C, por último L2 y 30°C). Así la **matriz de diseño** o **arreglo factorial** es el conjunto de **puntos experimentales o tratamientos** que pueden formarse considerando todas las posibles combinaciones de los niveles de los factores. Para el ejemplo, k=2 factores, y ambos factores a 2 niveles se forma un diseño factorial 2x2= 2^2, que consiste en cuatro tratamientos o puntos experimentales. Sin embargo, si se decide aumentar un nivel más el factor temperatura (digamos 26°), entonces se pueden construir 2x3 combinaciones y se tendría un diseño factorial 2x3, con 6 posibles **tratamientos**. A su vez si se decide correr el factor Levadura también a 3 niveles (L1, L2 y L3) entonces el diseño factorial sería 3X3 = 9 tratamientos, es decir un diseño 3^3.

En general, la familia de diseños factoriales 2k consiste en k factores, todos con 2 niveles, mientras que los diseños factoriales 3k consisten en k factores y todos ellos con 3 niveles.

Para obtener el número de corridas experimentales se multiplica el número de tratamientos por el número de replicas, donde cada replica se lleva a cabo cada vez que se corre el arreglo completo. Por ejemplo, del diseño factorial 2x3= 6 tratamientos, cada vez que se corren los 6 tratamientos, se está haciendo una replica.

Continuando con el ejemplo de la fermentación del tequila, se tienen dos factores a dos niveles: Tipo de levadura L1=1 y L2=2 y el factor Temperatura T1=22°C y T2=30°C, ambos factores y la respuesta del rendimiento del proceso de fermentación se muestran en la siguiente tabla.

Levadura (L)	Temperatura (T)	Rendimiento (Y)
L1= 1 (-1)	T1= 22° (-1)	28
L2= 2 (1)	T1= 22° (-1)	41
L1= 1 (-1)	T2= 30° (1)	63
L2= 2 (1)	T2= 30° (1)	45

Cada renglón es un tratamiento o punto del diseño factorial 2^2 resultante, y entre paréntesis se ha indicado cada nivel con códigos (1,-1) y que por practicidad se denominan nivel alto y bajo respectivamente, los cuales ayudan a identificar los niveles de los factores. Observe además que cada tratamiento se corrió una sola vez, es decir no hay replicas.

Efecto principal y efecto de interacción.

El efecto de un factor se define como el cambio observado en la variable de respuesta debido a un cambio de nivel de tal factor. A su vez, los **efectos principales** son los cambios en la media de la variable de respuesta que se deben a la acción individual de cada factor. En términos matemáticos, el efecto principal de un factor con dos niveles, es la diferencia entre la respuesta media observada cuando el factor estuvo en su primer nivel, y la respuesta media observada cuando el mismo factor estuvo en su segundo nivel. Entonces los efectos principales para Levadura y Temperatura son:

$$Efecto\ Levadura(L) = \frac{41+45}{2} - \frac{28+63}{2} = \mathbf{-2.5}$$

$$Efecto\ Temperatura\ (T) = \frac{63+45}{2} - \frac{28+41}{2} = \mathbf{19.5}$$

En términos absolutos, el efecto principal T, es mayor con respecto a L. Por otra parte, se dice que dos factores interactúan entre sí o **tienen un efecto de interacción** sobre la variable de respuesta, cuando el efecto de un factor depende del nivel en que se encuentra el otro. Por ejemplo los factores L y T interactuan si el efecto de L es muy diferente en cada nivel de T, o viceversa. Veamos

Efecto L (con T bajo) = 41-28= 13

Y cuando la temperatura es alta, el efecto de L es

Efecto L (con T alto) = 45-64= -18

Como estos dos efectos de L en función de T son muy diferentes, entonces es evidencia de que la elección más conveniente del nivel L depende del nivel en que está T, y viceversa. Es decir, eso es evidencia de que los factores L y T interactúan sobre Y. En la práctica, el cálculo del efecto L en cada nivel de T no se hace, y más bien se calcula el efecto global de la interacción de los dos factores, que es denotado por la multiplicación de ambos, es decir LT, y se calculan como la diferencia entre la respuesta media cuando ambos factores se encuentran en el mismo nivel: (-1, -1) (1, 1), y la respuesta media cuando ambos se encuentran en niveles opuestos: (-1, 1) (1, -1). Para nuestro ejemplo tenemos

$$LT = \frac{28+45}{2} - \frac{41+63}{2} = \mathbf{-15.5}$$

Los valores absolutos de los efectos principales y del efecto de interacción son una medida de importancia de su efecto sobre la variable de respuesta. Sin embargo, como se tienen estimaciones muestrales, para saber si los efectos son estadísticamente significativos (diferente de cero) se requiere el análisis de varianza (ANOVA)

Efectos con más de una replica.

Si el diseño factorial tuviera replicas, digamos n, los cálculos de los efectos utilizando la notación de yates sería:

$$Efecto\ A = \frac{1}{2n}[a + ab - b - (1)] = \frac{[a+ab]}{2n} - \frac{[b+(1)]}{2n}$$

Y el efecto de B sería

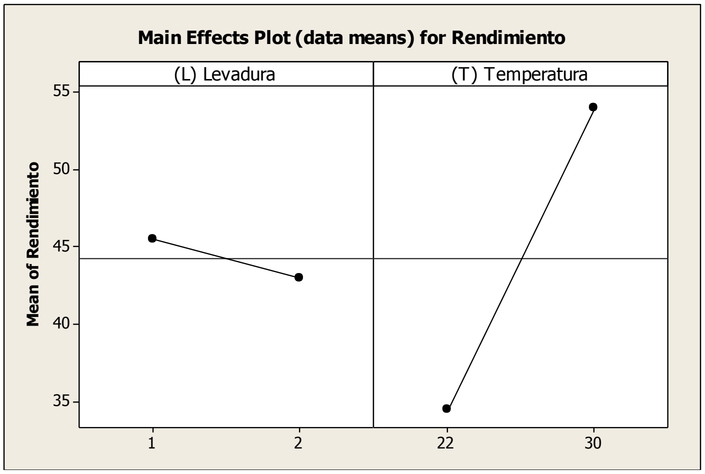
$$Efecto\ B = \frac{1}{2n}[b + ab - a - (1)] = \frac{[b+ab]}{2n} - \frac{[a+(1)]}{2n}$$

Por su parte el efecto de interacción sería

$$Efecto\ AB = \frac{1}{2n}[ab + (1) - a - b] = \frac{[ab-b]}{2n} - \frac{[a-(1)]}{2n}$$

Gráficas de efectos principales y de interacción.

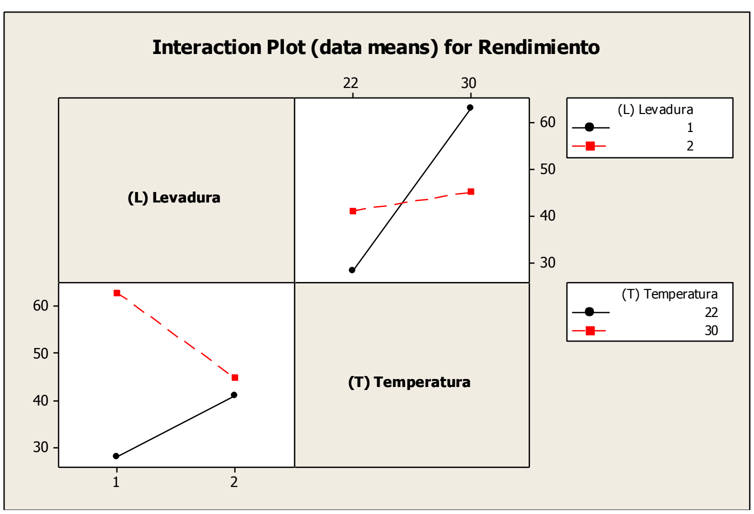
Para nuestro caso, podemos graficar los efectos principales (main effects) de los dos factores L y T siguiendo la secuencia en Minitab: *Stat/Anova / Main Effects Plot/* ingresar la respuesta (rendimiento) y los factores (Levadura y Temperatura) y después dar *ok.*



Observe que en el eje horizontal se ubican los niveles de cada factor y en el vertical se encuentra la media de la respuesta observada. A su vez se aprecia que el efecto del factor temperatura es mucho mayor que Levadura, y es que cuando la temperatura está en 22° el rendimiento es alrededor de 35 mientras que cuando sube a 30° el rendimiento se acerca a 55°, es decir el efecto principal es grande. En cambio, el efecto principal de Levadura se aprecia pequeño.

El efecto de interacción de los dos factores se representa con sus respectivos dos niveles en el eje horizontal y en dirección vertical de cada uno de estos niveles, se anota un punto que representa la respuesta media (promedios) en cada nivel del otro factor. Resulta que cuando existe interacción, las líneas obtenidas tienen una pendiente muy diferente. Si no hay interacción las líneas tienen pendientes similares, que son aproximadamente paralelas.

A su vez, podemos graficar en minitab también la interacción LT como sigue: *Stat/Anova / Interactions Plot*/ ingresar la respuesta (rendimiento) y los factores (Levadura y Temperatura) y después dar *ok*.



Observe la gráfica superior derecha donde Temperatura se representa en el eje horizontal y Levadura en el vertical. Claramente se aprecia que ambas líneas tiene pendientes diferentes, por lo que hay evidencia de que existe interacción entre T y L. Observe que si T se incrementa de 22° a 30°, cuando L esta en el nivel 1, la respuesta se incrementa significativamente, sin embargo si T se incrementa en el nivel 2 de la levadura, el aumento en Y es ligero.

