CAPITULO IV

PRUEBAS DE HIPOTESIS ESTADISTICA

La recolección y análisis de datos son componentes muy importantes de los *métodos científicos de investigación*. Los datos se utilizan para confirmar o rechazar teorías e hipótesis, y estas deben ser confirmadas con datos provenientes de experimentos cuidadosamente planificados. Si los datos contradicen las teorías formuladas, se deberá buscar adecuadas explicaciones o nuevas teorías. Estas ideas introductoras ilustran el hecho de que los datos obtenidos a partir de muestras se utilizan para decidir entre hipótesis elaboradas.

La inferencia estadística estudia los métodos que señalan como emplear datos obtenidos de muestras aleatorias para *inducir*, por generalización, características a las poblaciones de las que se ha obtenido las muestras. Más recientemente, los métodos de inferencia estadística se han unificado bajo los conceptos generales de la teoría de la decisión, es decir, bajo los conceptos generales de la manera de tomar decisiones en condiciones de incertidumbre.

Las pruebas de hipótesis estadística se encargan de analizar procedimientos para confirmar o rechazar hipótesis acerca de la distribución de variables aleatorias. Las hipótesis están confinadas a términos de parámetros de población; esto es, hecha una determinada afirmación sobre una población, deseamos saber si los resultados de una muestra contradicen o no tal afirmación. Por ejemplo, dos hipótesis pueden especificar que el promedio de la tensión de ruptura de cierta aleación es mayor (no es mayor) que 2000 lb.; o que la variabilidad de un proceso *A* es mayor (no es mayor) que aquella de un proceso *B*; etc.

Así como ocurrió con la construcción de los estimadores por intervalos de confianza, las pruebas de hipótesis también se apoyan en la distribución de probabilidad de los estimadores. De esta manera, las distribuciones de probabilidad de la media muestral y la varianza muestral, serán utilizadas para las respectivas pruebas sobre la media y la varianza poblacional.

4.1. Ejemplo ilustrativo.

Supóngase que se quiere determinar si ciertos *cambios* en un proceso productivo reducen el tiempo que le toma a un obrero completar una tarea de ensamblaje de un producto. Supóngase que en condiciones normales, en promedio, un trabajador emplea 30 minutos en completar la tarea de ensamblaje. También se sabe que ese tiempo promedio presenta variación; esto es, el tiempo de ensamblaje es una *variable aleatoria* que se puede representar por *X*. Mas aún, existe suficiente información para asegurar que *X* esta normalmente distribuida y tiene promedio μ = 30 y una desviación estándar σ = 1. Esto implica que aproximadamente 95% de los tiempos de ensamblaje se encuentran entre 28 y 32 minutos.

En este ejemplo se va a evaluar una hipótesis relacionada al parámetro μ, con el supuesto de que $\sigma = 1$ aún después de los cambios efectuados. Luego de concretados los cambios en el proceso de ensamblaje, se plantean dos hipótesis:

- La hipótesis del *no cambio* (hipótesis nula), que establece que μ sigue siendo 30 minutos; y
- La hipótesis alternativa que sugiere que μ < 30 minutos:

Para determinar la validez de una de estas hipótesis se procede a la recolección de información.

Primeramente se elige aleatoriamente a un solo trabajador para evaluar el efecto de los cambios realizados. Se observa que este trabajador emplea 29 minutos en la tarea normal de ensamblaje. Con este dato, y teniendo en consideración que *X* proviene de *N*(30, 1) se evalúa:

$$P[X \le 29] = \phi\left(\frac{29-30}{1}\right) = 0.1587$$

es decir, hay aproximadamente 16 % de probabilidades de que el tiempo de ensamblaje sea 29 minutos, siendo que *X* proviene de *N*(30, 1). Esto indica de que un tiempo de 29 min. no es un evento muy raro en esa distribución. Continuando la toma de datos se evalúa *n* = 5 trabajadores, para los cuales $\bar{X} = 29\,\mathrm{min.}$ Luego se toma *n* = 25 trabajadores que presentan $\bar{X} = 29\,\mathrm{min.}$ Reiterando que X proviene de *N*(30, 1), la probabilidad de que $\bar{X} = 29\,\mathrm{min.}$ para 25 trabajadores es:

$$P[\bar{X} \le 29] = \phi\left(\frac{29-30}{\frac{1}{\sqrt{25}}}\right) = P\left[\bar{X} \le -5\right] = 0.0000003$$

Esto hace que para una $N(30, 1)$, con n = 25, $\bar{X} = 29\,\text{min}.$ es un evento extremadamente improbable. Ciertamente, esto quita consistencia a la hipótesis de que μ sigue siendo 30 minutos. Por ello se puede afirmar que, efectivamente, el tiempo de ensamblaje de cada trabajador se ha reducido.

En términos técnicos, se dice que: los cambios efectuados en el proceso productivo originan una diferencia *estadísticamente significativa* en el tiempo de ensamblaje; o bien, que la diferencia en el tiempo, 30 - 29 = 1 es estadísticamente significativa.

Es necesario aclarar que una diferencia estadísticamente significativa no siempre es de utilidad práctica; ya que por ejemplo, la reducción de un minuto en el promedio μ no podría representar ahorros importantes en el proceso productivo. La decisión final de adoptar los cambios concierne a una evaluación de estructura de costos, que estas técnicas estadísticas no pueden resolver.

4.2. Definición de hipótesis

Para llegar a tomar decisiones, conviene hacer determinados supuestos o conjeturas acerca de las poblaciones que se estudian. Tales supuestos, que pueden ser ciertos ó no, se denominan *hipótesis estadísticas*. Por ejemplo, si se quiere decidir si un procedimiento es mejor que otro, se formulan las hipótesis:

- No hay diferencia estadística entre los dos procesos; y
- Si existe diferencia estadística entre los dos procesos.

La primera se define como la hipótesis del *"no cambio"* o “*hipótesis nula”* y se simboliza por H_o. La segunda se denomina “*hipótesis alternativa”* y se simboliza por H_1.

Por ejemplo, una empresa produce circuitos impresos con probabilidad histórica de producto fallado igual al 3% (P = 0.03). Un ingeniero de producción sugiere cambios en el proceso y asegura que tales cambios causarán una reducción de la probabilidad de falla; es decir, asegura que P < 0.03. Por lo tanto, se deberá evaluar:

$H_o : P = 0.03$ *(Hipótesis Nula)*

$H_1 : P < 0.03$ *(Hipótesis Alternativa)*

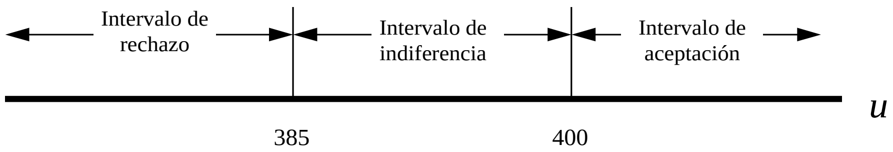
4.2 Nivel de significación

Se acostumbra a fijar niveles de significación de 0.1, 0.05 y 0.01, aunque se puede fijar otros valores. Si por ejemplo, se fija un nivel de significación del 0.05 (5%) al diseño de un ensayo de hipótesis, entonces hay aproximadamente 5 ocasiones en 100 en que se rechazaría una hipótesis cuando debería ser aceptada. En otras palabras, se estaría 95% seguro de que se toma la decisión adecuada.