Otra forma de ver la interacción es invirtiendo la posición en los ejes, tal como en la gráfica inferior izquierda, donde levadura representa el eje X y en Y Temperatura. Aquí también se sigue viendo que ambas líneas tienen pendientes diferentes (hay interacción). Observe que en esta gráfica si L se cambia de su nivel 1 a 2, cuando T es 22°, la respuesta Y también se incrementa; pero si T es 30°, la respuesta decrece considerablemente. En otras palabras, el factor L tiene un efecto positivo o negativo sobre Y, dependiendo del nivel de T.

Como se puede ver, ya sea con una u otra gráfica, la interacción está presente para estos dos factores, si se quisiera maximizar, minimizar o llevar a un valor objetivo la respuesta Y, no se debe mover al factor L sin tomar en cuenta en que nivel está T, y viceversa. Esta es la importancia de las interacciones.

Ventajas de los diseños factoriales.

1. Permite estudiar el efecto individual y de interacción de los distintos factores.
2. Son diseños que se pueden aumentar para formar diseños compuestos en caso de que se requiera una exploración más completa. Por ejemplo es útil aumentar el diseño si el comportamiento de la respuesta no es lineal.
3. Se pueden correr fracciones de diseños factoriales, las cuales son de gran utilidad en las primeras etapas de una investigación que involucra a muchos factores y que interesa descartar de manera económica los que no son importantes, antes de hacer un estudio más detallado con los factores que si son importantes.
4. Pueden utilizarse en combinación con diseños en bloques en situaciones en las que no puede correrse todo el diseño factorial bajo las mismas condiciones.
5. La interpretación y el cálculo de los efectos en los experimentos factoriales se puede hacer con aritmética elemental.

Ejemplo 2.

Se desea investigar los factores *tamaño de broca* (factor A) y *velocidad de broca* (factor B) que se cree influyen en la respuesta *vibración de taladro*. Para ello, se decide utilizar un diseño factorial 2^2 con 4 replicas, lo cual da un total de $2^2 \times 4 = 16$ corridas del proceso, las cuales se realizan en completo orden aleatorio. El tamaño de la broca se prueba en 1/16 y 1/8 de pulgada y la velocidad en 40 y 90 revoluciones por segundo, según la siguiente tabla:

Factor	Niveles		Unidad
A: Broca	1/16	1/8	Pulgadas
B: Velocidad	40	90	RPS

Posteriormente se escogió un diseño factorial 2^2, con el arreglo que se muestra en la siguiente tabla, junto con los 16 datos de las 16 corridas aleatorizadas. Cabe señalar que el experimento, además de considerar 4 repeticiones o replicas, se corrió en igualdad de circunstancias con respecto al resto de factores no estudiados. En la última columna se muestra el total por tratamientos utilizando la notación de Yates.

	Factores				Velocidad				Total en notación de Yates
Tratamiento	**A: Broca**	**B: Velocidad**	**A**	**B**					
1	1/16	40	-	-	18.2	18.9	12.9	14.4	(1)= 64.4
2	1/8	40	+	-	27.2	24	22.4	22.5	a= 96.1
3	1/16	90	-	+	15.9	14.5	15.1	14.2	b= 59.7
4	1/8	90	+	+	41	43.9	36.3	39.9	ab=161.1

Comenzamos por calcular los efectos de los factores así como de la interacción, para esto nos apoyamos en la notación de Yates, pero básicamente es la misma forma utilizada al inicio.

$$Efecto\ A = \frac{[a+ab]}{2n} - \frac{[b+(1)]}{2n} = \frac{[96.1+161.1]}{2x4} - \frac{[59.7+64.4]}{2x4} = \mathbf{16.64}$$

Y el efecto de B sería

$$Efecto\ B = \frac{[b+ab]}{2n} - \frac{[a+(1)]}{2n} = \frac{[59.7+161.1]}{2x4} - \frac{[96.1+64.4]}{2x4} = \mathbf{7.54}$$

Por su parte el efecto de interacción sería

$$Efecto\ AB = \frac{[ab-b]}{2n} - \frac{[a-(1)]}{2n} = \frac{[161.1+64.4]}{2x4} - \frac{[96.1+59.7]}{2x4} = \mathbf{8.71}$$

Se observa claramente que el efecto del factor A (Broca) es prácticamente el doble que los otros dos. Aunque Broca parece ser que es significativo dado el tamaño de su efecto, probaremos si alguno de los 3 es estadísticamente significativo mediante una tabla Anova, tal como se muestra a continuación.

Fuente de variación	Suma de cuadrados (SC)	grados de libertad (gl)	Cuadrados medios (CM)	Estadístico F_0	Valor P
Tratamiento A	$SC_A = \sum_{i=1}^{a} \frac{Y_{i..}^2}{bn} - \frac{Y_{...}^2}{N}$	$a-1$	$CM_A = \frac{SC_A}{a-1}$	$\frac{CM_A}{CM_{Error}}$	$P(F > F_0)$
Tratamiento B	$SC_B = \sum_{j=1}^{b} \frac{Y_{.j.}^2}{an} - \frac{Y_{...}^2}{N}$	$b-1$	$CM_B = \frac{SC_B}{b-1}$	$\frac{CM_B}{CM_{Error}}$	$P(F > F_0)$
Interacción AB	$SC_{AB} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{Y_{ij.}^2}{n} - \frac{Y_{...}^2}{N} - SC_A - SC_B$	$(a-1)(b-1)$	$CM_{AB} = \frac{SC_{AB}}{(a-1)(b-1)}$	$\frac{CM_{AB}}{CM_{Error}}$	$P(F > F_0)$
Error	$SC_{Error} = SC_T - SC_A - SC_B - SC_{AB}$	$ab(n-1)$	$CM_{Error} = \frac{SC_{Error}}{ab(n-1)}$		
Total	$SC_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} Y_{ijk}^2 - \frac{Y_{...}^2}{N}$	$abn-1$			

ANOVA DISEÑO FACTORIAL DE DOS FACTORES

a= número de niveles del factor A
n= número de replicas de cada tratamiento
b= número de niveles del factor B
N= número total de corridas = a x b x n

Nota: esta tabla de Anova funciona para cualquier diseño factorial con dos factores (A y B) y estos inclusive con diferentes niveles (por ejemplo 3 x 4, 2 x 2, etc.), aunque en la práctica el más utilizado es el 2^2, es decir un diseño 2 x 2, donde a=2 y b=2

Los resultados de la tabla Anova serían:

Fuente de variación	Suma de cuadrados (SC)	grados de libertad (gl)	Cuadrados medios (CM)	Estadístico F_0	Valor P
Broca (A)	$SC_A = \frac{(64.4+59.7)^2+(96.1+161.1)^2}{2x4} - \frac{381.3^2}{2x2x4} = \mathbf{1107.2}$	1	$CM_A = 1107.22$	185.25	0.0000
Velocidad (B)	$SC_B = \frac{(64.4+96.1)^2+(59.7+161.1)^2}{2x4} - \frac{381.3^2}{2x2x4} = \mathbf{227.25}$	1	$CM_B = 227.25$	38.02	0.0000
Interacción AB	$SC_{AB} = \mathbf{303.63}$	1	$CM_{AB} = 303.63$	50.80	0.0000
Error	$SC_{Error} = 1709.83 - 1107.2 - 227.25 - 303.63 = \mathbf{71.73}$	12			
Total	$SC_T = 1709.83$	15			

Interpretación:
Como el objetivo es minimizar la vibración del taladro, observe que tanto el efecto Broca (A) como velocidad (B), así como la interacción AB tienen un efecto significativo sobre la vibración, ya que sus pvalue fueron mucho menores que el nivel de significancia común ($pvalue < 0.05$), por lo que los dos factores y la interacción se debe tomar en cuenta para disminuir la vibración. En la práctica, las interacciones tienen prioridad sobre los efectos para controlarse, por lo que se le debe poner más atención.

Para verificar mediante Minitab, los datos se deben ordenar en columnas y especificando a Minitab cuales son niveles bajos y altos (por ejemplo 1=bajo, 2=alto) como sigue:

Broca (A)	Velocidad (B)	Vibración
1	1	18.2
2	1	27.2
1	2	15.9
2	2	41
1	1	18.9
2	1	24
1	2	14.5
2	2	43.9
1	1	12.9
2	1	22.4
1	2	15.1

2	2	36.3
1	1	14.4
2	1	22.5
1	2	14.2
2	2	39.9

Observe que son los mismos datos pero ordenados en columnas.

Ya en Minitab ingresamos las columnas con los datos y utilizamos la secuencia de comandos: *Stat / DOE / Factorial / Analize Factorial Designs,* posteriormente ingresamos en *factors*, las columnas de broca y velocidad, dado que es un 2^2 activamos la celda de *General Full Factorial*, y damos *ok*. Luego ingresamos la columna de respuesta y damos **ok**.

La hoja de sesión arrojada es la siguiente

General Linear Model: Vibración versus Broca (A), Velocidad (B)

Factor	Type	Levels	Values
Broca (A)	fixed	2	1, 2
Velocidad (B)	fixed	2	1, 2

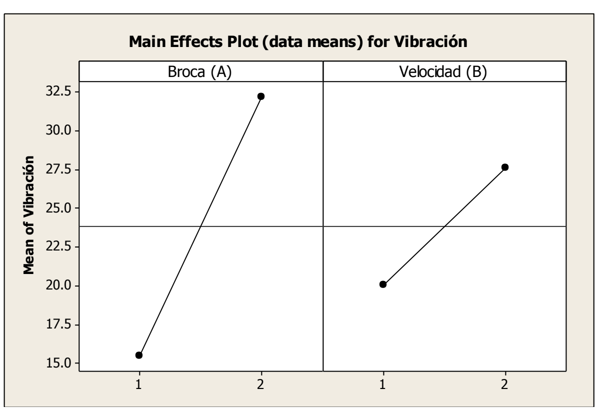
Analysis of Variance for Vibración, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Broca (A)	1	1107.23	1107.23	1107.23	185.25	0.000
Velocidad (B)	1	227.26	227.26	227.26	38.02	0.000
Broca (A)*Velocidad (B)	1	303.63	303.63	303.63	50.80	0.000
Error	12	71.72	71.72	5.98		
Total	15	1709.83				

S = 2.44476   R-Sq = 95.81%   R-Sq(adj) = 94.76%

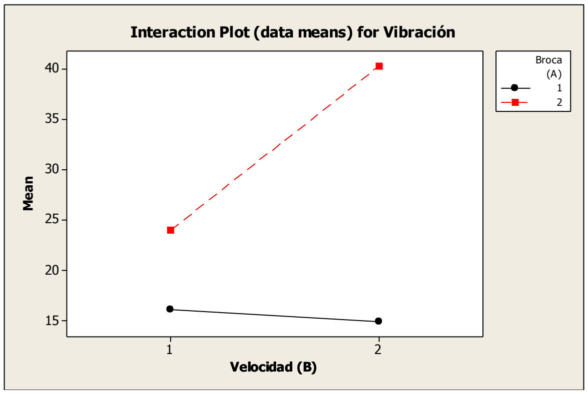
Esto concuerda con la anterior tabla de anova. Observe que tanto Broca, velocidad, como la interacción entre ambos, tienen un p value menor de 0.05, por lo cuál se concluye que los tres son significativos, es decir, los dos factores y la interacción influyen en la vibración del taladro

Dado que se busca minimizar la variable de respuesta (vibración del taladro), hay que definir cual combinación de niveles da un mejor resultado. Por lo que nos podemos apoyar en la gráfica de efectos y en la de interacción que genera Minitab siguiendo la secuencia *Stat / ANOVA / Main Effects Plot*, ingresamos la respuesta y posteriormente los dos factores y damos *ok*.



Observe que ambos factores tienen un efecto similar. En el caso de Velocidad en el nivel bajo (1) es cuando produce mejores resultados. A su vez broca en nivel bajo (1), también produce valores bajos de vibración.

A su vez, la gráfica de interacción se genera Minitab siguiendo la secuencia *Stat / ANOVA / Interaction plot*, ingresamos la respuesta y posteriormente los dos factores y damos *ok*.



El hecho de que la interacción fuera significativa ($< \alpha$), se aprecia visualmente cuando las líneas tienen diferentes pendientes. Como lo que interesa es minimizar la variable de respuesta, se observa que los mejores resultados se obtienen con la broca en nivel bajo, a pesar del cambio de la velocidad, el cambio en la respuesta es muy pequeño. Note que a mayor velocidad con la broca en nivel alto, si hay una tendencia a obtener peores resultados, además se ve que cuando velocidad es baja, el efecto de broca es menor (los puntos de velocidad en nivel bajo están más cerca). Por lo tanto, las condiciones de operación o dicho de otra forma, los tratamientos que dan mejor resultados la broca está en nivel bajo y cuando velocidad están en nivel alto.

B) DESARROLLO

Practique analizando los problemas que proporcione el profesor

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados.

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

- Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
- Douglas C. Montgomery (2009). Design and Analysis of Experiments
- Humberto Gutiérrez y Román De La Vara, (2009). Control Estadístico de Calidad y Seis Sigma. Editorial Mc Graw Hill, segunda edición.

6.- ANEXOS:

NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**DISEÑO FACTORIAL CON DOS FACTORES PERO A DIFERENTES NIVELES**	**PRÁCTICA NÚMERO**	**11**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

SOFTWARE REQUERIDO
MINITAB y EXCEL
OBSERVACIONES-COMENTARIOS

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

En ciertas ocasiones se necesita probar dos factores, pero no necesariamente a dos niveles, es decir, si tenemos los factores A y B y sus niveles $a\ y\ b$, en ciertas ocasiones puede ser que $(a, b \geq 2\ niveles\ de\ prueba)$, y esto depende del fenómeno a estudiar.

A continuación se analizan experimentos con diseños factoriales con dos factores pero diferentes niveles por medio del empleo del software Minitab y Excel.

2.- OBJETIVO (COMPETENCIA):

Analizar experimentos por medio de estudio del efecto de varios factores sobre una respuesta pero condiferentes niveles y decidir cual es la mejor opción utilizando el software Minitab.

3.- TEORIA:

Suponga que interesa estudiar el efecto del <u>factor A: profundidad de corte y B: velocidad de alimentación</u>, sobre la respuesta Y: acabado de metal, la cual se mide en gramos.

Los investigadores desean probar a 3 niveles el factor velocidad de alimentación (0.20, 0.25 y 0.30) y a 4 niveles el factor profundidad (0.15, 0.18, 0.21y 0.24). A su vez, desean correr 3 replicas de cada tratamiento, por lo que aleatorizan 36 pruebas y se obtienen los siguientes datos.

		Factor B (Velocidad)					
		0.2		0.25		0.3	
Factor A (Profundidad)	0.15	74 64 60	198	92 86 88	266	99 98 102	299
	0.18	79 68 73	220	98 104 88	290	104 99 95	298
	0.21	82 88 92	262	99 108 95	302	108 110 99	317
	0.24	99 104 96	299	104 110 99	313	114 111 107	332

Desarrollando la tabla de anova para diseño factorial de dos factores, tenemos el siguiente resultado:

Fuente de variación	Suma de cuadrados (SC)	grados de libertad (gl)	Cuadrados medios (CM)	Estadístico F_0	Valor P
Profundidad (A)	$SC_A = \frac{763^2 + 808^2 + 881^2 + 944^2}{3x3} - \frac{3396^2}{4x3x3} = \mathbf{2125.1}$	3	$CM_A = 708.37$	24.66	0.0000
Velocidad (B)	$SC_B = \frac{979^2 + 1171^2 + 1246^2}{4x3} - \frac{3396^2}{4x3x3} = \mathbf{3160.5}$	2	$CM_B = 1580.25$	55.02	0.0000
Interacción AB	$SC_{AB} = \frac{198^2 + 220^2 + \cdots + 332^2}{3} - \frac{3396^2}{4x3x3} = \mathbf{557.07}$	6	$CM_{AB} = 92.84$	3.23	0.0180
Error	$SC_{Error} = 6532 - 2125.1 - 3160.5 - 557.07 = \mathbf{689.33}$	24			
Total	$SC_T = 6532$	35			

Para verificar mediante Minitab, los datos se deben ordenar como sigue:

profundidad	Velocidad	Resultado
0.15	0.2	74
0.15	0.2	64
0.15	0.2	60
0.15	0.25	92
0.15	0.25	86
0.15	0.25	88
0.15	0.3	99
0.15	0.3	98
0.15	0.3	102
0.18	0.2	79
0.18	0.2	68
0.18	0.2	73
0.18	0.25	98
0.18	0.25	104
0.18	0.25	88
0.18	0.3	104
0.18	0.3	99
0.18	0.3	95
0.21	0.2	82

0.21	0.2	88
0.21	0.2	92
0.21	0.25	99
0.21	0.25	108
0.21	0.25	95
0.21	0.3	108
0.21	0.3	110
0.21	0.3	99
0.24	0.2	99
0.24	0.2	104
0.24	0.2	96
0.24	0.25	104
0.24	0.25	110
0.24	0.25	99
0.24	0.3	114
0.24	0.3	111
0.24	0.3	107

Observe que son los mismos datos pero ordenados en columnas.

En Minitab ingresamos las columnas con los datos y utilizamos la secuencia de comandos: *Stat / Doe / Factorial / Analize Factorial Designs,* posteriormente ingresamos en *factors*, las columnas de profundidad y velocidad, dado que no es un 2^2 activamos la celda de *General Full Factorial* y damos *ok*. Luego ingresamos la columna de respuesta y damos **ok**.

La hoja de sesión arrojada es la siguiente

Results for: profundidad vs velocidad factorial 3x4.MTW

General Linear Model: Resultado versus profundidad, Velocidad

Factor	Type	Levels	Values
profundidad	fixed	4	0.15, 0.18, 0.21, 0.24
Velocidad	fixed	3	0.20, 0.25, 0.30

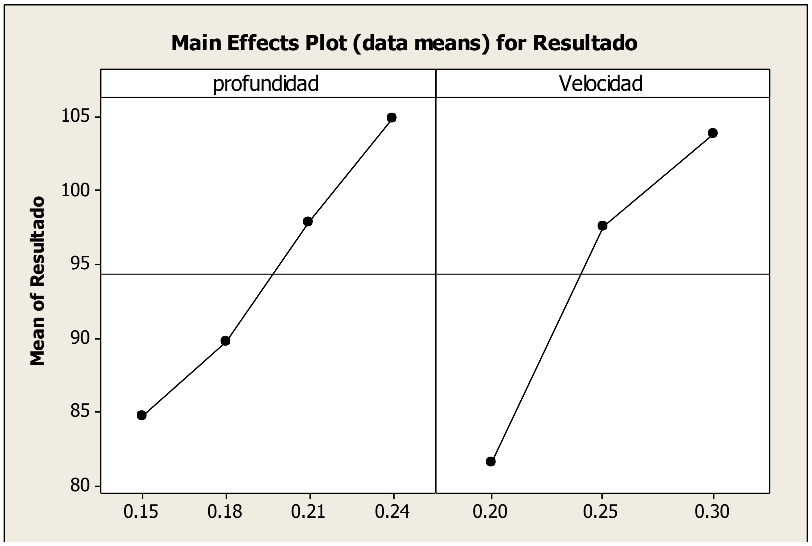
Analysis of Variance for Resultado, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
profundidad	3	2125.11	2125.11	708.37	24.66	0.000
Velocidad	2	3160.50	3160.50	1580.25	55.02	0.000
profundidad*Velocidad	6	557.06	557.06	92.84	3.23	0.018
Error	24	689.33	689.33	28.72		
Total	35	6532.00				