4.3 Tipos de errores

Definidas las hipótesis es posible cometer dos tipos de error. Se ilustra estos tipos de errores con el siguiente ejemplo.

Un fabricante de pintura asegura que un galón de pintura producida en su planta puede cubrir en promedio 400 pies cuadrados de superficie. Se toma una muestra de 36 galones y se observa que en promedio cubren 385 ft^2. Siendo que el promedio μ es desconocido, se establece el siguiente criterio técnico para tomar decisiones con respecto a esa pintura:



Según esto, es importante hacer notar que $\bar{x}$ puede exceder los 400 ft^2 aunque μ sea menor que 385 ft^2 o que $\bar{x}$ sean menor que 385 ft^2 y que μ exceda los 400 ft^2 . Esto ilustra el hecho de que los errores son inevitables cuando las decisiones se toman apoyándose en los resultados de muestras aleatorias. Consecuentemente, estos errores pueden ser de dos tipos diferentes y se describen esquemáticamente en la siguiente figura:

	u se encuentra en la zona de aceptación	*u* se encuentra en la zona de rechazo
Se acepta la afirmación	No hay Error	Error TIPO II
Se rechaza la afirmación	Error TIPO I	No hay Error

Nótese que si μ esta en el intervalo de indiferencia, no hay ningún error grave, cualquiera sea la decisión tomada.

Para juzgar las ventajas de cualquier criterio de decisión, es esencial conocer las probabilidades de tener errores del Tipo I y del Tipo II. Estas probabilidades se denotan con α y β respectivamente:

α = *P[Error tipo I]*

β = *P[Error tipo II]*

4.4. <u>Hipótesis referente a la media</u>

Se trata de contrastar la hipótesis de la <u>media,</u> $\bar{x}$, de una población igual a un valor determinando frente a otra alternativa conveniente. Es decir, hay que contrastar:

$H_o : \mu = \mu_o$ $\quad\bar{x}$ es igual que μ_o

Frente a una de las siguientes hipótesis:

$H_1 : \mu < \mu_o$ $\quad\bar{x}$ es significativamente menor que μ_o

$H_1 : \mu > \mu_o$ $\quad\bar{x}$ es significativamente mayor que μ_o

$H_1 : \mu \neq \mu_o$ $\quad\bar{x}$ es significativamente diferente que μ_o

Para efectuar una prueba con un nivel de significación α se escoge una región crítica (es decir una región de rechazo de H_o) para $\bar{x}$ < c, tal que:

$P[\bar{x} \le c;\ \mu = \mu_o] = \alpha$

Ya que bajo H_o $\bar{x}$ es $N(\mu_o, \sigma^2/n)$, la región crítica esta dada por:

$$\frac{\bar{X}-\mu_o}{\frac{\sigma}{\sqrt{n}}} \le -z(\alpha)$$

donde $z(\alpha)$ es el punto superior de porcentaje de $N(0, 1)$.

La cantidad:

$$\frac{\bar{X}-\mu_o}{\frac{\sigma}{\sqrt{n}}} = Z$$

se denomina el estadístico estandarizado o tipificado. Si este estadístico es un valor menor de - $z(\alpha)$, se rechaza H_o y se acepta H_1. Por el contrario, si Z es mayor que $-z(\alpha)$, se dice que no hay suficiente evidencia para rechazar H_1, por lo que se acepta H_o.

En general, las regiones críticas para contrastar $H_o : \mu = \mu_o$ se pueden expresar como se indica en la Tabla 4.1:

Tabla 4.1: Regiones críticas para contrastar $H_O : \mu = \mu_O$

Hipótesis alternativa, H_1	Rechazar $H_O : \mu = \mu_O$ Si:
$\mu < \mu_O$	$Z < -z(\alpha)$
$\mu > \mu_O$	$Z > +z(\alpha)$
$\mu \neq \mu_O$	$Z < -z(\alpha/2)$ ó $Z > +z(\alpha/2)$

4.4.1. Hipótesis referente a la media: Cola inferior

Para entender la correcta orientación de la región crítica, la Figura 4.1 es muy útil para el caso de

$H_O : \mu = \mu_O$ y;

$H_1 : \mu < \mu_O$.

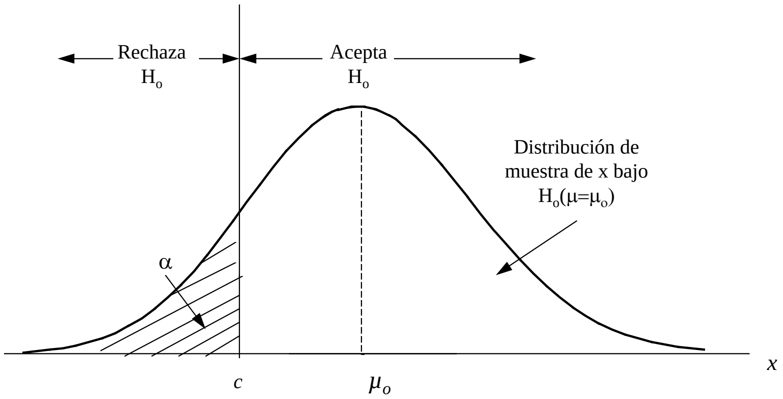


Figura 4.1: Ilustración gráfica de la prueba $H_O : \mu = \mu_O$ versus $H_1 : \mu < \mu_O$

En la Figura 4.1,"*c*" representa el valor crítico. La región a la derecha de este punto es la región de aceptación de H_O y la región a la izquierda de "*c*" es la de aceptación de H_1 . El punto *c* es tal que la probabilidad de rechazar H_O es α. El gráfico muestra que el valor crítico "*c*" debe ser menor que μ_O y es determinado por:

$$c = \mu_o + [z(\alpha)]\frac{\sigma}{\sqrt{n}}$$

El valor de "α" depende de cada aplicación. Un valor muy frecuentemente utilizado es α = 0.05. si altos costos están asociados a la evaluación, conviene utilizar valores más pequeños de α. Sin embargo, para valores pequeños de α implica un valor alto de $z(\alpha)$ y la prueba se torna muy conservadora, tal que se rechaza H_O solo en casos muy extremos.