S = 5.35931   R-Sq = 89.45%   R-Sq(adj) = 84.61%

Tanto Profundidad, velocidad, como la interacción entre ambos, tienen un p value menor de 0.05, por lo cuál se concluye que los tres son significativos, es decir, los dos factores y la interacción influyen en el acabado del metal

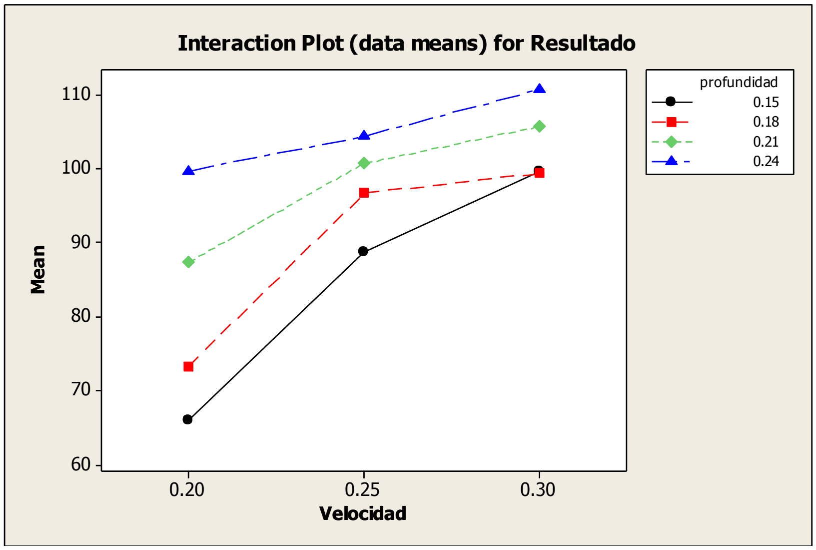
Dado que se busca minimizar la variable de respuesta (peso en gramos del metal) hay que definir cual combinación de niveles da un mejor resultado. Por lo que nos podemos apoyar en la gráfica de efectos y en la de interacción que genera Minitab siguiendo la secuencia *Stat / ANOVA / Main Effects Plot*, ingresamos la respuesta y posteriormente los dos factores y damos *ok*.



Observe que ambos factores tienen un efecto similar. En el caso de Velocidad en el nivel bajo (0.20) es cuando produce mejores resultados, alrededor de 83 gramos. A su vez profundidad en nivel bajo (0.15), también produce valores cercanos a 84.

A su vez, la gráfica de interacción se genera en Minitab siguiendo la secuencia *Stat / ANOVA / Interaction plot*, ingresamos la respuesta y posteriormente los dos factores y damos *ok*.

Observe en la gráfica de interacción que hay 4 líneas, esto porque profundidad tiene 4 niveles. El hecho de que la interacción fuera significativa ($< \alpha$), se aprecia visualmente cuando las líneas tienen diferentes pendientes. Como lo que interesa es minimizar la variable de respuesta, se observa que a mayor velocidad y profundidad hay una tendencia a obtener perores resultados, además se ve que cuando velocidad es alta, el efecto de profundidad es menor (los puntos están más cerca). Por lo tanto, las condiciones de operación o dicho de otra forma, los tratamientos que dan mejor resultados es cuando velocidad y profundidad están en niveles bajos (velocidad 0.20 y profundidad 0.15)



B) DESARROLLO

practique analizando los problemas que proporcione el profesor

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

- Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
- Douglas C. Montgomery (2009). Design and Analysis of Experiments
- Douglas C. Montgomery y George C. Runger, (2005). Probabilidad y Estadística aplicadas a la Ingeniería. Editorial Limusa Wiley, segunda edición.
- Seymour Lipschutz y John Schiller, (2000). Introducción a la Probabilidad y Estadística. Editorial Mc Graw Hill.
- Humberto Gutiérrez y Román De La Vara, (2009). Control Estadístico de Calidad y Seis Sigma. Editorial Mc Graw Hill, segunda edición.

6.- ANEXOS:

NOMBRE DE LA MATERIA	**DISEÑO DE EXPERIMENTOS**	**CLAVE**	**9020**
NOMBRE DE LA PRÁCTICA	**CALCULO DE CONTRASTE Y MODELO DE REGRESION**	**PRÁCTICA NÚMERO**	**12**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

SOFTWARE REQUERIDO
MINITAB y EXCEL
OBSERVACIONES-COMENTARIOS

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

Hasta ahora, hemos utilizado la tabla Anova (y sus respectivas sumas de cuadrados SC) para determinar si un efecto de un factor es significativo, es decir, si el efecto de un factor es diferente de cero. Sin embargo, la suma de cuadrados de esta tabla se puede obtener por otros medios, por ejemplo por medio de **contrastes**, los cuales están ligados a los efectos estimados ya vistos. A su vez, con la finalidad de predecir el valor de la variable de respuesta cuando los factores estudiados tienen diferentes valores, se puede generar un modelo de regresión. En este sentido a continuación se analizan los procedimientos a seguir para la obtención tanto de los contrastes como de un modelo de regresión por medio del empleo del software Minitab y Excel.

2.- OBJETIVO (COMPETENCIA):

Analizar experimentos por medio de estudio de los contrastes y predecir los valores que pueden tener una variable de respuesta por medio de regresión lineal con la finalidad de decidir cual es la mejor opción utilizando el software Minitab.

3.- TEORIA:

3.1 Definición de contraste

Hasta ahora, hemos utilizado la tabla Anova (y sus respectivas sumas de cuadrados SC) para determinar si un efecto de un factor es significativo, es decir, si el efecto de un factor es diferente de cero. Sin embargo, la suma de cuadrados de esta tabla se puede obtener por otros medios, por ejemplo por medio de **contrastes**, los cuales están ligados a los efectos estimados ya vistos. Los **contrastes** son combinaciones lineales donde los coeficientes suman cero, un ejemplo de contraste es 1+1-1-1=0

Método para calcular contrastes.

Para encontrar los contrastes de los efectos, es común utilizar la tabla de signos, la cual se construye a partir de la matriz de diseño, multiplicando las columnas que intervienen en la interacción que se quiere calcular. Por ejemplo para la interacción AB, se multiplica la columna de signos de A por la columna de signos de B y el resultado son los signos de AB.

A	B	AB	Yates
-	-	+	(1)
+	-	-	a
-	+	-	b
+	+	+	ab

Una vez que se tiene la columna de signos de A, B como de AB, **el contraste de cada efecto** resulta de multiplicar su columna de signos por la columna de los datos de la notación de Yates (la notación de Yates representa la suma de las observaciones en cada tratamiento), por ejemplo, el contraste de AB se obtiene de multiplicar +(1)-a-b+ab, es decir:

$$contraste\ de\ AB = ab + (1) - a - b$$

El contrastes de A se obtiene de multiplicar los signos de A por la columna de Yates.

$$Contraste\ A = [a + ab - b - (1)]$$

Y así, B se obtiene de multiplicar los signos de B por la columna de Yates

$$Contraste\ B = [b + ab - a - (1)]$$

De hecho, observe que los contrastes definen a los efectos que ya hemos vistos, sólo se divide el contraste entre $2n$ y se obtienen los efectos de los factores (principales y de interacción), descritos anteriormente:

$$Efecto\ A = \frac{1}{2n}[a + ab - b - (1)] = \frac{[a + ab]}{2n} - \frac{[b + (1)]}{2n}$$

$$Efecto\ B = \frac{1}{2n}[b + ab - a - (1)] = \frac{[b + ab]}{2n} - \frac{[a + (1)]}{2n}$$

$$Efecto\ AB = \frac{1}{2n}[ab + (1) - a - b] = \frac{[ab - b]}{2n} - \frac{[a - (1)]}{2n}$$

Observe que lo que está en rojo, es en realidad los contrastes divididos entre 2n, y a su vez se obtienen los efectos.

Cálculo de las sumas de cuadrados.

Para obtener las **sumas de cuadrados** para cada efecto, se eleva al cuadrado, tanto el numerador como el denominador.

$$SC_A = \frac{[a + ab - b - (1)]^2}{(2n)^2}$$

$$SC_B = \frac{[b + ab - a - (1)]^2}{(2n)^2}$$

$$SC_{AB} = \frac{[ab + (1) - a - b]^2}{(2n)^2}$$

Donde cada suma de cuadrados (SC) tiene solamente un grado de libertad, debido a que cada factor tiene 2 niveles (2-1=1), y la **suma de cuadrados totales** se calcula de la siguiente manera:

$$SC_{Total} = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{l=1}^{n} Y_{ijl}^2 - \frac{Y_{...}^2}{(2n)^2}$$

Con esta información se puede obtener la tabla de Anova como la siguiente:

Donde: a= número de niveles del factor A

b= número de niveles del factor B
n= número de replicas de cada tratamiento
N= número total de corridas = a x b x n

Fuente de variación	**Suma de cuadrados (SC)**	**grados de libertad (gl)**	**Cuadrados medios (CM)**	**Estadístico F_0**	**Valor P**
Tratamiento A	$SC_A = \frac{[\boldsymbol{a} + \boldsymbol{ab} - \boldsymbol{b} - (1)]^2}{(2n)^2}$	$a-1$	$CM_A = \frac{SC_A}{a-1}$	$\frac{CM_A}{CM_{Error}}$	$P(F > F_0)$
Tratamiento B	$SC_B = \frac{[\boldsymbol{b} + \boldsymbol{ab} - \boldsymbol{a} - (1)]^2}{(2n)^2}$	$b-1$	$CM_B = \frac{SC_B}{b-1}$	$\frac{CM_B}{CM_{Error}}$	$P(F > F_0)$
Interacción AB	$SC_{AB} = \frac{[\boldsymbol{ab} + (1) - \boldsymbol{a} - \boldsymbol{b}]^2}{(2n)^2}$	$(a-1)(b-1)$	$CM_{AB} = \frac{SC_{AB}}{(a-1)(b-1)}$	$\frac{CM_{AB}}{CM_{Error}}$	$P(F > F_0)$
Error	$SC_{Error} = SC_T - SC_A - SC_B - SC_{AB}$	$ab(n-1)$	$CM_{Error} = \frac{SC_{Error}}{ab(n-1)}$		
Total	$SC_{Total} = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{l=1}^{n} Y_{ijl}^2 - \frac{Y_{...}^2}{(2n)^2}$	$abn-1$			

Nota 1: esta tabla de Anova funciona para cualquier diseño factorial con dos factores (A y B) y estos inclusive con diferentes niveles (por ejemplo 3 x 4, 2 x 2, etc.), aunque en la práctica el más utilizado es el 2^2, es decir un diseño 2 x 2, donde a=2 y b=2

Nota 2: las vocales $\boldsymbol{a}$, $\boldsymbol{b}$ y $\boldsymbol{ab}$, solamente se refieren a las sumas o totales de la notación de YATES, de hecho lo único que cambia en esta tabla de Anova es la forma de encontrar la suma de cuadrados (SC_A, SC_B y SC_{AB})

3.2 Modelo de regresión y predicción de valores.

Con la finalidad de predecir el valor de la variable de respuesta cuando los factores estudiados tienen diferentes valores, se puede generar un modelo de regresión del tipo

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_1 x_2$$

Donde $\hat{Y}$ es la respuesta estimada cuando se fijan ciertos valores para x_1 y x_2. A su vez $\boldsymbol{x_1}$ **es el factor A** y $\boldsymbol{x_2}$ **es el factor** B que se quiere analizar. En el caso de los diseños 2^k, los coeficientes del modelo ($\hat{\beta}_1, \hat{\beta}_2$ y $\hat{\beta}_3$), se obtienen al dividir entre 2, los efectos estimados que resultaron significativos. Continuando con el ejemplo del taladro donde se tienen dos factores (A=Broca, B=Velocidad y AB=Interacción broca-velocidad) tenemos

$$\hat{\beta}_1 = \frac{Efecto\ A}{2} = \frac{16.64}{2} = 8.32$$

$$\hat{\beta}_2 = \frac{Efecto\ B}{2} = \frac{7.54}{2} = 3.77$$

$$\hat{\beta}_3 = \frac{Efecto\ AB}{2} = \frac{8.71}{2} = 4.35$$

A su vez, $\hat{\beta}_0 = \bar{Y}_{\ldots}$ es la media global de todos los datos, y para el ejemplo del taladro, representa la vibración predicha, en el centro de la región experimental ($x_1 = 0,\ x_2 = 0$), es decir, cuando el Factor A y B valen cero (están en el origen). Con esto podemos generar el modelo siguiente

$$\hat{Y} = 23.83 + 8.32 x_1 + 3.77 x_2 + 4.35 x_1 x_2$$

Por lo tanto, si se desea estimar el valor de vibración (Y), cuando $\boldsymbol{x_1}$ **(factor A) =-1** y $\boldsymbol{x_2}$ **(factor B)=1**, es decir el tratamiento A- y B+, se sustituye -1 en z x_1 y 1 en x_2 teniendo

$$\hat{Y} = 23.83 + 8.32(-1) + 3.77(1) + 4.35(-1)(1) = 14.92$$

En otras palabras, si se corriera en el nivel bajo A (broca 1/16) y el nivel alto de B(Velocidad 90 rps), se esperaría tener en promedio 14.92 unidades de vibración. Así pues, podemos sustituir todos los tratamientos analizados y ver cuál sería el valor estimado de Y.

En un sentido similar, **un Residual** (e), es la diferencia entre el valor observado en cierto tratamiento y el valor estimado o ajustado por el modelo de regresión para tal tratamiento $e_{ijk} = Y_{ijk} - \hat{Y}_{ijk}$

Los residuos permiten evaluar la calidad del modelo, ya que un residuo grande denotaría que el modelo estuvo alejado de lo que es el valor observado (original), por tanto, residuales pequeños significaría que el modelo ajusto bien a los datos observados y a su vez la calidad del modelo es buena.

Para ejemplificar un residual, revisemos los datos observados al correr A- y B+, fueron: 15.9, 14.5, 15.1 y 14.2

Por lo tanto el residual del primer dato en A- y B+ es $e = 15.9 - 14.92 = 0.975$

El resto de los valores estimados (ajustados) para el modelo y sus respectivos residuales son:

Broca (A)	Velocidad (B)	Vibración	Ajustado	Residual
-1	-1	18.2	16.09	2.11
1	-1	27.2	24.03	3.17
-1	1	15.9	14.93	0.97
1	1	41	40.27	0.73
-1	-1	18.9	16.09	2.81
1	-1	24	24.03	-0.03
-1	1	14.5	14.93	-0.43
1	1	43.9	40.27	3.63
-1	-1	12.9	16.09	-3.19
1	-1	22.4	24.03	-1.63
-1	1	15.1	14.93	0.17
1	1	36.3	40.27	-3.97
-1	-1	14.4	16.09	-1.69
1	-1	22.5	24.03	-1.53
-1	1	14.2	14.93	-0.73
1	1	39.9	40.27	-0.37

Calidad de un modelo de regresión.

Para medir la calidad global del modelo de regresión, se utiliza el **coeficiente de determinación R^2** y el **coeficiente de determinación ajustado R^2 aj**, los cuales se obtienen a partir de la tabla de Anova

$$R^2 = \left(\frac{SC_T - SC_{Error}}{SC_T}\right) x100 = \left(\frac{SC_{Modelo}}{SC_T}\right) x100$$

$$R^2_{aj} = \left(\frac{CM_T - CM_{Error}}{CM_T}\right) x100$$

Observe que ambos coeficientes comparan la variabilidad explicada por el modelo frente a la variación total y pueden tomar valores de 0.0 a 100, por lo que **R^2** y **R^2 aj** cuantifican el porcentaje de variabilidad presente en los datos y que es explicado por el modelo. Por tal motivo, es deseable valores cercanos a 100, ya eso significa que el modelo es bueno para explicar la variabilidad de la respuesta Y. en general se recomiendan coeficientes ≥ 70.

Cuando hay muchos factores, se prefiere el estadístico R^2 aj en lugar de R^2, ya que este último se incrementa de manera artificial con cada término que se agrega al modelo, aunque sea un término que no contribuya en mucho a la explicación de la respuesta (no sea significativo). Para nuestro ejemplo del taladro tenemos

$$R^2 = \left(\frac{1709.83 - 71.73}{1709.83}\right) x100 = 95.8$$

$$R_{aj}^2 = \left(\frac{\frac{1709.83}{15} - 5.98}{\frac{1709.83}{15}} \right) x100 = 94.76$$

Entonces, R_{aj}^2 indica que el modelo $\hat{Y} = 23.83 + 8.32x_1 + 3.77x_2 + 4.35x_1x_2$ explica hasta un 94.76% de la variabilidad de la vibración observada en el experimento. En otras palabras, los factores *Tipo de broca, velocidad de taladro y la interacción* de ambos, son responsables o explican un alto porcentaje de la variabilidad observada de la respuesta (vibración).

B) DESARROLLO

practique analizando los problemas que proporcione el profesor

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

- Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
- Douglas C. Montgomery (2009). Design and Analysis of Experiments
- Douglas C. Montgomery y George C. Runger, (2005). Probabilidad y Estadística aplicadas a la Ingeniería. Editorial Limusa Wiley, segunda edición.
- Seymour Lipschutz y John Schiller, (2000). Introducción a la Probabilidad y Estadística. Editorial Mc Graw Hill.
- Humberto Gutiérrez y Román De La Vara, (2009). Control Estadístico de Calidad y Seis Sigma. Editorial Mc Graw Hill, segunda edición.

6.- ANEXOS:

NOMBRE DE LA PRÁCTICA	**DISEÑO FACTORIAL 2^3**	**PRÁCTICA NÚMERO**	**13**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

SOFTWARE REQUERIDO
MINITAB y EXCEL

OBSERVACIONES-COMENTARIOS

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

En la vida real hay muchos fenómenos que interesa estudiar con tres factores o más. En la presente práctica se analizará un experimento donde intervienen tres factores y cada uno de ellos a dos niveles por medio del empleo del software Minitab y Excel.

2.- OBJETIVO (COMPETENCIA):

Analizar experimentos con tres factores a dos niveles y predecir el comportamiento de una variable de respuesta, con la finalidad de decidir cual es la mejor opción utilizando el software Minitab.

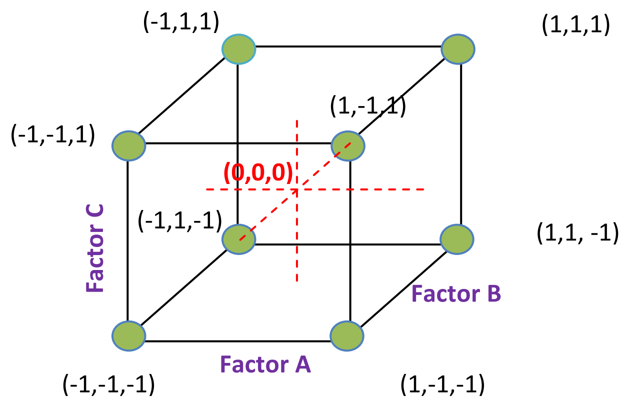
3.- TEORIA:

El diseño factorial completo 2^3 (3 factores con 2 niveles cada uno) consta de 2^3=2*2*2= 8 tratamientos diferentes, los cuales pueden identificarse con las mismas notaciones del 2^2 como sigue:

	Factores			Factores			Factores			Notación de Yates
Tratamiento	**A**	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**	
1	bajo	bajo	bajo	-	-	-	-1	-1	-1	(1)
2	alto	bajo	bajo	+	-	-	1	-1	-1	a
3	bajo	alto	bajo	-	+	-	-1	1	-1	b
4	alto	alto	bajo	+	+	-	1	1	-1	ab
5	bajo	bajo	alto	-	-	+	-1	-1	1	c
6	alto	bajo	alto	+	-	+	1	-1	1	ac
7	bajo	alto	alto	-	+	+	-1	1	1	bc
8	alto	alto	alto	+	+	+	1	1	1	abc

Representación geométrica.