Otra forma de conducir las pruebas de hipótesis es utilizando el concepto de ***valor de probabilidad*** o el valor ***'p'.*** Por ejemplo, en el caso de disminución de tiempos de ensamblaje de un producto de *n* = 34, si $\bar{x}$ = 29.68 para *N*(30, 1), se puede obtener un valor *p*, tal que:

$$p = P[\bar{X} \le \bar{x}] = P\left[Z \le \frac{\bar{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}}\right] = P\left[Z \le \frac{29.68 - 30}{\frac{1}{\sqrt{34}}}\right] = \phi(-1.866) = 0.031$$

En este caso se rechaza H_O si el valor de *p* es menor que α. = 0.05 (es decir para un nivel de significación de 0.05). Esto se ilustra en la Figura 4.2.

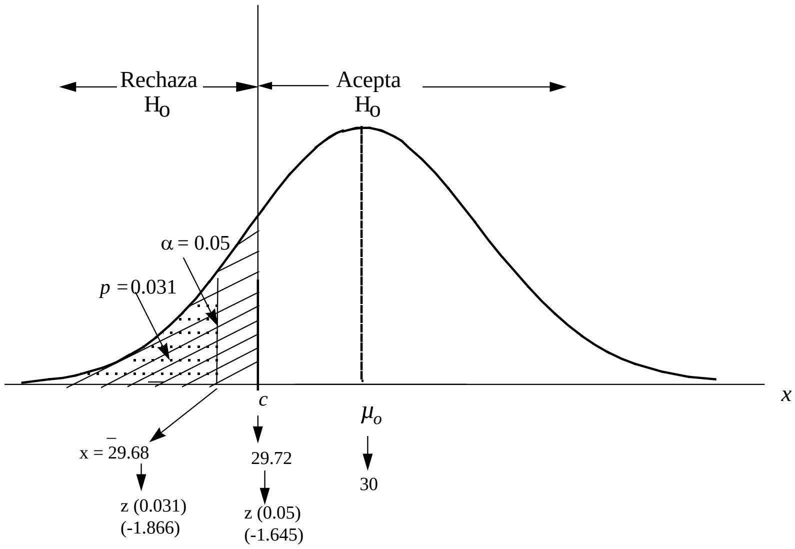


Figura 4.2: Ilustración del valor de probabilidad o el valor *'p'*.

4.4.2. Hipótesis referente a la media: Cola superior

Si ahora se trata de decidir si un procedimiento determinado ha producido un incremento en la circunstancia de probar H_o : $\mu = \mu_o$ versus H_1 : $\mu > \mu_o$. En este caso, se rechaza H_o en favor de H_1 cuando $\bar{x}$ > c, como se ilustra en la Figura 4.3.

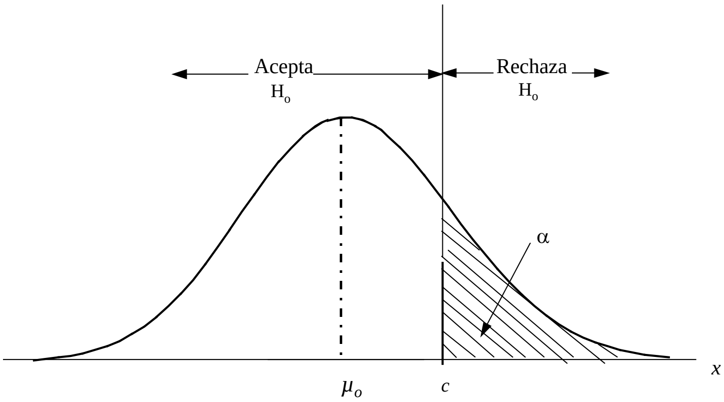


Figura 4.3: Ilustración gráfica de la prueba H_o : $\mu = \mu_o$ versus H_1 : $\mu > \mu_o$

En la Figura 4.3 se quiere probar a un nivel de significación α, se requiere que una probabilidad de falso rechazo de H_o sea α, es decir:

$$P[\bar{X} \geq c; \mu = \mu_o] = P\left[Z = \frac{\bar{X} - \mu_o}{\frac{\sigma}{\sqrt{n}}} \geq \frac{c - \mu_o}{\frac{\sigma}{\sqrt{n}}} \right] = \alpha$$

de modo que se rechaza H_o si el estadístico estandarizado:

$$\frac{\bar{X} - \mu_o}{\frac{\sigma}{\sqrt{n}}} \geq z(\alpha) \qquad o \ equivalentemente$$

$$valor\, p = P\left[\bar{X} \geq \bar{x};\ \mu = \mu_o \right] = 1 - \phi\left[\frac{\bar{X} - \mu_o}{\frac{\sigma}{\sqrt{n}}} \right]$$

4.3.3. Hipótesis referente a la media: Dos colas

Las pruebas $H_0: \mu = \mu_0$ versus $H_1: \mu < \mu_0$ y $H_0: \mu = \mu_0$ versus $H_1: \mu > \mu_0$ son llamadas pruebas a un *solo nivel o de una cola*. Hay ocasiones en que pruebas a dos lados son apropiadas, de modo que se contrasta:

$H_0: \mu = \mu_0$ versus $H_1: \mu \neq \mu_0$.

Esto se ilustra en la Figura 4.4.

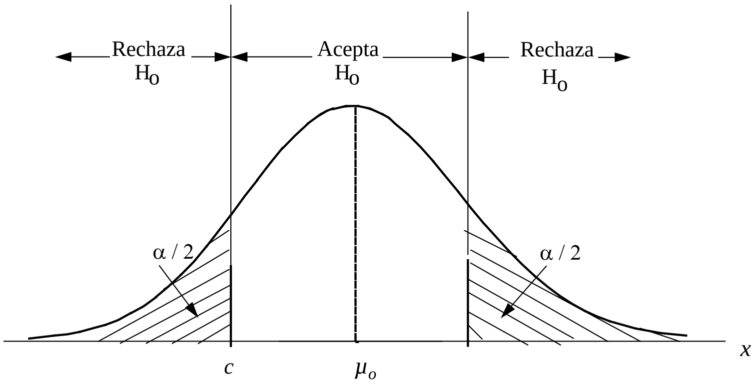


Figura 4.4: Ilustración gráfica de la prueba $H_0: \mu = \mu_0$ versus $H_1: \mu \neq \mu_0$

En este caso, con un nivel de significación α., se rechaza H_0 si:

$$\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| \geq z(\alpha/2) \qquad valor\ p = 2P\left[Z > \left|\frac{\bar{X}-\mu_o}{\frac{\sigma}{\sqrt{n}}}\right|\right]$$

4.4. <u>Hipótesis referente a medias de muestras pequeñas.</u>

Los anteriores conceptos asumen que se tiene una distribución *N*(0,1) con σ^2 conocida. Si ahora la varianza σ^2 es estimada por la varianza de muestra s^2, la distribución de muestra:

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$$

es una distribución *t* con n-1 = r grados de libertad y los valores críticos $z(\alpha)$ deben ser reemplazados por $t(\alpha; n\text{-}1)$. Si acaso *n* es grande (por lo menos 30) se puede utilizar la Tabla normal.

Las regiones críticas resultantes se muestran en la Tabla de *t-Student*, en donde por ejemplo $t(\alpha; n\text{-}1)$ es el área a la derecha bajo la curva de distribución *t* con *n*-1 grados de libertad y es igual a α, Figura 4.5

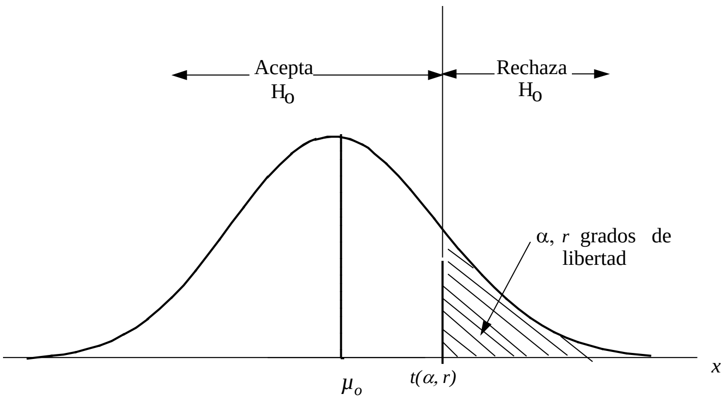


Figura 4.5: Ilustración gráfica de la prueba $H_0: \mu = \mu_0$ versus $H_1: \mu > \mu_0$ para muestras pequeñas.

<u>Tabla 4.2</u>: <u>Regiones críticas para contrastar</u> $H_0: \mu = \mu_0$

Hipótesis alternativa, H_1	Rechazar H_o : $\mu = \mu_o$ Si:
$\mu < \mu_o$	$t < -t(\alpha,r)$
$\mu > \mu_o$	$t > +t(\alpha,r)$
$\mu \neq \mu_o$	$t < -t(\alpha/2,r)$ ó $t > +t(\alpha/2,r)$

4.5. Hipótesis referente a dos medias

Una de las pruebas más importante que se hace en estadística es aquellas para las que se compara dos métodos diferentes. Por ejemplo, si se han considerado dos tipos de acero para ser usado en ciertas vigas de estructura metálica, se tomarán muestras y decidirá cual es mejor al comparar sus resistencias medias. Generalmente se compara las medias de dos distribuciones, de lo que resulta que se establecen las hipótesis nula y alternativa:

$H_o : \mu = \mu_o$ versus $H_1 : \mu < \mu_o$, ó $H_1 : \mu > \mu_o$ ó $H_1 : \mu \neq \mu_o$.

4.5.1 Comparación de dos muestras independientes grandes

Consideremos dos muestras aleatorias cada una con una distribución independiente y con promedios μ_1 y μ_2 , varianzas σ_1^2 y σ_2^2 respectivamente y de tamaño n_1 y n_2, denotadas por: $x_1, x_2, \dots, x_{n_1}$ y $y_1, y_2, \dots y_{n_2}$

y que son aproximadamente:

$$N(\mu_1, \sigma_1^2 / n_1) \text{ y } N(\mu_2, \sigma_2^2 / n_2)$$

Se asume que σ_1^2 y σ_2^2 son conocidas.

Bajo la hipótesis nula H_o : $\mu = \mu_o$ ó que $\mu - \mu_o = 0$, la variable aleatoria:

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

es $N(0,1)$. Se rechaza H_o y se acepta H_1 cuando $Z \geq z(\alpha)$ a una prueba de nivel de significación α.

4.5.2 Comparación con pequeñas muestras

Si las varianzas σ_1^2 y σ_2^2 son desconocidas, estas pueden ser reemplazadas por las varianzas de muestra, S_1^2 , S_2^2. En este caso se toma en consideración los grados de libertad $r_1 = n_1 - 1$ y $r_2 = n_2 - 1$, y bajo la hipótesis nula H_o : $\mu = \mu_o$, la variable aleatoria:

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{(n_1 - 1)S_x^2 + (n_2 - 1)S_y^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

tiene una distribución *t-Student* $t(\alpha; n_1 + n_2 - 2)$. Si se rechaza H_o: $\mu = \mu_o$ y se acepta $H_1 : \mu > \mu_o$, cuando $T \geq +t(\alpha; n_1 + n_2 - 2)$ se obtiene una prueba con nivel de significación α. Similarmente, en una prueba H_o: $\mu = \mu_o$ versus $H_1 : \mu < \mu_o$, se acepta H_1 cuando $T \leq -t(\alpha; n_1 + n_2 - 2)$. En una prueba H_o: $\mu = \mu_o$ versus $H_1 : \mu \neq \mu_o$, se acepta H_1 cuando $|T| \geq t(\alpha; n_1 + n_2 - 2)$.

4.5.3 Pruebas de comparación por pares

Hasta aquí se ha asumido que las observaciones provienen de dos grupos independientes entre si. En la mayoría de aplicaciones esto es válido, pero hay circunstancias en las que se debe asumir *dependencia* entre las observaciones. Por ejemplo, considérese una máquina que mide la dureza de un metal por la profundidad de la marca hecha por la punta de la máquina en la probeta a evaluar. Esta máquina dispone de dos puntas, diferentes la una a la otra, aun cuando la variabilidad de las dos parece ser la misma.

Si se desea evaluar si existe diferencia en las medidas de dureza de ambas puntas, se puede optar por ejemplo en tomar 20 probetas de metal y aleatoriamente escoger 10 probetas para cada una de las puntas. Este sencillo y lógico procedimiento puede tener desventajas. Supóngase que las 20 probetas fueron obtenidas de distintas existencias del mismo material, que podrían tener algunas diferencias físicas entre si, (por ejemplo, diferente dureza entre los distintos lotes). Esto contribuiría a incrementar el error experimental, variabilidad y haría más difícil detectar las diferencias entre ambas puntas.

Para salvar esta posibilidad de obtener una conclusión errónea, se puede optar por otro procedimiento. Si cada probeta es lo suficientemente grande, se puede utilizar cada una de las probetas para medir la dureza con las dos puntas. El orden y ubicación de medición se deberá ejecutar aleatoriamente. En este caso las mediciones están mutuamente relacionadas y se dice que hay dependencia entre las observaciones.