El diseño factorial 2^3 se representa de manera geométrica por los 8 vértices (tratamientos) del cubo regular, centrado en el origen (0,0,0). El área limitada por este cubo se conoce como **región experimental** y, en principio, las conclusiones que se obtengan del experimento, sólo tienen validez sobre esta región.



Con este diseño se pueden estudiar los 2^3-1=7 efectos: 3 efectos principales (A, B y C), 3 interacciones dobles (AB, AC y BC) y una interacción triple (ABC). Por lo general el interés se enfoca en estudiar los efectos principales y las interacciones dobles, y no tanto en el triple, ya que normalmente su valor es pequeño. Habrá pues que verificar que la interacción triple es pequeña, para así ignorarla.

A su vez de la tabla de signos completa para un diseño 2^3, muestra los signos para las interacciones dobles y la triple, observe que estos signos son resultado de multiplicar signos de las columnas de factores, por ejemplo el signo de la interacción ABC es – en el renglón de (1), y esto se da al multiplicar (-)(-)(-).

Tratamiento	Total	**A**	**B**	**C**	**A B**	**AC**	**BC**	**ABC**
1	(1)	-	-	-	+	+	+	-
2	a	+	-	-	-	-	+	+
3	b	-	+	-	-	+	-	+
4	ab	+	+	-	+	-	-	-
5	c	-	-	+	+	-	-	+
6	ac	+	-	+	-	+	-	-
7	bc	-	+	+	-	-	+	-
8	abc	+	+	+	+	+	+	+

A su vez, los contrastes de cada factor se obtienen como ya hemos visto, al multiplicar la columna de totales (Yates), por la columna de signos de cada factor. Por ejemplo el contraste de A es

$$Contraste\ A = [a + ab + ac + abc - (1) - b - c - bc]$$

Ejercicio: obtenga el resto de los contrastes B,C,AB,AC,BC y ABC

Si se hacen $n$replicas de cada tratamiento, los efectos de un diseño 2^3 se estiman dividiendo los contrastes entre $4n$, por ejemplo:

$$Efecto\ A = \frac{contraste\ A}{n2^{k-1}}$$

Las sumas de cuadrados para cada efecto que se requiere para la tabla de Anova, se calculan a partir de los contrastes de la siguiente forma,

$$SC_{efecto} = \frac{\left(contraste_{efecto}\right)^2}{n2^k}$$

La suma total de cuadrados se obtiene de la manera usual

$$SC_T = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{l=1}^{2}\sum_{m=1}^{n} Y_{ijlm}^2 - \frac{Y_{....}^2}{n2^k}$$

Con esta información se obtiene la tabla de Anova para el diseño 2^3.

Ejemplo diseño factorial 2^3.

En una empresa que fabrica dispositivos electrónicos se identificó mediante un análisis Pareto, que las fracturas de las obleas de silicio por choques térmicos, era la principal causa de obleas rotas en las etapas de procesamiento conocidas como “grabado mesa” y “piraña”. Un grupo de estas áreas identificó a tres factores principales (temperaturas) como las probables causas del problema, por ello se utilizó un experimento factorial 2^3 con el objetivo de localizar una combinación de temperaturas (tratamientos), en la cual se rompan un mínimo de obleas por efecto térmico. Los tres factores controlados y sus niveles en unidades originales, son:

Factor A=T1: Temperatura de grabado (-3°C y -1°C)

Factor B=T2: Temperatura de piraña (60°C y 98°C)

Factor C=T3: Temperatura de agua (20°C y 70°C)

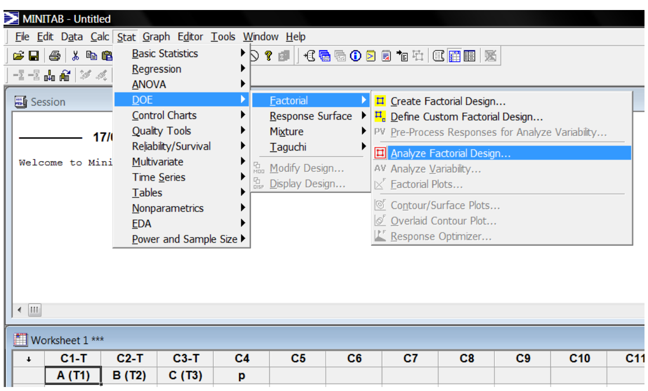
La empresa utiliza normalmente el tratamiento T1=-3°C T2=98°C y T3=20°C, es decir (A^-, B^+ y C^-) en el proceso diario. Así pues uno de los dos niveles de cada factor, es la temperatura usual y el otro es una temperatura que se supone reduce el efecto térmico sobre la oblea.

Se decide correr el experimento2^3 con 1 replicas es decir $2*2^3$= 8 tratamientos *2=16 corridas. Como la variable de respuesta es el número de obleas rotas observadas, se decide utilizar proporciones de obleas rotas, así, una proporción p=0.040 indica que el 4% salieron rotas. Los 16 resultados obtenidos son:

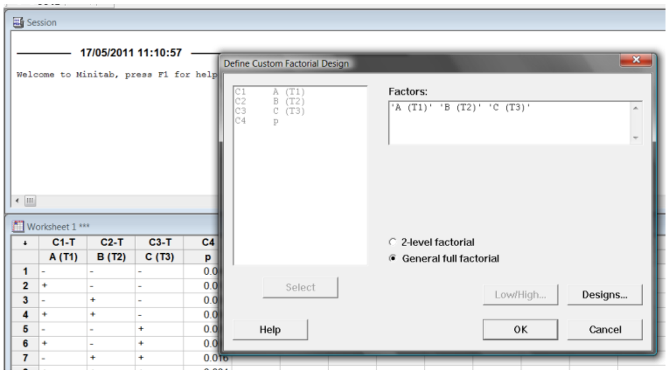
A (T1)	B (T2)	C (T3)	p
-	-	-	0.040
+	-	-	0.012
-	+	-	0.036
+	+	-	0.000
-	-	+	0.020
+	-	+	0.000
-	+	+	0.016
+	+	+	0.004
-	-	-	0.032
+	-	-	0.008
-	+	-	0.028
+	+	-	0.000
-	-	+	0.020
+	-	+	0.016
-	+	+	0.008
+	+	+	0.004

Diseño factorial 2^3 en minitab.

Para realizarlo en minitab ingresamos las columnas y con la secuencia *Stat / DOE / Factorial / Analize Factorial Design*



Ingresamos los factores A, B y C, y especificamos que es un General full factorial (caso general de diseño factorial), damos *ok*



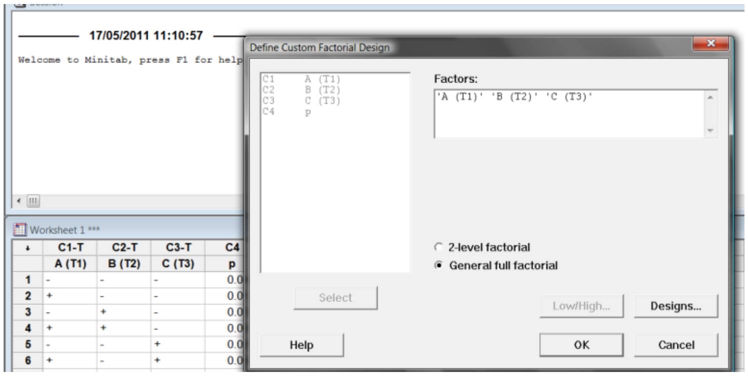
Ingresamos la variable de respuesta y verificamos la casilla *Terms*, que estén activados los 3 efectos y las 4 interacciones posibles (A, B, C, AB, AC, BC y ABC), luego damos ok al análisis, teniendo los siguientes resultados

General Linear Model: p versus A (T1), B (T2), C (T3)

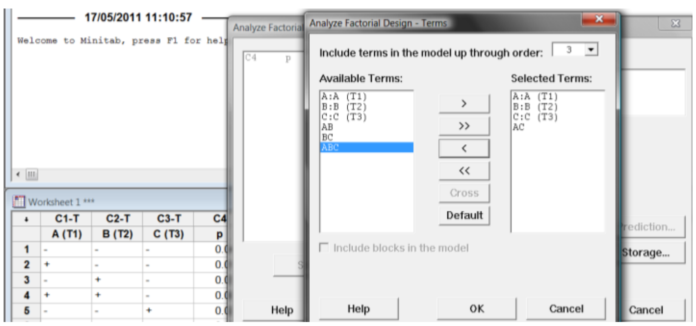
```
Factor  Type   Levels  Values
A (T1)  fixed       2  +, -
B (T2)  fixed       2  +, -
C (T3)  fixed       2  +, -


Analysis of Variance for p, using Adjusted SS for Tests

Source                DF     Seq SS     Adj SS     Adj MS      F      P
A (T1)                 1  0.0015210  0.0015210  0.0015210  52.45  0.000
B (T2)                 1  0.0001690  0.0001690  0.0001690   5.83  0.042
C (T3)                 1  0.0002890  0.0002890  0.0002890   9.97  0.013
A (T1)*B (T2)          1  0.0000010  0.0000010  0.0000010   0.03  0.857
A (T1)*C (T3)          1  0.0003610  0.0003610  0.0003610  12.45  0.008
B (T2)*C (T3)          1  0.0000010  0.0000010  0.0000010   0.03  0.857
A (T1)*B (T2)*C (T3)   1  0.0000250  0.0000250  0.0000250   0.86  0.380
Error                  8  0.0002320  0.0002320  0.0000290
Total                 15  0.0025990


S = 0.00538516   R-Sq = 91.07%   R-Sq(adj) = 83.26%
```

Observe en la tabla de Anova, que los efectos A, B, C y AC son significativos, es decir, el hecho de que estén presentes en nivel bajo o en nivel alto, si hace que cambie el comportamiento de la variable de respuesta Y (proporción de obleas rotas), por lo tanto se pudiera descartar las interacciones AB, BC y ABC, ya que no producen efecto sobre Y. Además observe que el pvalue del efecto B es 0.042, es decir, muy cerca del nivel de significancia común $\alpha = 0.05$, por lo que con muy poca fuerza (potencia) se rechaza la hipótesis nula de que su efecto es cero.