En tales situaciones, se analiza las diferencias $W_i = X_i - Y_i$. Estas diferencias:

$W_1 = X_1 - Y_1$; $W_2 = X_2 - Y_2$; ...; $W_n = X_n - Y_n$ son muestras aleatorias de tamaño n provenientes de una distribución con promedio $\mu_1 - \mu_2$ y varianza σ_w^2 Las variables $W_1, W_2, ..., W_n$ son independientes (provienen de diferentes objetos). Asumiendo que W es normal, se puede utilizar la prueba estadística:

$$T = \frac{\overline{W}}{S_w / \sqrt{n}}$$

donde S_W es la desviación estándar de las diferencias, W. Una prueba de hipótesis por ejemplo puede ser H_0: $\mu = \mu_0$ versus H_1 : $\mu > \mu_0$. Se acepta H_1 si $\overline{W} / \left(S_w / \sqrt{n} \right) \geq t(\alpha, n-1)$ Si n es suficientemente grande, se puede utilizar la distribución normal.

4.6 <u>Prueba para $\sigma^2 = \sigma_0^2$</u>

Si se desea probar la hipótesis de que la varianza de una población normal σ^2 es igual a un valor específico, por ejemplo, σ_0^2. Sea $X_1, X_2, ..., X_n$ una muestra aleatoria de n observaciones tomadas de esta población. Para probar:

$$H_o : \sigma^2 = \sigma_0^2$$
$$H_1 : \sigma^2 \neq \sigma_0^2$$

se utiliza el estadístico de prueba:

$$\chi_o^2 = \frac{(n-1) S^2}{\sigma_o^2}$$

donde S^2 es la varianza muestral. Ahora si H_o es verdadera, entonces el estadístico de prueba χ_o^2 sigue una distribución *chi cuadrada* con *n-1* grados de libertad. Por consiguiente, se calcula el valor de la estadística de prueba χ_o^2 y la hipótesis $H_o : \sigma^2 = \sigma_0^2$ debe rechazarse si:

$\chi_o^2 > \chi_{\alpha/2,\, n-1}^2$ o si $\chi_o^2 < \chi_{1-\alpha/2,\, n-1}^2$

donde $\chi_{\alpha/2,\, n-1}^2$ $\chi_{1-\alpha/2,\, n-1}^2$ son los puntos que corresponden a los porcentajes 100α/2 inferior y superior a la distribución *chi cuadrada* con *n-1* grados de libertad respectivamente.

El mismo estadístico se utiliza para la hipótesis alternativas unilaterales. Para la hipótesis unilateral:

$$H_o : \sigma^2 = \sigma_0^2$$
$$H_1 : \sigma^2 > \sigma_0^2$$

se rechaza H_o si $\chi_o^2 > \chi_{\alpha/2,\, n-1}^2$

Para la otra hipótesis unilateral:

$$H_o : \sigma^2 = \sigma_0^2$$
$$H_1 : \sigma^2 < \sigma_0^2$$

se rechaza H_o si $\chi_o^2 < \chi^2_{1-\alpha/2,\, n-1}$

4.8 Prueba para $\sigma_1^2 = \sigma_2^2$

Supóngase que se tiene interés en dos poblaciones normales independientes, donde las medias y varianzas de la población, $\mu_1, \sigma_1^2, \mu_2 \ y, \sigma_2^2$ son desconocidos. Se desea probar las hipótesis sobre la igualdad de las dos varianzas H_o :_ $\sigma_1^2 = \sigma_2^2$ por ejemplo. Si se tienen dos muestras aleatorias de tamaño n_1 tomada de la población 1, y otra de tamaño n_2 proveniente de la población 2 y sean S_x^2 y S_y^2 las respectivas varianzas muestrales. Para probar las alternativas:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 > \sigma_2^2$$

se utiliza el estadístico:

$$F = \frac{S_1^2}{S_2^2}$$

tiene una distribución F con n_1-1 grados de libertad en el numerador y n_2-1 grados de libertad en el denominador.

Se rechaza $H_0 : \sigma_1^2 = \sigma_2^2$ y se acepta $H_1 : \sigma_1^2 > \sigma_2^2$ si $S_x^2 / S_y^2 \geq F(\alpha; n_1 - 1, n_2 - 1)$

4.9 Comparación de muestras cuando $\sigma_1^2 \neq \sigma_2^2$

Hasta ahora las inferencias referentes a la diferencia entre medias de dos poblaciones se han presentado bajo el supuesto de que las varianzas poblacionales son conocidas o desconocidas, pero iguales. Si las varianzas no son iguales, al usar los estimadores insesgados $S_1^2 \ y \ S_2^2$ en vez de $\sigma_1^2 \ y \ \sigma_2^2$ la cantidad:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

no sigue la distribución *t-student* con $n_1 + n_2 - 2$ grados de libertad . Pero si se estima los grados de libertad por la fórmula:

$$r = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1+1} + \frac{(S_1^2/n_1)^2}{n_1+1}} - 2$$

Donde "*r*" corresponde a los grados de libertad. Entonces, la expresión

$$t^* = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Se aproxima a la distribución *t-student* con *r* grados de libertad calculados en la anterior ecuación. Si el valor de *r* no es un valor entero, se le aproxima al entero más cercano. Con esto, las pruebas de hipótesis se ejecutan de igual manera a lo descrito para muestras pequeñas de varianza desconocida.

Ejemplo:

Para probar la efectividad de dos pegamentos se utilizan 41 moldes pegados con cada uno de esos pegamentos. Se mide la fuerza para romper los moldes pegados con los siguientes resultados:

Pegamento 1:	$n_1 = 41$	$\bar{x}_1 = 45\ ft/lb$	$S_1 = 6.2$ ft/lb
Pegamento 2:	$n_2 = 41$	$\bar{x}_1 = 42\ ft/lb$	$S_2 = 4.3$ ft/lb

Probar la hipótesis Ho: $\mu_1 = \mu_2$ con $\alpha = 0.05$

4.10. ***Ejercicios de aplicación***

Se realizaron ensayos estándar de Bond repetidos sobre dos muestras A y B. Se obtuvieron los siguientes valores:

	Wi (kWH/t corta)			
Muestra	1	2	3	4
A	12,0	11,7	12,5	11,9
B	12,3	11,5	11,8	12,0

Existe diferencia significativa en dureza en las muestras A y B al nivel de confianza del 95 %? Suponga que $\sigma_A^2 = \sigma_B^2$

4.10.1 Comparando dos muestras

Las medidas de tensión de compresión de cierto tipo de concreto se ha realizado en dos máquinas. Los resultados de 10 muestras son como sigue:

X_1	83	87	85	82	86	83	91	85	79	87
X_2	88	89	90	84	85	84	89	79	81	90

a) Encuentre el promedio y varianza de X_1 y X_2.

b) Verificar si las dos máquinas dan el mismo promedio de tensión a la compresión para el concreto. Es decir, verificar si $\mu_1 = \mu_2$

4.10.2. Inferencia con muestras pequeñas y varianza desconocida

La Empresa Uranitec utiliza un proceso de extracción por solventes (*SX, Solvent Extraction*) para recuperar Uranio y vanadio de soluciones de lixiviación. La tecnología implementada inicialmente por Uranitec recupera Uranio tal que la fracción en masa de Uranio en el solvente , m_u, es 0.47 con una varianza de 0.003

En una investigación reciente se demostró que una configuración diferente de los deflectores en el tanque de extracción podía incrementar el valor de m_u. El metalurgista de planta decidió el cambio de los deflectores en uno de los tanques de *SX* en planta de acuerdo al diseño propuesto de la investigación. Por varias semanas procedió con ello a recolectar información para comparar el antiguo y nuevo diseño. Los datos se presentan en la Tabla de a continuación, (Fracción de masa de uranio).