Nota: en estas situaciones, es recomendable quitar del análisis, a los factores que no fueron significativos (AB, BC y ABC) y mandar su variación al error aleatorio. Con esto podemos apreciar mejor a los que si fueron significativos (A, B, C y AC) y corroborar si se mantienen como tal. A esta técnica de experimentación se le conoce como el **Mejor Anova**, y se realiza de manera similar con la secuencia *Stat / DOE / Factorial / Analize Factorial Design. Luego* Ingresamos los factores A, B y C, y especificamos que es un General full factorial (caso general de diseño factorial), damos *ok*



Ingresamos la variable de respuesta y **en la casilla *__Terms__*, dejamos solamente los factores significativos (A, B, C y AC)** y mandamos a la izquierda a los no significativos (AB, BC y ABC) para posteriormente dar *__ok__* al análisis.



La hoja de sesión muestra ahora la nueva tabla Anova (**Mejor Anova**) ya sin las interacciones AB, BC, y ABC

General Linear Model: p versus A (T1), B (T2), C (T3)

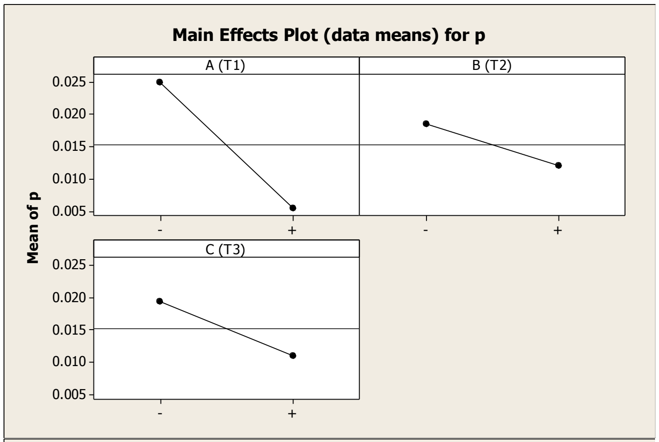
Factor	Type	Levels	Values
A (T1)	fixed	2	+, -
B (T2)	fixed	2	+, -
C (T3)	fixed	2	+, -

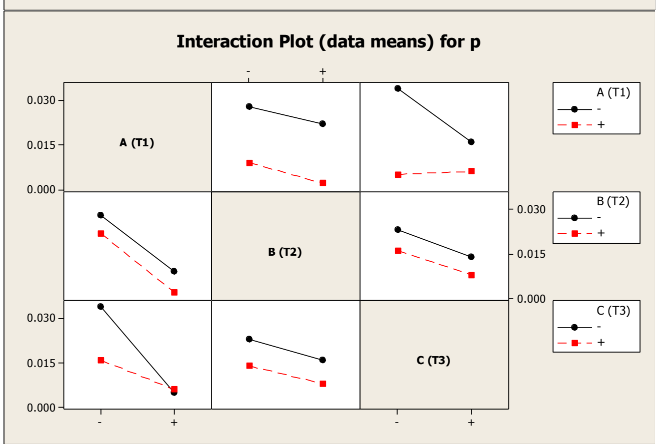
Analysis of Variance for p, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
A (T1)	1	0.0015210	0.0015210	0.0015210	64.60	0.000
B (T2)	1	0.0001690	0.0001690	0.0001690	7.18	0.021
C (T3)	1	0.0002890	0.0002890	0.0002890	12.27	0.005
A (T1)*C (T3)	1	0.0003610	0.0003610	0.0003610	15.33	0.002
Error	11	0.0002590	0.0002590	0.0000235		
Total	15	0.0025990				

Observe que los factores A, B, C y la interacción AC, siguen siendo significativos, y que ahora el factor B tiene un pvalue de 0.021, por tal motivo se confirma que estos cuatro efectos son los que realmente se tienen que controlar para que la cantidad de obleas rotas sea mínima.

Revisando los **efectos principales**, observamos que la cantidad de proporción de obleas rotas es menor cuando A+, B+ y C+. Sin embargo, cuando revisamos la gráfica de **interacción** se aprecia claramente que sólo AC tienen pendientes diferentes, el resto son prácticamente líneas paralelas (sin interacción). Además se observa que cuando A+ prácticamente el nivel de C no cambia el resultado de las proporciones, sin embargo si se usa A- si influye el nivel de C que se emplea. Por lo tanto se puede utilizar tanto C- como C+, siempre y cuando A este en su nivel alto (A+).





Así los efectos principales son

Efecto principal	**Estimación**
A (T1)	-0.0195
B (T2)	-0.0065
C (T3)	-0.0085
AB	-0.0005
AC	0.0095
BC	0.0005
ABC	0.0025

Observe que en términos absolutos, A, AC, C y B son los que tienen mayor efecto, esto coincide con los efectos significativos encontrados en la tabla de Anova.

Tenemos pues que los mejores tratamientos son (A+, B+, C+) y (A+, B+, C-). En términos de las unidades utilizadas, diríamos que los mejores tratamientos son (T1=-1°C, T2=98°C y T3=70°C) y (T1=-1°C, T2=98°C y T3=20°C).

B) DESARROLLO

practique analizando los problemas que proporcione el profesor

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

- Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
- Douglas C. Montgomery (2009). Design and Analysis of Experiments
- Douglas C. Montgomery y George C. Runger, (2005). Probabilidad y Estadística aplicadas a la Ingeniería. Editorial Limusa Wiley, segunda edición.
- Seymour Lipschutz y John Schiller, (2000). Introducción a la Probabilidad y Estadística. Editorial Mc Graw Hill.
- Humberto Gutiérrez y Román De La Vara, (2009). Control Estadístico de Calidad y Seis Sigma. Editorial Mc Graw Hill, segunda edición.

6.- ANEXOS:

NOMBRE DE LA PRÁCTICA	**DISEÑO FACTORIAL 3^k**	**PRÁCTICA NÚMERO**	**14**
PROGRAMA EDUCATIVO	**INGENIERIA INDUSTRIAL**	**PLAN DE ESTUDIO**	**2007-2**
NOMBRE DEL PROFESOR/A	**M.I. DIEGO ALFREDO TLAPA MENDOZA**	**NÚMERO DE EMPLEADO**	**20673**
LABORATORIO	**LABORATORIO DE COMPUTACION**	**FECHA**	20/09/11

EQUIPO-HERRAMIENTA REQUERIDO	**CANTIDAD**
LAP-TOP	1
CAÑON DE PROYECCION	1
COMPUTADORA DE ESCRITORIO	18

SOFTWARE REQUERIDO
MINITAB y EXCEL
OBSERVACIONES-COMENTARIOS

NOMBRE Y FIRMA DEL PROFESOR	**NOMBRE Y FIRMA DEL COORDINADOR DE PROGRAMA EDUCATIVO**

1.- INTRODUCCIÓN:

Continuando con los proceso reales y su comportamiento, en muchas ocasiones el ingeniero está interesado en probar varios factores, pero estos a tres niveles. En la presente práctica se analizará un experimento donde intervienen varios factores y cada uno de ellos a tres niveles por medio del empleo del software Minitab y Excel.

2.- OBJETIVO (COMPETENCIA):

Analizar experimentos con varios factores a tres niveles y predecir el comportamiento de una variable de respuesta, con la finalidad de decidir cual es la mejor opción utilizando el software Minitab.

3.- TEORIA:

3.1 Diseño Factorial 3^k

El diseño factorial 3^k considera k factores con 3 niveles cada uno y 3^k tratamientos. La primera desventaja de estos diseños, es la cantidad de pruebas que requiere, por ejemplo un diseño de 4 factores a 3 niveles sin replica implicaría 3^4=81 tratamientos. Mientras un 2^4 = 16 tratamientos, por lo que la diferencia de usar los mismos 4 factores ya sea a 2 niveles o 3 niveles es de 65 pruebas más.

En la práctica, usar 4 o más factores con 3 niveles, es demasiado complicado y consume muchos recursos. Por lo general un diseño factorial 3^k se justifica bajo alguna de las dos siguientes condiciones:

1. Se tienen factores de tipo continuo e interesa estudiar efectos cuadráticos (efectos de curvatura) como A^2, B^2, … , A^2B, B^2A, A^2B^2, etc. Esto se realiza cuando se cree que la variable de respuesta no es lineal.
2. Los factores son categóricos o discretos y de manera natural tienen tres niveles cada uno. Por ejemplo, si uno de los factores es un reactivo del cual existen tres marcas, resulta natural que se desee experimentar las tres marcas con la idea de comparar su desempeño.