Tabla: Fracción en masa de Uranio con en antiguo sistema y el nuevo en los tanques de *SX.*

Fecha,	Diseño antiguo	Diseño nuevo	Fecha	Diseño antiguo	Diseño nuevo	Fecha	Diseño antiguo	Diseño nuevo
1	0.535	0.502	12	0.458	0.518	22	0.502	0.536
2	0.522	0.561	13	0.480	0.507	23	0.494	0.518
3	0.417	0.614	14	0.464	0.405	24	0.464	0.532
4	0.402	0.478	15	0.437	0.481	25	0.480	0.523
5	0.460	0.506	16	0.466	0.485	26	0.470	0.462
6	0.477	0.562	17	0.480	0.595	27	0.417	0.551
7	0.493	0.469	18	0.486	0.443	28	0.520	0.503
8	0.408	0.559	19	0.458	0.564	29	0.520	0.507
9	0.379	0.476	20	0.505	0.534	30	0.368	0.488
10	0.548	0.481	21	0.518	0.459	31	0.435	0.475
11	0.504	0.420						

Con esta información determine:

1. Después de que los deflectores fueron cambiados, se mejoró la recuperación de Uranio?.

a) Compare los resultados en la primera semana (día 1 - 7)

b) Compare los resultados en las dos primeras semanas (días 1 - 14).

c) Compare los resultados para el mes completo (días 1 - 31).

2. ¿Cuál es el mejor estimado de recuperación con el nuevo diseño de deflector? En otras palabras, ¿la recuperación de Uranio se incremento progresivamente en el mes, disminuyó en el mes o se mantuvo constante en el mes?. Las siguientes pruebas pueden dar respuesta a la pregunta:

a) $\mu_1 \geq \mu_2$?

b) $\mu_1 \geq \mu_3$?

c) $\mu_2 \geq \mu_3$?

Verificar esto con el 95% de confianza.

3. Para el nuevo diseño de deflector, determine el mejor estimado de recuperación junto con el intervalo que contenga 97.5% de veces el verdadero promedio.

4. Determine el número de días necesarios de prueba para producir un intervalo especificado en la sección (3) que se encuentre a ± 0.06 del estimado de recuperación.

4.10.3 Aplicación de la Inferencia Estadística para evaluación de mejoras en un proceso de flotación de minerales

Usted en su posición de Jefe Metalurgista recibe la visita de un vendedor que dice que la compañía que representa tiene un reactivo de flotación nuevo al que le denominan "*Float All*". El vendedor asegura que con este reactivo se puede incrementar la recuperación de sulfuro de cobre en el circuito de flotación desde un 89 % hasta un 95%. El precio de "*Float All*" (US $/lb.) y la adición de colector (lbs/ton de mineral) son equivalentes al que está en actual uso en la planta concentradora. Tales perspectivas anunciadas por el vendedor se ven muy atractivas y ameritan una investigación adecuada, así que usted asigna a un metalurgista de planta para que efectúe las evaluaciones del nuevo colector y aconseje sobre el cambio de reactivo.

El metalurgista de planta realiza pruebas con "*Float All*" durante 06 días. Luego compara estos resultados con los correspondientes a un período de 20 días previos de operación con su reactivo rutinario. Los datos obtenidos figuran en la siguiente Tabla:

Datos obtenidos por el metalurgista de planta

(A) Reactivo de Planta			
Día	*Recuperación total, %*	*Día*	*Recuperación total, %*
1	88.5	11	87.3
2	86.1	12	82.6
3	93.2	13	86.2
4	85.9	14	86.5
5	90.6	15	93.6
6	93.1	16	90.3
7	96.2	17	86.2
8	88.3	18	90.3
9	87.8	19	87.4
10	85.5	20	87.8
(B) Reactivo Float All			
Día	*Recuperación total, %*	*Día*	*Recuperación total, %*
1	89.5	4	91.8
2	92.7	5	90.7
3	93.3	6	96.6

Al analizar los datos obtenidos, calculando el promedio y varianza de los dos grupos de datos, el metalurgista de planta se plantea las siguientes pruebas de hipótesis:

1. La varianza de las recuperaciones con ambos colectores es igual.
2. El promedio de recuperación utilizando "Float All" es de 95 % (ó mayor).
3. El promedio de recuperación con ambos reactivos es igual

Luego del análisis respectivo, concluyó:

1. El promedio de recuperación en planta con “Float All” será menor al 95%.
2. Sin embargo, “Float All” mejorará las recuperaciones con respecto al reactivo en actual uso.
3. El reactivo actual debe ser reemplazado por “Float All”.

Usted sabe por experiencia que las características de mineral son cambiantes incluso de un día al otro. Con esto en mente, usted revisa los datos obtenidos y decidirá si las conclusiones del metalurgista de planta son válidas. Luego de la revisión del trabajo del metalurgista de planta, usted decide que las conclusiones deben ser más contundentes. Para ello usted planifica una campaña de experimentación durante dos semanas. En este período los dos reactivos son utilizados conjuntamente y con el mismo tipo de mineral en dos bancos de celdas. De este modo, las conclusiones que usted obtenga no estarán afectadas por las naturales fluctuaciones en la calidad del mineral. Los datos obtenidos son los siguientes;

Recuperación Total, %			
Día	Reactivo de Planta	“Float All”	$D = R_F - R_P$
1	87.9	87.3	-0.6
2	92.7	94.8	2.1
3	87.7	89.2	1.5
4	88.7	89.3	0.6
5	91.3	87.0	-4.3
6	90.0	89.0	-1.0
7	91.7	92.6	0.9
8	95.1	93.4	-1.7
9	84.7	83.1	-1.6
10	95.7	97.3	1.6
11	97.0	95.4	-1.6
12	83.0	86.2	3.2
13	88.6	89.1	0.5
14	83.8	84.5	0.7

Para el análisis de la información puede tratar las hipótesis del metalurgista de planta con 95% de confiabilidad:

A) El promedio de recuperación utilizando “Float All” es de 95 % (ó mayor).

B) La varianza de las recuperaciones con ambos colectores es igual.

C) El promedio de recuperación con ambos reactivos es igual

El análisis de los datos como muestras en pares a un nivel de significación del 95% será el procedimiento adecuado y en base a ello se deberán obtener las conclusiones con respecto a las ventajas de “Float All”.

EJERCICIOS GRUPO 4

1. Se desea contrastar la hipótesis μ = 0 frente a la alternativa μ >0 sobre la base de una muestra aleatoria de tamaño 9 obtenida de una población normal con $\sigma^2 = 1$ a un nivel de significación de 0.05. Verificar cual de las alternativas es la correcta. Rpta.: H_1 es cierta si $\bar{x}$ >0.55
2. Se desea contrastar la hipótesis μ = 0 frente a la alternativa μ >0 sobre la base de una muestra aleatoria de tamaño 100 obtenida de una población normal con $\sigma^2 = 1$. Si la hipótesis se debe rechazar cuando $\bar{x}$ > 0.233, cuál es el nivel de significación empleado? Rpta.: α = 0.01
3. Queremos contrastar la hipótesis μ = 0 frente a la alternativa $\mu \neq 0$ sobre la base de una muestra aleatoria de tamaño 25 obtenida de una población normal con $\sigma^2 = 1$. Si α = 0.05 hallar el conjunto de valores de x para los que la hipótesis debe aceptarse. Rpta.: [-0.392; 0.392]
4. La duración media de una muestra de 100 tubos fluorescentes producidos por una compañía resulta ser de 1,570 horas con una desviación estándar de 120 horas. Si el promedio histórico de los tubos es de 1,600. Verificar si la duración de los tubos es inferior a 1,600 horas a un nivel de significación de 0.05. Rpta.: La duración de focos es inferior.
5. Un diseñador de productos esta interesado en reducir el tiempo de secado de una pintura. Se prueban dos fórmulas de pintura; la fórmula 1 tiene el contenido químico estándar, y la fórmula 2 tiene un nuevo ingrediente secante que debe reducir el tiempo de secado. De la experiencia se sabe que la desviación estándar del tiempo de secado es de 8 minutos, y esta variabilidad inherente no debe verse afectada por la adición del nuevo ingrediente. Se pintan 10 especímenes con la fórmula 1 y otros 10 con la fórmula 2. Los dos tiempos de secado muestrales son 121 min y 112 min respectivamente. ¿A qué conclusiones puede llegar el diseñador del producto sobre la eficacia del nuevo ingrediente, utilizando α = 0.05?

   Rpta.: El tiempo de secado es menor.
6. Se requiere que la tensión de ruptura de un hilo utilizado en la fabricación de material de tapicería sea al menos 100 psi. La experiencia ha indicado que la desviación estándar de la tensión de ruptura es 2 psi. Se prueba una muestra aleatoria de 9 especímenes y la tensión promedio observada es de 98 psi.

   a) ¿Debe considerarse la fibra como aceptable con α = 0.05?.

   b) ¿Cuál es el valor *p* de esta prueba?

   Rpta.: (a)No, la resistencia es menor; (b) 0.00135
7. Se estudia el rendimiento de un proceso químico. De la experiencia previa con este proceso se sabe que la desviación estándar del rendimiento es 3. En los cinco días anteriores de operación de la planta, se han observado los siguientes rendimientos:

91.6%	88.75%	90.8%	89.95%	91.3%

   a) ¿Existe evidencia de que el rendimiento no es del 90% con α = 0.05?.

   b) Cuál es el valor *p* de esta prueba?

   Rpta.; (a) No; (b)0.3594]
8. El tiempo para reparar desperfectos de máquinas fotocopiadora tiene un promedio de 93 min. La compañía fabricante sostiene que con su nuevo modelo, aquellos tiempos serán inferiores. Para probar esto se evalúan 73 paradas por desperfectos y se observa un promedio de reparación de 88.8 minutos y una desviación estándar de 26.26 minutos. A un nivel de significación de 0.05 verificar la aseveración de la compañía.

   Rpta.: No hay diferencia.
9. Se tomaron aleatoriamente 64 fusibles y fueron sometidos a un 20% de sobrecarga y el tiempo de falla fue registrado. Para esta muestra se encontró $\bar{x}$ = 8.5 minutos y s = 2.4. A un nivel de significación del 0.05 probar si μ > 8.0.

   Rpta.: Si, μ> 8.0
10. En el pasado una máquina ha producido arandelas con un grosor de 0.05 pulgadas y una desviación estándar, σ, de 0.003 pulgadas. Para determinar si la máquina sigue en buenas condiciones de producción, se toma una muestra de 10 arandelas que resultan tener un grosor medio de 0.053 pulgadas. Ensayar la hipótesis de que la máquina este en buenas condiciones de producción a un nivel de significación del 0.05.
    Rpta.: , $\mu \neq 0.05$

11. Se analizan dos catalizadores para determinar la forma en que afectan el rendimiento promedio de un proceso químico. De manera específica, el catalizador 1 es el que se está empleando en este momento; pero el catalizador 2 también es aceptable. Debido a que el catalizador 2 es más económico, éste puede adoptarse siempre y cuando no cambie el rendimiento del proceso. Se hace una prueba en una planta piloto; los resultados obtenidos son:

Observación #	Catalizador 1	Catalizador 2
1	91.5	89.19
2	94.18	90.95
3	92.18	90.46
4	95.39	93.21
5	91.79	97.19
6	89.07	97.04
7	94.72	91.07
8	89.21	92.75

¿Existe alguna diferencia entre los rendimientos promedio de ambos catalizadores al 0.05 de significación?

Rpta.: No existe diferencia.

12. Diez individuos participan en un programa de modificación de dieta diseñado para estimular la pérdida de peso. En la siguiente tabla se indica el peso de cada participante antes y después de haber participado en el programa. ¿Existe evidencia que apoye la afirmación de que este programa de modificación de dieta es eficaz en reducir el peso? Utilice $\alpha = 0.05$

Sujeto	Antes	Después
1	195	187
2	213	195
3	247	221
4	201	190
5	187	175
6	210	197
7	215	199
8	246	221
9	294	278
10	310	285

Rpta.: Si es eficaz.

13. Se analiza una marca particular de margarina dietética para determinar el nivel de ácido graso poli-insaturado (en porcentaje). Se toma una muestra de seis paquetes y se obtiene los siguientes datos: 16.8; 17.2; 17.4; 16.9; 16.5; 17.1. Pruebe la hipótesis H_o $\mu = 17.0$ contra H_1: $\mu \neq 17.0$ con $\alpha = 0.01$. ¿Cuáles son sus conclusiones?

Rpta.: μ=170

14. Un ensayo sobre la resistencia a la rotura de 6 cuerdas fabricadas por una compañía mostró una resistencia media de 7,750 lb y una desviación estándar de 145 lb, mientras que el fabricante sostenía que la resistencia media de sus cuerdas era de 8,000 lb. Se puede admitir la afirmación del fabricante a un nivel de significación a) del 0.05; y b) del 0.01?

Rpta.: (a) μ=8000 lb.

15. Un laboratorio de pruebas desea contrastar si el promedio de vida de cierta herramienta de corte es de 2,000 piezas, frente a la alternativa de que es menos de 2,000. Qué conclusión se deberá obtener a un nivel de significación del 0.01, si 6 pruebas mostraron como vidas de las herramientas 2010, 1980, 1920, 2005, 1975, y 1950 piezas?

Rpta. : μ = 2000

16. Se afirma que la resistencia de un alambre eléctrico se puede reducir como mínimo en 0.05 ohms aleando el material. Se hacen 25 pruebas en alambre aleado y otras tantas en alambre sin aleación, dando los siguientes resultados:

	Media, ohms	Desviación estándar, ohms
Alambre aleado	0.089	0.003
Alambre sin aleación	0.141	0.002

Empleando un nivel de significación del 0.05 de determinar si esa afirmación es cierta.

Rpta.: $\mu_1 - \mu_2 > 0.05$

17. Una compañía desea comparar las vidas de dos piedras abrasivas y encuentra que el promedio de vida de 10 piedras de la primera clase es de 58 piezas con una desviación estándar de 6 piezas; y que la vida promedio de 12 piedras de la segunda clase es de 66 piezas con una desviación estándar de 4 piezas. Contrastar la hipótesis de que no hay diferencia significativa entre los promedios de vida de las dos piedras con un nivel de significación de 0.01.

    Rpta.: Si hay diferencia.

18. Dos productores de focos sostienen que sus productos son los de más larga duración. Se toma 100 focos de cada uno de estos fabricantes y son probados. Los resultados obtenidos son $\bar{x}$ = 798 y $\bar{y}$ = 826 horas respectivamente con $s_x^2 = 7982$ $s_y^2 = 9001$. Probar si estos dos tipos de focos son diferentes en cuanto su duración a un nivel de significación del 0.005.

    Rpta.: $\mu_1 = \mu_2$

19. Dos diferentes tipos de telas son comparados. Las perdidas de tela en miligramos en 7 pruebas son como sigue:

	P r u e b a						
Tela	1	2	3	4	5	6	7
A	36	26	31	38	28	37	22
B	39	27	35	42	31	39	21

Analizar los datos y definir si alguna de las telas es mejor que la otra.

Rpta. No son diferentes.

20. Un producto dietético líquido afirma en su publicidad que el empleo del mismo durante un mes produce una pérdida promedio de 3 libras de peso. Ocho sujetos utilizan el producto por un mes, y los datos sobre pérdida de peso son los siguientes:

	Sujeto							
	1	2	3	4	5	6	7	8
Peso inicial, lb	163	201	195	198	155	143	150	187
Peso final, lb	161	195	192	197	150	141	146	183

¿Los datos apoyan la afirmación hecha por la publicidad?

Rpta.: Sí.

21. Para evaluar el nivel de polución dentro de un hogar, se mide la cantidad de partículas sólidas suspendidas durante un período de 24 horas. Se toma muestras de hogares en donde no hay fumadores y en hogares en donde hay al menos un fumador. Siendo los resultados $n_1 = 16$ $\bar{x}$ = 67.1; $s_x^2 = 7.82$.; $n_2 = 13$ $\bar{y}$ = 132.3; $s_y^2 = 24.12$. Probar :

    H_o: $\sigma_1^2 = \sigma_2^2$ contra H_1: $\sigma_1^2 < \sigma_2^2$ con α = 0.05

    Rpta.: H_1 es cierta.

22. Un fabricante de detergente líquido esta interesado en la uniformidad de la máquina utilizada para llenar botellas. De manera específica, es deseable que la desviación estándar del proceso de llenado sea menor que 0.5 onzas de líquido; de otro modo, existe un porcentaje mayor del deseable de botellas con un contenido menor de detergente. La distribución del volumen de llenado es normal. Al tomar una muestra aleatoria de 20 botellas, se obtiene una varianza muestral de 0.0153 (onzas de fluido)^2.

    a) Definir el intervalo de confianza al 95% que contenga el verdadero valor de la desviación estándar.

b) Si la varianza de llenado es mayor que 0.01 (onzas de fluido)^2, entonces existe una proporción inaceptable de botellas que serán llenadas con una cantidad menor de líquido. Existe evidencia en los datos muestrales que sugiera que el fabricante tiene un problema con el llenado de las botellas? Utilice $\alpha = 0.05$.

Rpta.: (a) [0.1807; 0.094]; (b) $\sigma^2 = 0.01$

23. El contenido de azúcar del almíbar de los duraznos enlatados tiene una distribución normal, donde $\sigma^2 = 18$ (mg)^2 . Pruebe la hipótesis H_o $\sigma^2 = 18$ contra H_1: $\sigma^2 > 18$ si al tomar una muestra de 10 latas la desviación estándar muestral es $s = 4.8$ mg con $\alpha = 0.05$

Rpta.: $\sigma^2 = 18$

24. Un ingeniero que trabaja para un fabricante de llantas investiga la duración promedio de un compuesto nuevo de caucho. Para ello construye 16 llantas y las prueba en una carretera hasta alcanzar el fin de la vida útil de éstas. Los datos en Km. obtenidos son los siguientes:

60,613	59,836	59,554	60,252
59,784	60,221	60,311	50,040
60,545	60,257	60,000	59,997
69,947	60,135	60,220	60,523

¿Puede concluirse que la desviación estándar de la duración de las llantas excede los 200 Km con $\alpha = 0.05$?

Rpta.: $\sigma^2 > 200$.

25. Dos compañías de compuestos químicos pueden surtir materia prima. La concentración de un elemento en particular es importante. La concentración promedio ambos proveedores es la misma, pero se sospecha que la variabilidad en la concentración puede diferir entre las dos compañías. La desviación estándar de la concentración en una muestra aleatoria de $n_1 = 15$ lotes producidos por la compañía 1 es $s_1 = 4.7$ g/l, mientras que para la compañía 2 , una muestra aleatoria $n_2 = 20$ lotes proporciona una $s_2 = 5.8$ g/l. ¿Existe evidencia suficiente para concluir que las varianzas de las dos poblaciones son diferentes? Utilice $\alpha = 0.05$.

Rpta.: $\sigma_1^2 = \sigma_2^2$