Ejemplo de Diseño Factorial 3^k

En un proceso de fabricación de cajas se utiliza pegamento; con la idea de mejorar el desempeño de las cajas se realiza un experimento para estudiar la fuerza de adhesión del pegamento en diferentes condiciones de humedad y temperatura. La variable de respuesta es la fuerza necesaria en libras para poder despegar la caja. Los datos obtenidos en cada una de las nueve combinaciones de un Diseño Factorial 3^k con 2 replicas son:

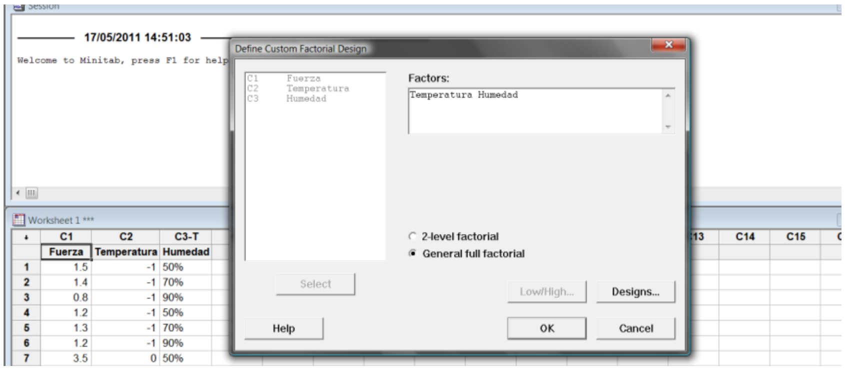
	Factor A (Temperatura)			
	Frio (-1)	**Ambiente (0)**	**Caliente (1)**	
Factor B	**Replica**	**Replica**	**Replica**	

(Humedad)	**1**	**2**	**1**	**2**	**1**	**2**	
50%	1.5	1.2	3.5	3.2	4.0	4.2	**17.6**
70%	1.4	1.3	2.9	2.5	3.8	3.4	**15.3**
90%	0.8	1.2	1.8	2.0	2.7	3.0	**11.5**
Totales	**7.4**		**15.9**		**21.1**		**44.4**

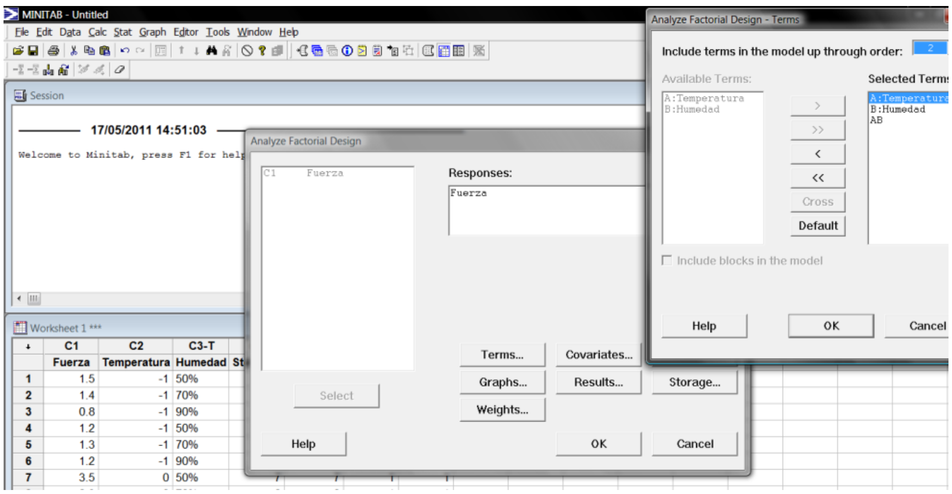
Observe que otra forma de acomodar los datos es

Fuerza	Temperatura	Humedad
1.50	-1	50%
1.40	-1	70%
0.80	-1	90%
1.20	-1	50%
1.30	-1	70%
1.20	-1	90%
3.50	0	50%
2.90	0	70%
1.80	0	90%
3.20	0	50%
2.50	0	70%
2.00	0	90%
4.00	1	50%
3.80	1	70%
2.70	1	90%
4.20	1	50%
3.40	1	70%
3.00	1	90%

Para realizarlo en minitab, básicamente se sigue el mismo procedimiento que cuando se realiza un 2^k; ingresamos los datos a una hoja de cálculo y seguimos la secuencia *Stat / DOE / Analyze Factors Designs*. Ingresamos los factores Temperatura y Humedad dentro del casillero *General Full Factorial* y damos *ok*



Posteriormente ingresamos la variable de respuesta (Fuerza) y verificamos en el casillero de *__Terms__*, que estén seleccionados los factores A, B y AB, para luego dar *__ok__* y posteriormente correr el análisis.



El resultado se muestra en la siguiente hoja de sesión

General Linear Model: Fuerza versus Temperatura, Humedad

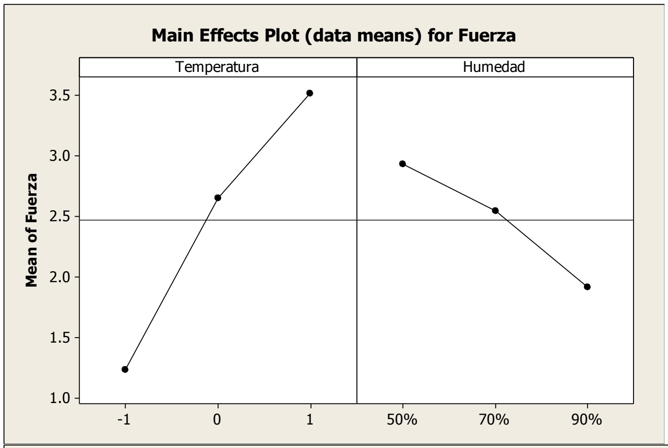
Factor	Type	Levels	Values
Temperatura	fixed	3	Ambiente, Caliente, Frio
Humedad	fixed	3	50%, 70%, 90%

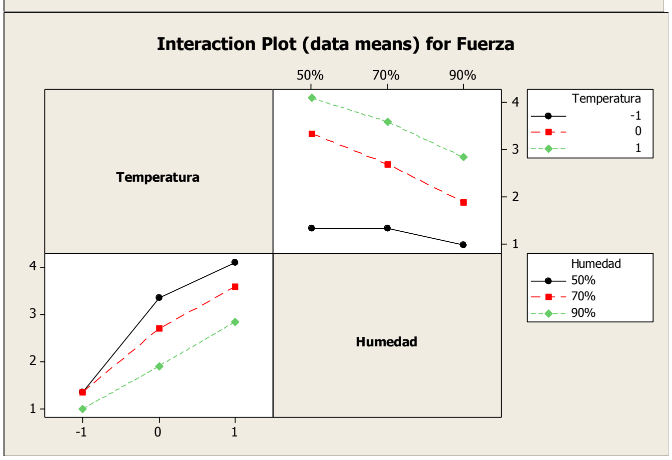
Analysis of Variance for Fuerza, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Temperatura	2	15.9433	15.9433	7.9717	170.82	0.000
Humedad	2	3.1633	3.1633	1.5817	33.89	0.000
Temperatura*Humedad	4	0.6933	0.6933	0.1733	3.71	0.047
Error	9	0.4200	0.4200	0.0467		
Total	17	20.2200				

S = 0.216025   R-Sq = 97.92%   R-Sq(adj) = 96.08%

Observe que los factores A=Temperatura, B=Humedad y la interacción AB son significativos, ya que su pvalue está por debajo del nivel de significancia $\alpha = 0.05$, sin embargo, la interacción es apenas significativa (es decir, su efecto es apenas diferente de cero), por lo que ayudaría bastante revisar las gráficas de efectos principales y de interacción y ver el comportamiento. Recuerde la secuencia en minitab para estas gráficas es <u>*Stat / Anova / Main Effects Plots*</u> y <u>*Stat / Anova / Interactions Plots*</u>. En ambos casos ingresar los factores y la variable de respuesta, para tener así la siguientes gráficas.





Observe que los efectos principales, tanto de Temperatura como de Humedad no son del todo lineales, es decir parece que siguen una pequeña curva, sin embargo predomina más el componente lineal que la curva. Se observa además que a mayor temperatura y menor humedad, es más efectivo es el proceso de pegado.

En la gráfica de interacción se corrobora que el mejor resultado es humedad 50% y temperatura caliente. Además se aprecia un pequeño efecto de interacción, ya que las líneas tienen pendiente similar, lo que comprueba que el efecto de interacción prácticamente no influye, tal como se vio en la tabla de Anova. De esta manera, para efectos prácticos se debería escoger **A** en nivel alto (temperatura caliente), **B** en nivel bajo (humedad al 50%) para así conseguir la mayor fuerza de adhesión.

B) DESARROLLO

Practique analizando los problemas que proporcione el profesor

C) CÁLCULOS Y REPORTE:

El reporte de la práctica consistirá en entregar electrónicamente el análisis de los problemas planteados previamente

D) RESULTADOS:

Mencione los hallazgos encontrados del análisis de los problemas tratados.

E) CONCLUSIONES:

Mencione los aprendizajes y competencias adquiridos en esta práctica

5.- BIBLIOGRAFÍA:

- Humberto Gutiérrez y Román De La Vara (2009). Análisis y diseño de experimentos.
- Douglas C. Montgomery (2009). Design and Analysis of Experiments
- Douglas C. Montgomery y George C. Runger, (2005). Probabilidad y Estadística aplicadas a la Ingeniería. Editorial Limusa Wiley, segunda edición.
- Seymour Lipschutz y John Schiller, (2000). Introducción a la Probabilidad y Estadística. Editorial Mc Graw Hill.
- Humberto Gutiérrez y Román De La Vara, (2009). Control Estadístico de Calidad y Seis Sigma. Editorial Mc Graw Hill, segunda edición.

6.- ANEXOS: