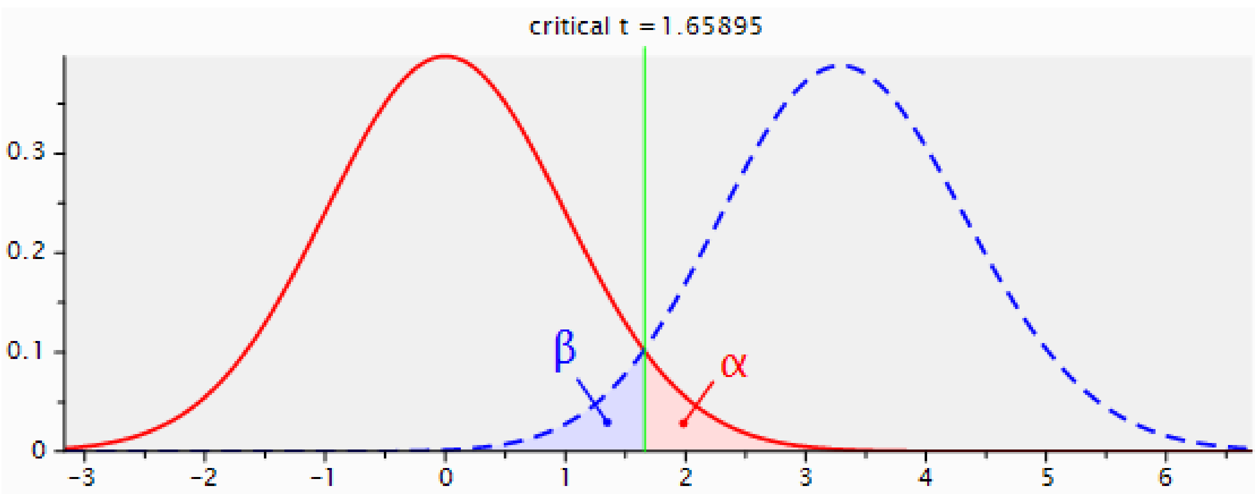
PRUEBA DE HIPÓTESIS

Rechazar o no Ho: he ahí el dilema



JORGE FALLAS

2012

Cuando las leyes de la matemática se refieren a la realidad, no son ciertas; cuando son ciertas, no se refieren a la realidad. - Albert Einstein (1879 - 1955)

Índice

1. Introducción 1
2. Datos experimentales y distribuciones de frecuencia teóricas 1
3. Hipótesis nula, alternativa, nivel de significancia y error tipo I y II 2
3.1. Error tipo I, II y potencia de la prueba 4
3.2. Prueba de hipótesis direccionada y no direccionada 6
4. Comparación de medias 10
4.1. Comparación de una media con varianza desconocida y muestra pequeña 10
4.2. Efecto de la variabilidad de los datos en la prueba de hipótesis 20
4.3. Comparación de dos medias con varianzas desconocidas y muestras pequeñas 20
4.3.1. La prueba t de Estudiante para dos medias independientes 21
4.3.2. Datos no normales o simétricos: ¿qué hacer? 34
4.3.3. Prueba de hipótesis de una cola o unilateral 36
4.4. Medias pareadas o dependientes 38
4.5. Bibliografía..........43
4.6. Ejercicios..........45
Anexo 2: Prueba de hipótesis: flujograma 48
Anexo 3: Guía para prueba de hipótesis..........49
Anexo 4: Fórmulas..........52

1. Introducción

Las investigaciones son diseñadas para responder a múltiples preguntas; sin embargo en el presente capítulo nos concentraremos en una de las preguntas más simples: ¿existe una diferencia estadísticamente significativa entre un estimador y el parámetro de la población o entre dos estimadores? Para responder a esta pregunta recurrimos a la prueba de hipótesis o contrastes.

Al someter a prueba una hipótesis determinamos si dos valores numéricos--obtenidos de un diseño estadísticamente válido--son diferentes a un nivel de significancia dado. Por ejemplo, podemos preguntarnos ¿es la precipitación media anual en Cartago diferente a la precipitación media de Liberia para el periodo 1950-1980? o ¿es el tratamiento pregerminativo "A" superior al "B"? En estos casos, el objetivo del estudio es estimar las diferencias y su error, para luego determinar si existe una diferencial estadística entre las variables medidas. Con frecuencia a este tipo de estudios se les denomina *comparativo* ya que involucran sólo dos grupos experimentales.

En el presente capítulo se trata el tema de prueba de significancia, prueba de hipótesis o contrastes y se retoma del tema de intervalo de confianza y nivel de significancia. Estos conceptos se utilizarán en los siguientes capítulos para analizar datos provenientes de diseños más complejos.

A continuación se presentan los supuestos y procedimientos estadísticos utilizados para comparar una media con un estándar o norma, dos medias independientes y dos medias dependientes. A lo largo del capítulo usted se familiarizará con términos como distribución de referencia, distribución de probabilidad, prueba de significancia, distribución de t, normal y F.

2. Datos experimentales y distribuciones de frecuencia teóricas

Al analizar un set de datos la primer interrogante a la que se enfrenta el o la investigadora es ¿cómo saber si los valores son muy grandes, muy pequeños ó promedios? Por ejemplo, ¿cómo saber si la producción media anual de frutos de un parque de bosque es excepcional, normal o escasa? Para responder a esta pregunta se necesita un patrón o valor de referencia contra el cual se pueda comparar el set de datos. En el mundo de la estadística se conoce a dicho set de referencia como la "*distribución esperada de la variables en estudio*", e indica cuales valores podría tomar la misma. En una prueba de hipótesis se declaran como *significativas* aquellas *diferencias que son excepcionalmente grandes ó pequeñas* con respecto a la distribución esperada de las diferencias para la variable en estudio. Por ejemplo, la producción histórica de frutos (esperada) para el parche de bosque es 100kg± 3kg y en el año "X" registró 90kg± 5kg: ¿es la diferencia significativa?

En un diseño experimental el investigador(a) aplica un tratamiento a un grupo de sujetos experimentales y luego mide la respuesta de dichos sujetos al tratamiento. Una vez que obtiene los datos debe compararlos con un set de referencia (control) o sea aquellos valores que se obtendrían en ausencia del tratamiento. Una vez obtenida la diferencia debe decidir si la misma es muy grande ó muy pequeña (*estadísticamente significativa*) o si por el contrario las diferencias pueden atribuirse al *efecto del azar o ruido*. Un aspecto importante a la hora de seleccionar la distribución de referencia es que sea relevante para el caso en estudio. Por ejemplo, si estamos estudiando el crecimiento de un bosque de pochote no tiene mucho sentido compararlo con el crecimiento de plantaciones de cedro macho. Esto nos lleva al tema de la población de referencia o sea aquella a la cual esperamos aplicar los resultados (*validez externa del estudio*).

La figura 1 ilustra la distribución de frecuencia para 1000 observaciones obtenidas de una población normal con media 17.4 cm y varianza de 25 cm^2. Cualquier observación con un valor *superior a 27.5 cm* ó *inferior a 7.5 cm* se ubica a ± dos desviaciones estándares de la media y por tanto podría considerarse como un valor raro o poco frecuente comparado con la distribución de referencia, observe que dichos valores se ubican en las colas de la distribución.

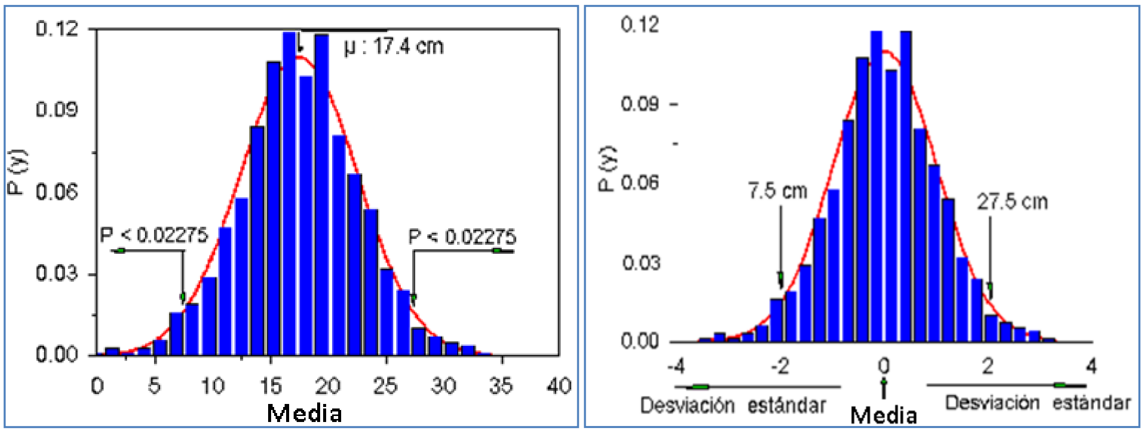


Figura 1: Concepto de distribución de frecuencia y desviación estándar de la media. Estas gráficas ilustran la distribución de frecuencia esperada para una población con un diámetro medio de 17,4 cm y una varianza de 2,5 cm^2. Cualquier muestra con un diámetro medio superior a 27,5 cm ó inferior a 7,5 cm se consideraría un evento raro dado la distribución de referencia.

3. Hipótesis nula, alternativa, nivel de significancia y error tipo I y II

La hipótesis nula (Ho) es la que se somete a prueba y sobre ella se hace la decisión. Para los propósitos de la prueba se asume como verdadera y se rechaza ó no se rechaza como resultado del proceso de análisis. En la vida cotidiana, la pregunta o razón por la cual se hace la prueba de hipótesis está más relacionada con la hipótesis alternativa (Ha) que con la nula.

Por ejemplo, si estamos interesados en saber si un <u>*nuevo tratamiento pregerminativo es mejor*</u> que el utilizado actualmente, la hipótesis nula se plantea en términos de no diferencia entre el método actual y el nuevo. Por su parte, la hipótesis alternativa se plantearía de tal forma que indique que el <u>*nuevo método es mejor que el utilizado actualmente*</u>.

Ho es sometida a prueba en lugar de la hipótesis alternativa (Ha) porque la serie estadística provee la información necesaria para estimar los parámetros de su distribución muestral; en tanto que Ha no ofrece esta ventaja. Por ejemplo, si sometemos a prueba la hipótesis Ho: μ=0,87 gr/cm^3, asumimos que la distribución muestral de las medias está centrada en el valor 0,87. Conociendo esto podemos determinar si la media muestral corresponde o no a dicha distribución y además si el valor de la media muestral es suficientemente raro (muy grande o muy pequeño) y por lo tanto poco probable como para que deba rechazarse la hipótesis nula.

Por otra parte si intentamos probar directamente la hipótesis alternativa Ha:μ <> 0.87, nos encontraríamos con el inconveniente de que no sabríamos donde se centra la distribución muestral de las medias, lo único que podríamos afirmar es que no se centra en el valor 0.87. Lo anterior imposibilita someter a prueba la hipótesis alternativa y a la vez justifica la necesidad de probar la validez de la hipótesis nula.

Los términos *no rechazar* y *rechazar* sólo deben utilizarse cuando nos referimos a la hipótesis nula; pues ésta es la que sometemos a prueba. El hecho de rechazar Ho significa que los datos muestrales brindan suficiente evidencia como para pensar que lo planteado por la hipótesis nula es estadísticamente improbable a un nivel de significancia dado. De la misma manera cuando no rechazamos Ho significa que los datos muestrales no brindan suficiente evidencia como para pensar que lo planteado por la hipótesis nula sea improbable a un nivel de significancia dado. Al analizar los resultados de una prueba de hipótesis siempre debe considerarse el *efecto de confusión* derivado de la presencia de variables no consideradas en el diseño original y que no se están sometiendo a prueba.

El *nivel de significancia* se designa con la letra griega α e indica cuan rara (muy grande o muy pequeña) deber ser la diferencia con respecto a lo planteado por la hipótesis nula como para que sea rechazada *dado que sea correcta* (Fig. 2). Por ejemplo, si el volumen medio por hectárea de un bosque es 200 m^3/ha (Ho: μ = 200 m^3/ha) en cuánto debe diferir el volumen/ha de una muestra para que se considere diferente de 200 m^3/ha.

Significancia estadística: Esta es una regla que permite afirmar que la diferencia observada entre dos o más sets de datos es el resultado del efecto del “tratamiento” y no del azar. Con frecuencia se declaran como significativas aquellas diferencias con una probabilidad inferior a 0,05 (o sea 5%) de observarse en forma aleatoria. En algunos textos de estadística se recomienda utilizar un asterisco (*) para designar diferencias significativas a un 5% ($P < 0.05$), dos asteriscos (**) para designar diferencias significativas al 1% ($P < 0.01$) y tres asteriscos (***) para designar diferencias significativas al 0,1% ($P<0.001$). Sin embargo, dado que los paquetes estadísticos le brindan el valor de “p” se recomienda reportar dicho valor con los respectivos grados de libertad e indicar.

Es común observar en documentos científicos valores de α entre 0,05 y 0,001; al primero se le denomina *diferencia significativa* y al segundo *diferencia altamente significativa*. Cuando se rechaza Ho se utiliza el término "significativo" o "estadísticamente significativo" y el resultado puede interpretarse en el sentido de "haber aprendido algo nuevo sobre la población". Por otro lado, el término "no significativo" expresa el sentir de que la prueba no aportó nuevo conocimiento sobre la población en estudio. Al realizar una prueba de hipótesis debemos recordar que *muestras grandes tienden a generar diferencias significativas* aún cuando las diferencias sean *prácticamente insignificantes*.

3.1. Error tipo I, II y potencia de la prueba

El valor de α indica del riesgo a equivocarse que el investigador(a) asume cuando evalúa Ho. Por ejemplo, para un α de 0,05 y dado que Ho sea verdadera se espera que la misma sea rechazada, por razones de azar, un 5% de las veces que se ejecute la prueba. O sea, en *5 de cada 100 contrastes Ho será rechazada aún cuando sea verdadera* (recodemos que Ho se plantea siempre en términos de no diferencia). A este tipo de error se le denomina *error tipo I* y crea un falso positivo (declara algo como cierto cuando en realidad no lo es). Por otro lado, el no rechazar Ho cuando en realidad debe rechazarse se denomina error *tipo II* y se designa con la letra β. En este caso se crea un falso negativo al no declarar algo como cierto cuando en realidad lo es. (Fig. 2). La *probabilidad de un error de tipo II* depende de la media de la población que es desconocida; sin embargo se puede calcular para valores dados de σ^2, μ, y. n.

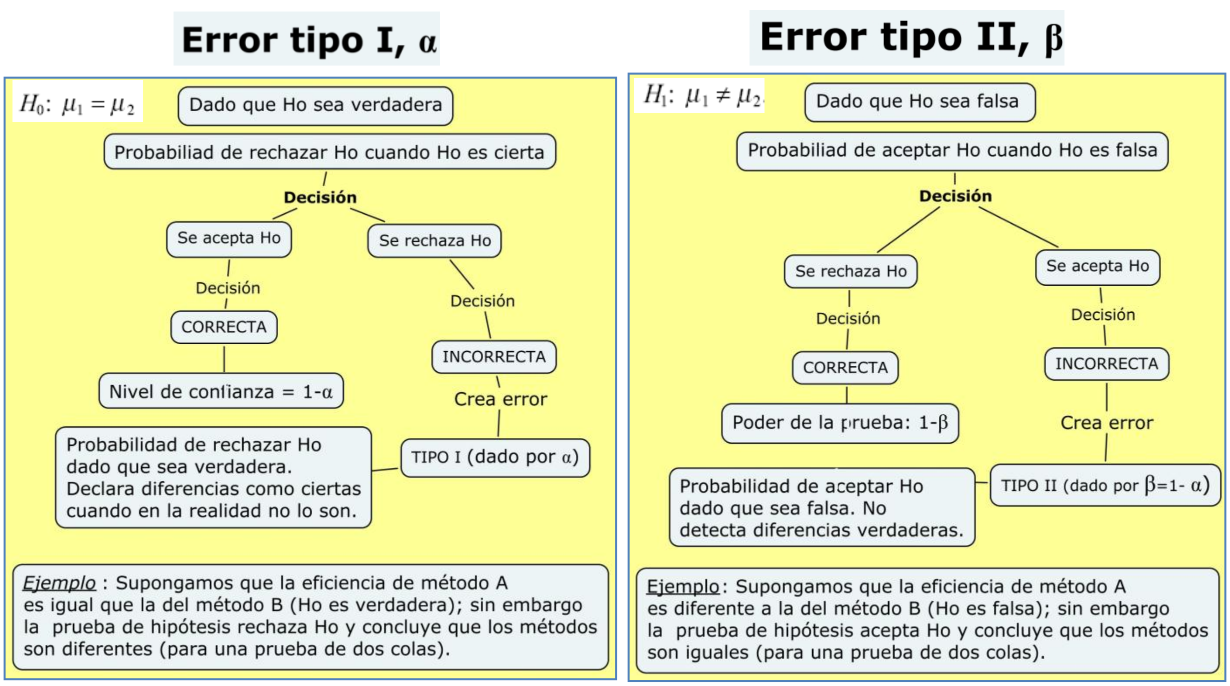
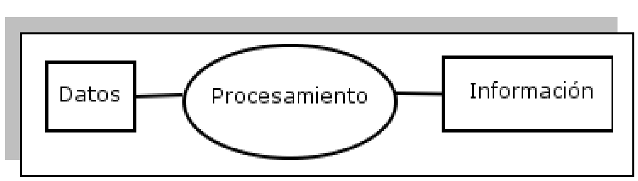


Figura 2: Error tipo I y II. . H_0 (hipótesis nula) H_1 (hipótesis alternativa).



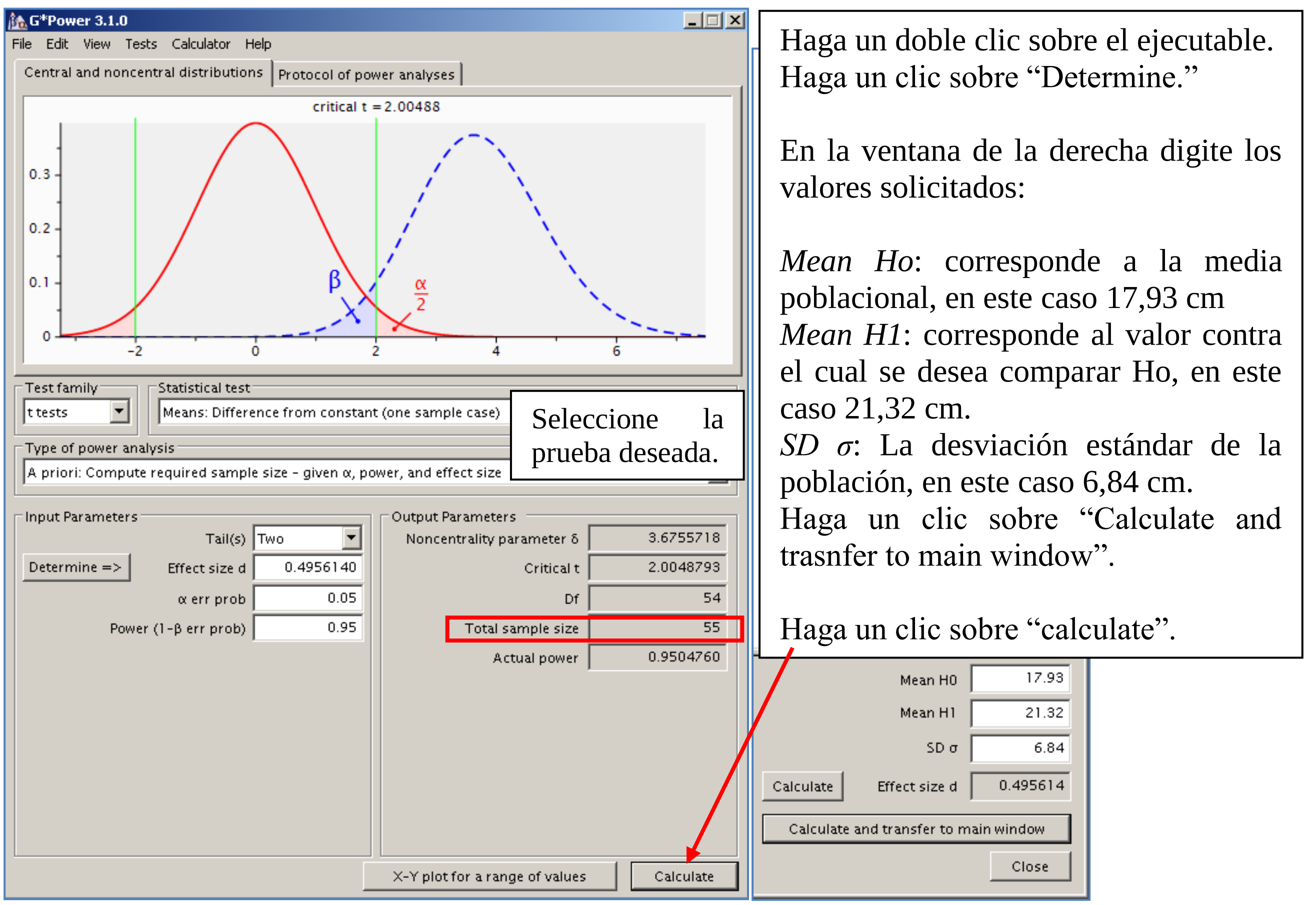
Fuente: http://es.wikipedia.org/wiki/Dato

La *potencia de una prueba de hipótesis* es igual al uno menos la probabilidad del error de tipo II (1-β) e indica la probabilidad de que se rechace Ho cuando la misma es falsa (es decir, se tome la decisión correcta). A continuación se ilustra el concepto utilizando XLStatitiscs y el archivo diámetros.xlsx.

Parámetros (corresponden a los 100 datos).
Desviación estándar (σ): 6,84 cm
Media poblacional (μ): 17,93 cm

Deseamos determinar el poder de una prueba hipótesis para rechazar Ho dado que sea falsa para una muestra con una media de 21,32 cm. Para realizar los cálculos utilizaremos los programas gratuitos Power (http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3) y XLStatistics (http://www.deakin.edu.au/~rodneyc/XLStatistics/).

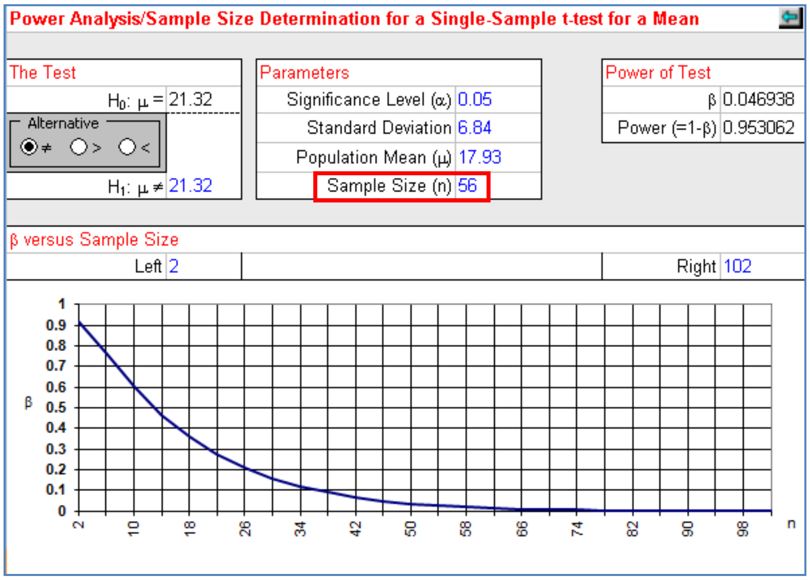
Power
Configure la ventana de diálogo como se muestra a continuación:



Mean Ho: corresponde a la media poblacional, en este caso 17,93 cm
Mean H1: corresponde al valor contra el cual se desea comparar Ho, en este caso 21,32 cm.
SD σ: La desviación estándar de la población, en este caso 6,84 cm.

El programa indica que la potencia de la prueba sería 0,95 con un tamaño de muestra de 55.

XLStatistics Para los mismos datos XLStatistics recomienda una muestra de 56 observaciones.

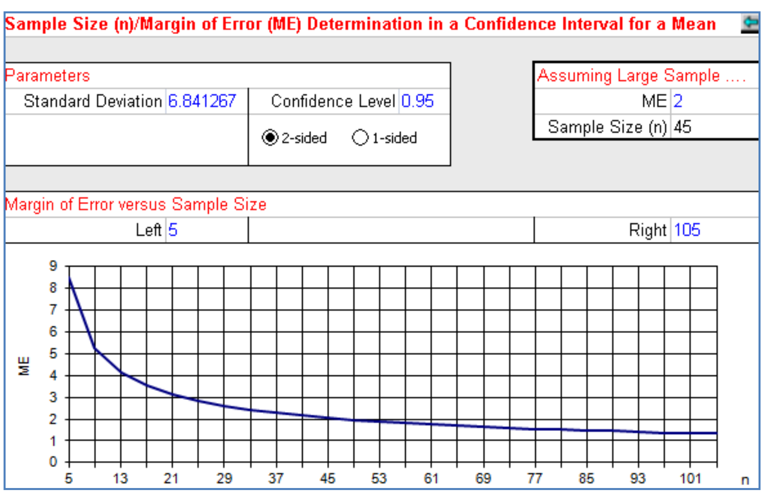


La gráfica permite apreciar que:

1. El error tipo II (β) se reduce rápidamente al aumentar el tamaño de muestra (n).
2. A partir de un tamaño de muestra 50, la ganancia en reducción del error tipo II es insignificante y por lo tanto adicionar más muestras retornan muy poco beneficio.

Error marginal y tamaño de muestra

El error marginal es otro término asociado a las estimaciones basadas en muestras e indica cuál esl tamaño de muestra requerido para logar un error marginal dado un intervalo de confianza y un valor de desviación estándar. Por ejemplo, para los 100 datos de diámetro del archivo diámetros.xlsx se requiere un tamaño de muestra de 45 para lograr un error marginal de ± 2cm alrededor de la media con una confianza de 95%.



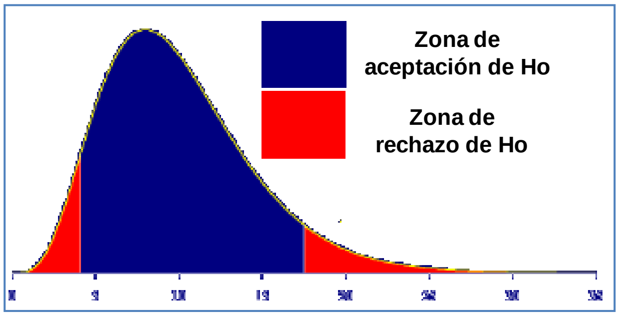
Como ejercicio se recomienda que extraiga varias muestras de tamaño 45 de los 100 datos de diámetro, calcule los IC al 95 y compruebe lo indicado por XLStatistics.

3.2. Prueba de hipótesis direccionada y no direccionada

Al plantear una prueba de hipótesis, usted debe decidir si desea realizar una prueba de dos colas o bilateral o por el contrario desea hacer una prueba de una cola (superior o inferior). A continuación se explica en qué consiste cada una de ellas.

Prueba de hipótesis no direccionada, de dos colas o bilateral

En este caso interesa determinar si el valor del estimador (muestra) es diferente al valor del parámetro (e.g. μ); sin importar si es mayor o menor. Este tipo de prueba se denomina también de *dos colas o bilateral*; ya que Ho se rechazará si el valor del estadístico de prueba se ubica *en cualquiera de las zonas de rechazo de* Ho como se muestra a continuación.



En una prueba de hipótesis no direccionada solo nos interesa saber si el valor muestral no se ubica en la zona de aceptación de Ho; lo cual permite concluir que es diferente al parámetro.

Por ejemplo, si existe una norma nacional en cuanto al rendimiento medio por hectárea de un cultivo, podemos comparar el crecimiento de una muestra (parcela) con dicha norma para determinar si es igual o diferente. En la práctica esta prueba de hipótesis es poco útil pues si rechazamos H_0 lo único que se puede afirmar es que la muestra es diferente al parámetro y que por lo tanto pertenece a otra población.

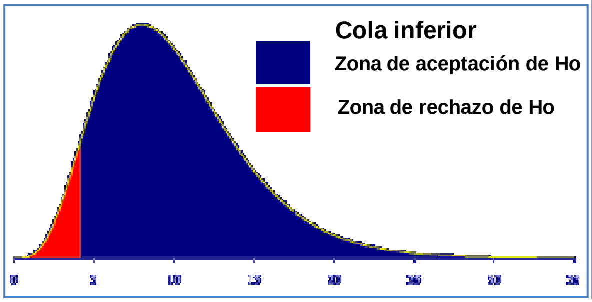
Prueba de hipótesis de una población	**Prueba de hipótesis de dos poblaciones**
H_0: estipula el valor del parámetro (e.g. μ=15) H_A: el estimador es diferente al parámetro	H_0: estipula el valor de los parámetros de las poblaciones (e.g. $\mu_1 = \mu_2$). H_A: los estimadores son diferentes

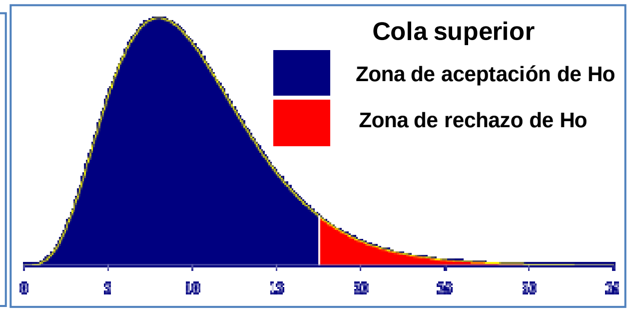
Un error frecuente en la interpretación de los resultados de esta prueba de hipótesis es el siguiente:

Suponga que usted somete a prueba la hipótesis "el rendimiento de mi parcela es igual a la norma nacional, e.g. μ=15" utilizando una prueba de dos colas. Una vez realizados los cálculos, H_0 es rechaza, pero al observar el valor del estimador usted se da cuenta de que es mayor que el parámetro y por lo tanto concluye: "la parcela muestra un rendimiento superior a la norma nacional". Aunque numéricamente su observación es correcta, esa no fue la hipótesis que usted sometió a prueba y por lo tanto lo único que puede afirmar es que es diferente.

Prueba de hipótesis direccionada, de una cola o unilateral

En este caso, a diferencia del anterior, sí estamos interesados en determinar si el valor del estimador (muestra) *es mayor* o *menor* que el valor del parámetro (e.g. μ). Por esta razón este tipo de prueba se denomina también de *una cola o unilateral*; ya que Ho se rechazará si el valor del estadístico de prueba se ubica en la cola inferior o superior de la distribución muestral como se ilustra a continuación:





En una prueba de hipótesis direccionada nos interesa saber si el valor muestral se ubica en la cola superior o inferior de la distribución y no solo si es diferente del paramétro.

Por ejemplo, si existe una norma nacional en cuanto al rendimiento medio por hectárea para un cultivo "X", podemos comparar el crecimiento de una muestra (parcela) con dicha norma para determinar si es *mayor* o *menor*. En la vida real, esta prueba de hipótesis se utiliza con mayor frecuencia ya que si rechazamos H_0 podemos afirmar es que la muestra es mayor o menor que el parámetro y no solo que es diferente. El planteamiento de la prueba de hipótesis puede hacerse de la siguiente manera:

Caso	**Prueba de hipótesis de una población**
1	H_0: estipula el valor del parámetro (e.g. μ=15) H_A: el estimador es ***mayor*** al parámetro
2	H_0: estipula el valor del parámetro (e.g. μ=15) H_A: el estimador es ***menor*** al parámetro

Caso	**Prueba de hipótesis de dos poblaciones**
1	H_0: estipula el valor de los parámetros de las poblaciones (e.g. $\mu_1 = \mu_2$). H_A: un estimador es mayor que el otro (e.g. $\mu_1 > \mu_2$)
2	H_0: estipula el valor de los parámetros de las poblaciones (e.g. $\mu_1 = \mu_2$). H_A: un estimador es menor que el otro (e.g. $\mu_1 < \mu_2$)

Observe que Ho es igual para ambos casos; sin embargo H_A nos lleva a conclusiones opuestas. En el primer caso podemos afirmar que el estimador es mayor que el parámetro y en el segundo que es menor. Dado que esto tiene importantes implicaciones prácticas usted debe decidir cuál caso someterá a prueba antes de iniciar la toma de datos.

Prueba de una cola: ¿Cómo decidir cuál utilizar?

Una de las preguntas más frecuentes al utilizar una prueba de hipótesis de una cola o unilateral es ¿cuál debe ser la hipótesis alternativa? Observe que la hipótesis nula siempre se plantea en términos de no diferencia; por ejemplo: el estimador es igual al parámetro o no existe diferencia entre dos tratamientos o muestras.

La decisión sobre la hipótesis alternativa debe guiarse por la pregunta que usted desea responder. Veamos un ejemplo. Usted está interesado(a) en saber si una nueva técnica para determinar oxígeno disuelto *es mejor* que la utilizada actualmente. Antes de realizar las mediciones usted debe plantear su hipótesis nula y alternativa; veamos las opciones disponibles:

Hipótesis	**Inferencia (conclusión)**
Plantea una hipótesis de dos colas	Si rechaza H_0 se puede afirmar que el nuevo método es diferente al anterior y por lo tanto no responde a la pregunta planteada en el estudio.
Plantea una hipótesis de una cola pero utiliza la cola inferior.	Si rechaza H_0 se puede afirmar que el nuevo método es inferior al anterior y por lo tanto tampoco responde a la pregunta planteada en el estudio.
Plantea una hipótesis de una cola pero utiliza la cola superior.	Si rechaza H_0 se puede afirmar que el nuevo método es superior al anterior y por lo tanto si responde a la pregunta planteada en el estudio.

¿Qué significa rechazarH_0?

La prueba de hipótesis es uno de los pilares de la investigación aplicada tanto en ciencias naturales como sociales. A diferencia de las ciencias exactas, en estos campos no existen leyes físicas que permitan establecer relaciones de causa-efecto y por lo tanto la prueba de hipótesis es una herramienta valiosa para lidiar con el efecto del azar. Pero ¿qué significa en la práctica “rechazar H_O?. Con frecuencia se acepta como una afirmación absoluta cuando en realidad es una afirmación acotada. Veamos por qué.

Asumiendo que se elige la prueba correcta y que se cumple con los supuestos de la misma, el rechazar H_0 depende de los siguientes factores:

1. De que efectivamente sea falsa. Si H_0 es verdadera no debe rechazarse.
2. El nivel de significancia elegido por el investigador(a). Dado por alfa (α).
3. La potencia de la prueba estadística utilizada (capacidad para rechazar H_0 dado que sea falsa).
4. El tamaño de la muestra utilizada para estimar el valor del parámetro a someter a prueba (e.g. la media). La muestra provee la evidencia contra la cual se somete a prueba lo planteado en la hipótesis nula. Por esta razón es importante utilizar un método valido para elegirla y técnicas de medición apropiadas.
5. La variabilidad natural de la población.

De los factores anteriores el investigador(a) puede decidir sobre el valor de alfa, el tamaño de la muestra (aunque en la práctica no siempre sea cierto) y reducir la variabilidad de la población restringiéndola. La decisión de aceptar o rechazar Ho siempre incluye un margen de error y lo que la estadística le indica es, dada ciertas condiciones, ¿cuál es dicho margen de error?

4. Comparación de medias

Cuando nos referimos a pruebas de hipótesis o contrastes sobre medias podemos estar interesados en analizar los siguientes escenarios:

1. ***Comparar una media con un valor específico***. Por ejemplo ¿es la temperatura media para el mes de marzo del 2000 igual a 30 ^{0}C?

2. ***Comparar medias obtenidas en forma independiente de dos poblaciones.*** Por ejemplo, dado un ensayo de especies nativas podemos preguntarnos ¿es el crecimiento medio en altura de laurel superior al de roble de sabana?

3. ***Comparar tres o más medias***. Por ejemplo, dado un experimento de inoculación de micorriza podemos comparar la eficiencia de los tratamientos A, B y C. En este caso la prueba se denomina análisis de varianza y se discutirá en otra sesión del curso.

En todos los casos, las pruebas pueden ser de una cola o de dos colas. En el primer caso nos interesa saber si la media es mayor o menor que un valor dado y en el segundo si es diferente.

4.1. Comparación de una media con varianza desconocida y muestra pequeña[1]

Esta es la prueba de hipótesis más simple y pretende responder a la siguiente pregunta:

¿Es el valor del estimador igual a la media poblacional μ?

Si asumimos que la media proviene de una población con una distribución normal y que la muestra se obtuvo en forma aleatoria se puede utilizar el estadístico **t** para someter a prueba esta hipótesis. Dicho estadístico tiene una distribución *t de Estudiante* con n-1 grados de libertad.

$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}$ donde S = desviación estándar, $\bar{X}$ = media muestral y n: tamaño de muestra.

Por ejemplo, si estamos estudiando la densidad de la madera de roble podemos utilizar el valor reportado en la literatura como media (μ) y comparar el valor de una muestra obtenida en Talamanca. Esto nos permite determinar cuán similar o disímil es la densidad de los robledales de dicha zona con respecto al valor poblacional o de referencia.

Dada una distribución muestral para la densidad media de roble como la observada en la figura 2, fácilmente podríamos concluir que una muestra con una media de 0,41 gr/cm^3 y una desviación estándar de 0.041 gr/cm^3 es diferente al valor poblacional (μ= 0,87 gr/cm^3) (Figs. 3); sin embargo no podríamos afirmar lo mismo para una muestra con un valor de 0,85 gr/cm^3 (Figs. 3).

---

[1] En los textos de estadística usted encontrará que también existe una prueba de hipótesis para la media de una población

Extendiendo este razonamiento a una prueba de hipótesis debemos preguntarnos cuan raro o alejado de μ debe estar la media muestral como para concluir que es diferente y por ende declarar la muestra como diferente de la población de referencia.

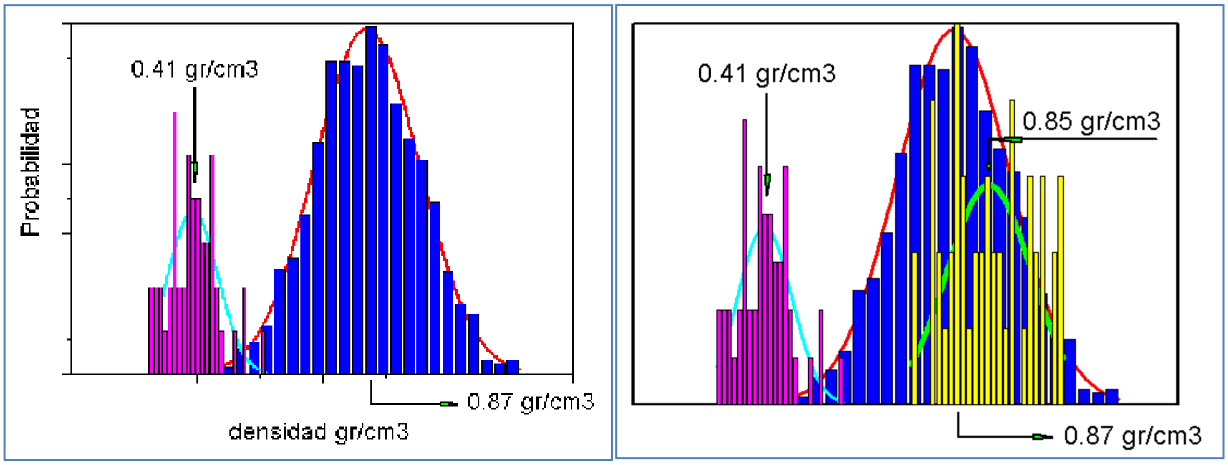


Figura 3: A. Distribución de frecuencia para una población normal con μ= 0,87 gr/cm3 y σ =0.087 gr/cm3 y una muestra con una media de 0.41 gr/cm3 y una desviación estándar de 0.041 gr/cm3. B. Distribución de frecuencia para: A) población normal con μ = 0,87 gr/cm3 y σ =0.087 gr/cm3. B) muestra con media de 0.41 gr/cm3 y desviación estándar de 0.041 gr/cm3. C. muestra con media de 0.85 gr/cm3 y desviación estándar de 0.085 gr/cm3.

EJEMPLO

Suponga que en los suelos fértiles y bien drenados de la Zona Norte el diámetro medio del bosque es de 15 cm a la altura del pecho en quince años (μ= 15 cm). Un inversionista desea comprar una finca con varios parches de bosque y desea saber si el crecimiento del bosque es comparable con el de los mejores sitios de la Zona Norte. Usted selecciona al azar una parcela con cien árboles en un parche de bosque de quince años ubicado en la finca y obtiene un diámetro medio de 14,23 cm con una desviación estándar de 1,26cm.

La pregunta que nos interesa responder es ¿muestran los datos de la parcela suficiente evidencia como para afirmar que el crecimiento del bosque en la finca es igual al de los mejores sitios de la Zona Norte del país? Del planteamiento de la pregunta se concluye que interesa saber si el crecimiento del bosque en la finca es igual al de los mejores sitios de la Zona Norte y por esta razón se plantea una prueba de dos colas o bilateral.

Ho: μ_1 = 15 cm
Ha: $\mu_1 \neq$ 15 cm

Para realizar esta prueba de hipótesis se recomienda seguir el siguiente procedimiento:

1. *Cálculo de estadísticos descriptivos*

   Los estadísticos descriptivos resumen lo relevante de los datos en términos de tendencia central, variabilidad y forma de la distribución. Normalmente se calcula la media, desviación estándar, coeficiente de variación, error estándar y los coeficientes de asimetría y curtosis.

2. *Análisis gráfico*

   El objetivo del análisis gráfico es detectar patrones o tendencias en el set de datos. Por ejemplo, se puede analizar la tendencia central, la variabilidad y la forma de la distribución que caracteriza al set de datos. Los gráficos de Box-Whisker y de barra de errores (desviación estándar, error estándar, intervalo de confianza) son apropiados para visualizar el comportamiento de dos o más sets de datos. Cuando se desea evaluar la normalidad de los datos puede utilizarse un histograma o un diagrama de probabilidad normal.

3. *Prueba de hipótesis*

   Una vez que usted se ha familiarizado con el set de datos puede proceder a realizar la prueba de hipótesis. El proceso involucra los siguientes pasos:

A. Plantear la hipótesis nula y alternativa.
B. Seleccionar el estadístico de prueba y definir el nivel de significancia. Para efectuar la prueba de hipótesis puede optar por un estadístico parámetrico o por uno no paramétrico. Toda prueba paramétrica requiere que los datos sean normales y por lo tanto antes de aplicar la prueba debe realizar un prueba para probar por la normalidad del set de datos. Con frecuencia las pruebas paramétricas son preferidas sobre las no paramétricas porque son más eficientes o sea requieren de un menor tamaño de muestra para decidir sobre H_0 con respecto al equivalente no paramétrico.
C. Efectuar la prueba de hipótesis.
D. Tomar una decisión
E. Proponer a una conclusión o explicación.

Uso de InfoStat

1. Cálculo de estadísticos descriptivos

Algunos aspectos relevantes del set de datos son: su media es 17,93 cm y su coeficiente de variación es 38,2%, por cuanto el set de datos puede considerarse como muy variable. Los valores de asimetría y curtosis: 1,13 y 1, 82, respectivamente, indican que los datos son no normales y con una cola hacia la derecha.

Resumen	d (cm)
n	100.00
Media	17.93
D.E.	6.84
Var(n-1)	46.80
E.E.	0.68
CV	38.16
Mín	6.80
Máx	40.00
Mediana	17.50
Q1	13.70
Q3	21.40
Asimetría	1.13
Kurtosis	1.82

2. ***Análisis gráfico***

El gráfico de probabilidad normal y el histograma indican que la variable diámetro (cm) no proviene de una población con una distribución normal. El set de datos se aparta de la curva de normalidad en sus valores extremos (pequeños y grandes) y además presenta una distribución con asimetría positiva.

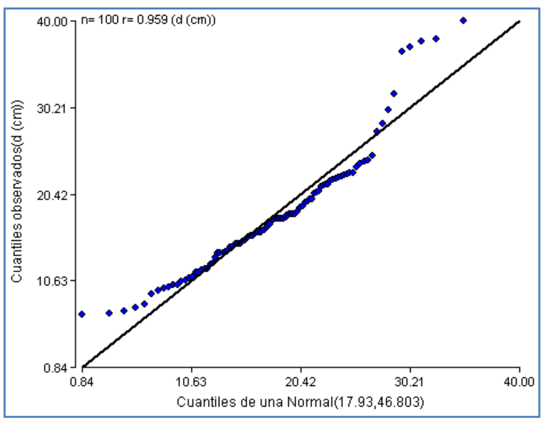
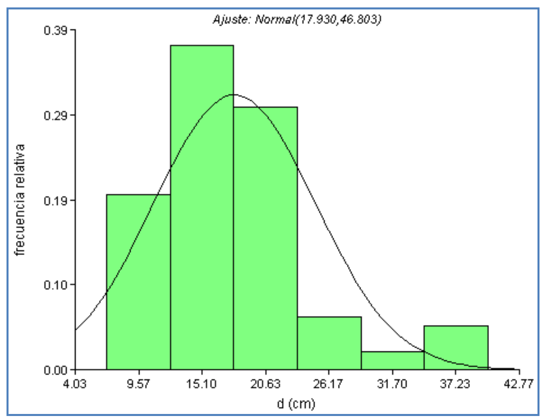


Gráfico de probabilidad normal



Histograma

3. ***Planteamiento de la prueba de hipótesis***

H_O: μ_0 = 15 cm (La media de la población es igual a 15cm)

H_A: $\mu_0 \neq$ 15 cm (La media de la población es diferente a 15cm)

Supuestos: normalidad, muestra aleatoria

NOTA: Recuerde que la decisión de realizar una prueba de una ó dos colas debe hacerse antes de colectar los datos. De lo contrario usted estaría sesgando su decisión.

4. Seleccionar el estadístico de prueba y definir el nivel de significancia

Se utilizará el estadístico t de estudiante y un nivel de significancia α= 0,01 (1%).

Estadístico de prueba

$t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$ donde S = desviación estándar, $\bar{X}$ = media muestral, n= tamaño de muestra y μ_0 la media poblacional. El estadístico "t" tiene una distribución "t" de Estudiante con n-1 grados de libertad.

5. Prueba de normalidad

Uno de los supuestos de la prueba "t" es que los datos provienen de una población con una distribución normal.

H_0: Los datos provienen de una distribución normal
H_A: Los datos no provienen de una distribución normal

InFostat utiliza la prueba W de Shapiro-Wilk (Shapiro 1965) para determinar la normalidad del set de datos. Dicho estadístico es apropiado tanto para muestras pequeñas (i.e. <50 observaciones) como para muestras grandes (e.g 1000 o más observaciones). Se le considera como una de las pruebas más poderosas para probar por la normalidad de un set de datos.
El estadístico de prueba es:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

Donde,

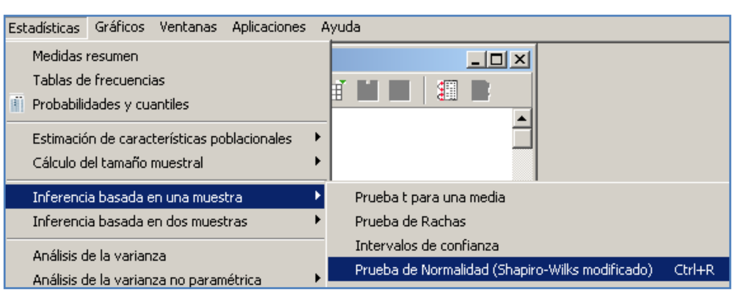
- $x_{(i)}$ es el número que ocupa la *i-ésima* posición en la muestra
- $\overline{x} = (x_1 + \ldots + x_n)/n$ es la media muestral
- las constantes a_i se calculan con:

$$(a_1, \ldots, a_n) = \frac{m^\top V^{-1}}{(m^\top V^{-1} V^{-1} m)^{1/2}}$$ Donde,

$$m = (m_1, \ldots, m_n)^\top$$

Donde, $m_1, \ldots, m_n$ son valores medios del estadístico ordenado, de variables aleatorias independientes e idénticamente distribuidas, muestreadas de distribuciones normales. V es la matriz de covarianzas de ese estadístico de orden.

Para realizar una prueba de normalidad en InFostat seleccione Estadísticas, Inferencia basada en una muestra, Prueba de Normalidad (Shapiro-Wilks modicado).



El resultado de la prueba es el siguiente:

Shapiro-Wilks (modificado)

Variable	n	Media	D.E.	W*	p(Unilateral D)
d (cm)	100	17.93	6.84	0.91	<0.0001

n = 100   Media = 17,93 cm
DE: 6,84 cm (desviación estándar)
W*: 0,91 (valor del estadísticoW de Shapiro-Wilks)
p (Unilateral D): <0,0001 (valor de p calculado)

Decisión: Dado que p es menor que 0,001 se rechaza Ho.

Conclusión: El diámetro medio del parche de bosque no proviene de una población con una distribución normal a un nivel de significancia de 0,01 y por lo tanto no cumple con el requisito de normalidad de la prueba t.

6. Prueba de hipótesis

Aún cuando el set de datos no cumple con el requisito de normalidad, con fines didácticos, se procede a mostrar el resultado de la prueba de hipótesis para la media. Recordemos que la hipótesis a probar es:

H_O: μ_0 = 15 cm (La media de la población es igual a 15cm)

$H_{A:}$ $\mu_0 \neq$ 15 cm (La media de la población es diferente a 15cm)

Y que el nivel de significancia es α= 0,01 (1%).

Prueba t para una media

Valor de la media bajo la hipótesis nula: 15

Variable	n	Media	DE	LI(95)	LS(95)	T	p(Bilateral)
d (cm)	100	17.93	6.84	16.57	19.29	4.28	<0.0001

n = 100 (tamaño de la muestra)
Media = 17,93 cm
DE: 6,84 cm
LI (95): 16,57 cm
LS (95): 19,29 cm
T: 4,28 (valor de estadístico t)
p (bilateral): <0,0001 (p calculado)

Decisión: Dado que “p” es menor que 0,01 se rechaza Ho.

Nota: Otra forma de responder a lo planteado en la hipótesis nula es observar el ámbito del intervalo de confianza (16,57 cm < μ <19,29 cm), si el mismo no contiene del valor de μ, como en este caso (μ =15 cm) se debe rechazar H_0.

Conclusión: El diámetro medio del parche de bosque es estadísticamente diferente al diámetro esperado para un bosque de quince años en suelos fértiles y bien drenados de la Zona Norte; o sea el estimador es diferente del parámetro (μ) y por lo tanto pertenece a otra población.

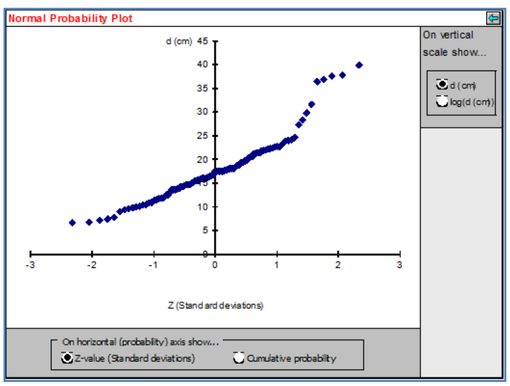
Uso de XLStatistics

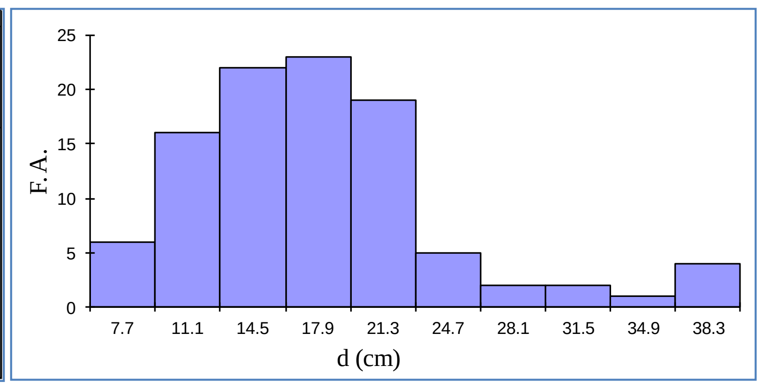
1. Active Excel, cargue el complemento XLSTatistics y lea los datos desde el archivo diámetros.xls.

2. Seleccione la columna d (cm) y haga un clic sobre 1Num. Antes de realizar la prueba de hipótesis procederemos a realizar una prueba de normalidad.

Prueba de normalidad

El gráfico de probabilidad normal y el histograma indican que la variable diámetro (cm) no proviene de una población con una distribución normal. El set de datos se aparta de la curva de normalidad en sus valores extremos (pequeños y grandes) y además presenta una distribución con asimetría positiva.





Estadístico de prueba: Prueba de bondad de ajuste de Chi-cuadrado

La prueba de chi-cuadrado (Snedecor y Cochran, 1989) se utiliza para probar si una muestra de datos (discreta o continua) proviene de una población con una distribución particular (e.g. normal, binomial, Poisson). Su principal ventaja es que puede aplicarse a cualquier distribución univariada para las que se puede calcular su función de distribución acumulada. Entre sus desventajas tenemos:

1. La prueba se aplica a datos agrupados y por lo tanto el valor del estadístico de prueba depende del número de clases utilizadas. Para que la aproximación Chi-cuadrado sea válida, la frecuencia esperada por clase debe ser al menos cinco.
2. Requiere de un tamaño de muestra suficientemente grande para que la aproximación de chi-cuadrado sea válida. La prueba no es válida para muestras pequeñas.

La prueba de chi-cuadrado es una alternativa a las pruebas de bondad de ajuste de Kolmogorov-Smirnov (K-S) y Anderson-Darling (la cual es una modificación de K-S), las cuales solo pueden utilizarse con distribuciones continuas.

La hipótesis y el estadístico de prueba de chi-cuadrado es el siguiente:

H_0: Los datos siguen la distribución especificada (en este caso la normal).
H_A: Los datos siguen la distribución especificada (en este caso la normal).

El estadístico de prueba es:

$$\chi^2 = \sum_{i=1}^{k} (O_i - E_i)^2 / E_i$$

Donde O_i es la frecuencia observada para la clase *i* y E_i es la frecuencia esperada para dicha clase. La frecuencia esperada se es igual a:

$$E_i = N(F(Y_u) - F(Y_l))$$

Donde F es la función de distribución acumulada para la distribución que se somete a prueba (en este caso la normal), Y_u es el límite superior de la clase ***i***, Y_l es el límite inferior de la clase ***i***, y *N* es el tamaño de la muestra.

El estadístico de prueba χ^2 sigue, aproximadamente, una distribución de chi-cuadrado con (k - c) grados de libertad, donde k es el número de clases no vacías y c = el número de parámetros estimados (i..e parámetros de localización, escala y forma) para la distribución + 1. Por ejemplo, para la distribución normal, c = 2. Por lo tanto, la hipótesis de que los datos provienen de una población con la distribución especificada se rechaza si

$$\chi^2 > \chi^2_{(1\text{-alfa, k-c})}$$

Donde, $\chi^2_{(1\text{-alfa, k-c})}$ es el valor crítico de chi-cuadrado con k - c grados de libertad y un nivel de significación α.

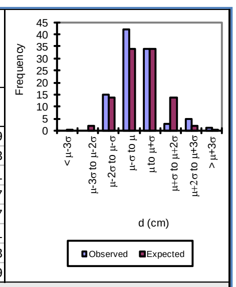
Los resultados de la prueba de hipótesis se presentan a continuación:

Goodness-Of-Fit Test for Normality of d (cm)

Data

Mean (μ)	17.93
St Dev (σ)	6.841267

	Proportions		Frequencies	
Range	Observed	Expected	Observed	Expected
< μ-3σ	0	0.00135	0	0.13499
μ-3σ to μ-2σ	0	0.0214	0	2.140023
μ-2σ to μ-σ	0.15	0.135905	15	13.59051
μ-σ to μ	0.42	0.341345	42	34.13447
μ to μ+σ	0.34	0.341345	34	34.13447
μ+σ to μ+2σ	0.03	0.135905	3	13.59051
μ+2σ to μ+3σ	0.05	0.0214	5	2.140023
> μ+3σ	0.01	0.00135	1	0.13499



Hypothesis Test

H_0: Population is normally distributed with the stated Mean and St Dev
H_1: Population is not normally distributed with the stated Mean and St Dev

Chisquare	21.85199
DF	7
p-value =	0.002694

Decisión: Dado que p es menor que 0,01 se rechaza Ho.

Conclusión: El diámetro medio del parche de bosque no proviene de una población con una distribución normal y por lo tanto no cumple con el requisito de normalidad de la prueba t.

Ejercicio: Como práctica transforme los datos utilizando log y realice nuevamente la prueba de hipótesis de normalidad. ¿Cuál es el resultado?

Prueba de hipótesis para la media

Aún cuando el set de datos no cumple con el requisito de normalidad, con fines didácticos, se procede a mostrar el resultado de la prueba de hipótesis para la media. Recordemos que la hipótesis a probar es:

H_O: μ_0 = 15 cm (La media de la población es igual a 15cm)
$H_{A:}$ μ_0 $\neq$ 15 cm (La media de la población es diferente a 15cm)

Y que el nivel de significancia es α= 0,01 (1%).

Tests on the Mean (μ) (t-tests)

Sample Data	
Sample Size	100
Mean	17.93
Standard Deviation	6.841267
SE Mean	0.684127

Hypothesis Tests	
H_0: μ =	15
Alternative: ? > <	
H_1: μ ≠	15
T	4.28283
DF	99
p-value =	4.3E-05

Confidence Intervals for μ		
Type (2,U,L)	2	
Confidence Level	0.95	
ME	Lower	Upper
1.357456	16.57254	19.28746

Sample size (n) = 100 (tamaño de la muestra)
Mean (Media) = 17,93 cm
Standard Deviation (DE): 6,84 cm
T: 4,28 (valor de estadístico t)
DF (grados de libertad; n-1) = 99
p-value (dos colas):4,3 E-05 o sea 0,000043 (p calculado)

Confidence level (nivel de confianza 95%)
ME (error medio): 1,36 cm. Su valor es igual a (19,28- 16,57) /2 = 1,36 cm
Lowe (LI, 95%): 16,57 cm
Upper (LS, 95%): 19,29 cm

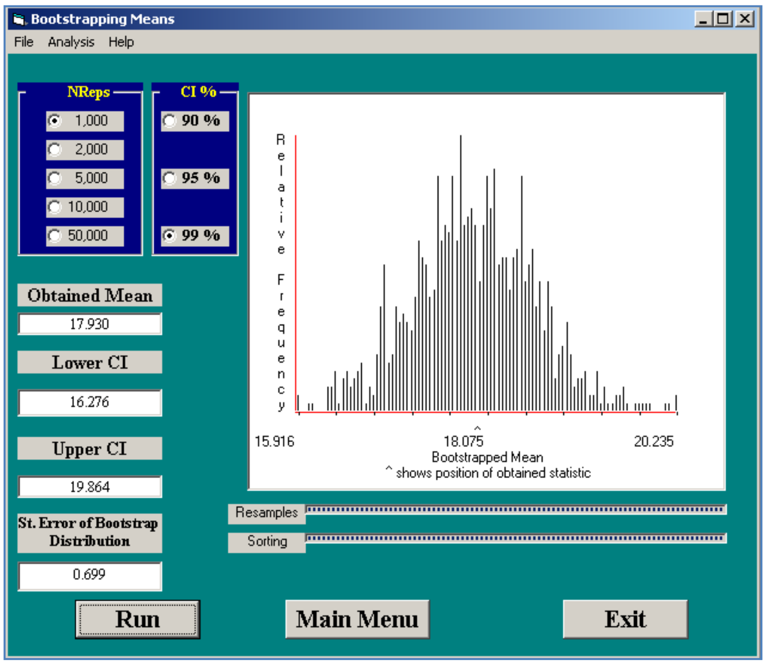
Decisión: Dado que p es menor que 0,01 se rechaza Ho.

Conclusión: El diámetro medio del parche de bosque es estadísticamente diferente al diámetro esperado para un bosque de quince años en suelos fértiles y bien drenados de la Zona Norte; o sea el estimador es diferente del parámetro (μ) y por lo tanto pertenece a otra población.

Uso de remuestreo: Resampling Stats

Cuando el set datos no cumple con el requisito de normalidad una de las mejoras opciones es utilizar un método de de remuestreo para realizar la prueba de hipótesis ya dichos métodos no asumen a priori que los datos provienen de una distribución particular. A continuación se muestran los resultados para el programa *Resampling Procedures*[2].

Para 1000 repeticiones las estimaciones de *bootraping* son:
Media es 17,93 cm
Error estándar de estimación: 0,70 cm
Intervalo de confianza (95%): 16, 28cm < μ <19,90 cm.



Conclusión: Dado que el intervalo de confianza no incluye el valor de μ (15 cm) se concluye que el diámetro medio del parche de bosque es estadísticamente diferente al diámetro esperado para un bosque de quince años en suelos fértiles y bien drenados de la Zona Norte; o sea el estimador es diferente del parámetro (μ) y por lo tanto pertenece a otra población.

Resumen

1. La prueba "t" asume que los datos fueron obtenidos al azar y que tienen una distribución normal.

2. Para muestras moderadamente grandes y una prueba de una cola, el estadístico "t" es relativamente robusto a violaciones moderadas del supuesto de normalidad.

3. Dado que los datos de diámetro son no normales, usted tiene las siguientes opciones:
   a) Transformar los datos (e.g. Log)
   b) Utilizar un prueba no paramétrica

[2] http://www.uvm.edu/~dhowell/StatPages/Resampling/Resampling.htm

c) Utilizar remuestreo y calcular la media y el intervalo de confianza esperado para el set de datos.

4. Del análisis realizado en el presente ejemplo, se concluye que el crecimiento medio en diámetro de la parcela es diferente del valor esperado para un sitio fértil de bosque segundario para un nivel de significancia de 0,001; sin embargo no es posible afirmar que el crecimiento en la parcela sea mejor o peor que un "buen sitio"; ya que se realizó una prueba de dos colas. Si desea clasificar su parcela como *mejor* ó *peor* que un buen sitio debe realizar una prueba de 1 cola.

5. La decisión de realizar una prueba de una cola debe hacerse antes de realizar hacer las mediciones para evitar que los datos de campo le indiquen en qué sentido debe plantear la prueba de hipótesis. Si se deja influenciar por los datos de campo estaría cometiendo un sesgo o decisión subjetiva a la hora de plantear su hipótesis alternativa y estaría haciendo una prueba de dos colas, esté o no consciente de ello.

4.2. Efecto de la variabilidad de los datos en la prueba de hipótesis

La variabilidad del set de datos es uno de los principales aspectos que el investigador(a) debe controlar en su experimento/estudio. Para ilustrar el efecto de la variabilidad en la prueba de hipótesis se presentan a continuación los resultados de pruebas de hipótesis de dos colas realizadas para otras dos parcelas de bosque natural con medias similares (14,2 cm y 14,4 cm) pero con variabilidad diferente (CV: 22.2% y CV: 30.2%; respectivamente).

Parcela	media muestral (cm)	μ (cm)	CV (%)	Ho	Conclusión (α = 0.05)
1	14.2	15.0	8.85	μ = 15	Rechazar Ho
2	14.4	15.0	22.15	μ = 15	No rechazar Ho
3	14.1	15.0	30.18	μ = 15	No rechazar Ho

Este ejemplo ilustra claramente el efecto de la variabilidad del set de datos en el resultado de la prueba de hipótesis. Por esta razón es esencial controlar dicho aspecto mediante un diseño apropiado de muestreo. En el presente ejemplo, el valor del "t crítico" se mantiene constante en todos los casos y que lo que cambia es el valor del error estándar. Si usted analiza la fórmula del estadístico de prueba "t" notará que el error estándar es el denominador de la ecuación y por tanto al aumentar dicho valor se reducirá el valor de "t" y por ende la prueba de hipótesis tenderá a ser no significativa.

4.3. Comparación de dos medias con varianzas desconocidas y muestras pequeñas[3]

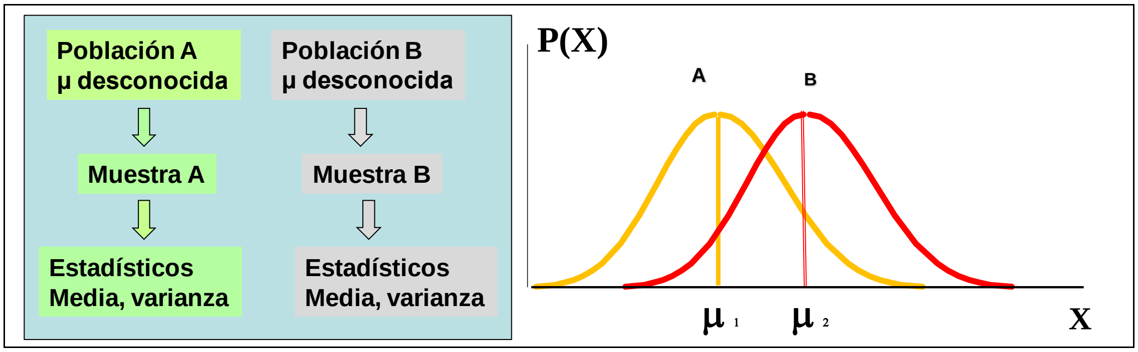
La prueba para dos medias con varianzas desconocidas y muestras pequeñas puede realizarse para medias independientes o para medias pareadas o dependientes.

[3] Al igual que para el caso de una media con varianza conocida también existe una prueba de hipótesis para dos medias de una población normal con varianzas conocidas, la cual utiliza el estadístico de prueba es Z. Dado que en la práctica es muy poco probable conocer dicho parámetro se ha omitido la prueba.

4.3.1. La prueba t de Estudiante para dos medias independientes

En la mayoría de los estudios el investigador(a) desconoce el valor real de la variable de interés (*parámetro*) y por tanto se trabaja con su *estimador*. En la terminología estadística los parámetros son constantes representados por letras griegas (e.g. α, μ, σ) en tanto que los estimadores son valores que cambian de muestra a muestra y que se representan con letras latinas (Ej. a, b, g).

Por ejemplo, podemos tener dos parcelas que representan una muestra de dos sitios diferentes, sin embargo no sabemos si ambas parcelas pertenecen a la misma población (μ) o a dos poblaciones diferentes (μ_1 y μ_2). Para responder a la interrogante se requiere medir la totalidad de la población; algo que no es práctico ni económico. A continuación se presenta el uso de la prueba t de Estudiante para dos medias independientes y con distribución normal como un método estadístico que permite responder a esta pregunta.



A y B representan dos poblaciones. Las curvas muestran la distribución muestral de las medias para cada una de las poblaciones. Cuanto más cerca se encuentren las respectivas medias, menos probable será detectar una diferencia significativa entre las poblaciones.

EJEMPLO

A continuación se ilustra cómo realizar una prueba de hipótesis para dos medias aleatorias provenientes de una población con una distribución normal. Suponga que los datos del archivo diametros_zn_zs.xlsx representan mediciones de diámetro a la altura del pecho para dos parcelas de igual edad y de la misma especie; una ubicada en la Zona Norte y la otra en la Zona Sur.

La pregunta que nos interesa responder es ¿muestran los datos alguna evidencia que nos permite suponer que la especie crece en forma diferente en el Norte y en el Sur del país? Del planteamiento de la pregunta se concluye que interesa saber si el crecimiento es igual en ambas zonas del país y por esta razón se plantea una prueba de do colas o bilateral.

Ho: $\mu_1 = \mu_2$

Ha: $\mu_1 \neq \mu_2$

Para realizar esta prueba de hipótesis se recomienda seguir el siguiente procedimiento:

1. *Cálculo de estadísticos descriptivos*
   Los estadísticos descriptivos resumen lo relevante de los datos en términos de tendencia central, variabilidad y forma de la distribución. Normalmente se calcula la media, desviación estándar, coeficiente de variación, error estándar y los coeficientes de asimetría y curtosis.

2. *Análisis gráfico*
   El objetivo del análisis gráfico es detectar patrones o tendencias en el set de datos. Por ejemplo, se puede analizar la tendencia central, la variabilidad y la forma de la distribución que caracteriza al set de datos. Los gráficos de Box-Whisker y de barra de errores (desviación estándar, error estándar, intervalo de confianza) son apropiados para visualizar el comportamiento de dos o más sets de datos. Cuando se desea evaluar la normalidad de los datos puede utilizarse un histograma o un diagrama de probabilidad normal.

3. *Prueba de hipótesis*
   Una vez que usted se ha familiarizado con el set de datos puede proceder a realizar la prueba de hipótesis. El proceso involucra los siguientes pasos:
   A. Plantear la hipótesis nula y alternativa.
   B. Seleccionar el estadístico de prueba y definir el nivel de significancia. Si opta por una prueba paramétrica debe:
      a. Realizar una prueba sobre normalidad de ambos sets de datos.
      b. Realizar una prueba de igualdad de varianzas para luego elegir entre una prueba de hipótesis con varianzas iguales o diferentes.
   C. Efectuar la prueba de hipótesis.
   D. Tomar una decisión
   E. Conclusión estadística y práctica.

Uso de Infostat

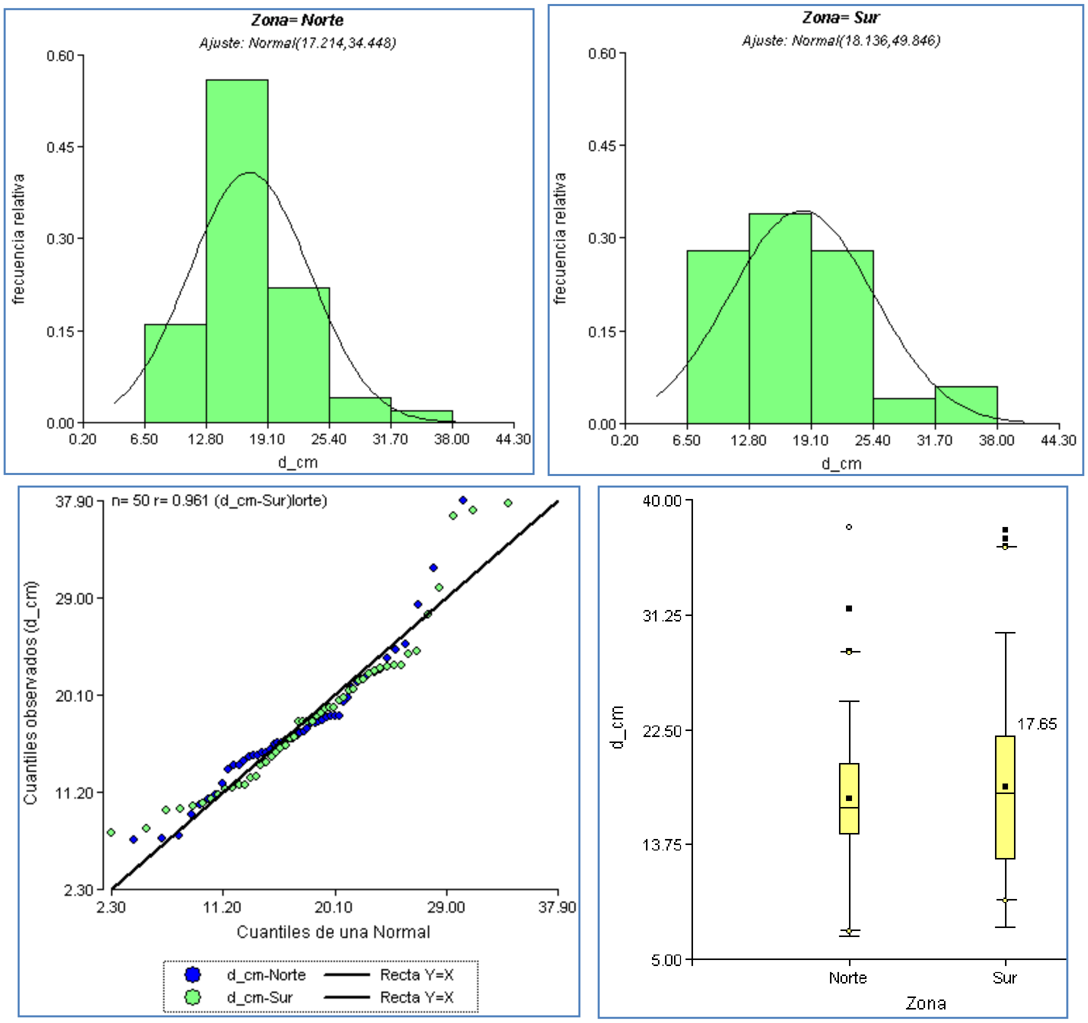
1. Estadísticos descriptivos

Resumen de estadísticos descriptivos

Zona	Variable	n	Media	D.E.	Var(n-1)	E.E.	CV	Mín	Máx	Mediana	Q1	Q3	Asimetría	Kurtosis
Norte	d_cm	50	17.21	5.87	34.45	0.83	34.10	6.80	37.90	16.50	14.40	19.90	1.03	2.60
Sur	d_cm	50	18.14	7.06	49.85	1.00	38.93	7.50	37.70	17.65	12.50	22.00	1.01	1.24

La media y mediana son muy similares en ambas parcelas aunque la del norte es un poco menos variable (CV%:34,4) que la del sur (CV%:38,9). La forma de la distribución es ligeramente asimétrica hacia la derecha y platicurtica, en especial para el set de datos del norte.

2. Análisis gráficos



El histograma de la distribución diamétrica indica que los datos de la Zona Norte tienden a una distribución normal; sin embargo los valores de la Zona Sur tienden a ser más aplanados (distribución platicúrtica) y con una cola hacia la derecha (asimetría positiva). La curva de distribución normal confirma la desviación de la normalidad de los datos de la Zona Sur. Para normalizar la variable diámetro se pueden aplicar las siguientes transformaciones: Log (d), raíz (d) o 1/d. La gráfica de cajas muestra que ambas series son muy similares en cuanto a tendencia central pero no en cuanto a variabilidad.

3. Prueba de hipótesis

3.1 Plantear la hipótesis nula y alternativa.

$H_0 : \mu_1 = \mu_2$ (las dos medias muestrales son iguales o sea provienen de la misma población)

$H_A : \mu_1 \neq \mu_2$ (las dos medias muestrales son diferentes o sea provienen de poblaciones diferentes)

3.2 Seleccionar el estadístico de prueba y definir el nivel de significancia.

El estadístico de prueba es t con un nivel de significancia de 0,05 (equivale a un nivel de confianza de 1-0,05= 95%)

a. Prueba de normalidad

H_0: los datos provienen de una distribución normal.

H_A: los datos provienen de una distribución no normal.

Shapiro-Wilks (modificado)

Zona	Variable	n	Media	D.E.	W*	p(Unilateral D)
Norte	d_cm	50	17.21	5.87	0.94	0.0520
Sur	d_cm	50	18.14	7.06	0.91	0.0023

No se rechaza Ho para α=0,05
Se rechaza Ho α=0,05

Para un nivel de significancia de 0,05% (α=0,05) la prueba de normalidad de Shapiro-Wilks indica que el set de datos del norte *es normal* ya que el "p" calculado (0,052) es mayor que el valor de "p" critico (0,050); en tanto que el set de datos del sur es *no normal* dado que el p calculado (0,002) es menor que el p crítico (0,05).

Es un buen momento para ilustrar el efecto del nivel de significancia elegido en el resultado de la prueba de hipótesis. Por ejemplo, los datos de la zona norte serían no normales si se hubiese elegido un alfa de 0,1 ya que el valor de p calculado (0,052) hubiese sido menor que el valor del p critico (0,1).
En tanto que en la zona sur si se hubiese elegido un α de 0,001 se hubiese aceptado Ho ya que el p calculado (0,002) hubiese sido mayor que α.

b. Prueba de igualdad de varianzas

La hipótesis a probar es la siguiente:

Ho: $\boldsymbol{\sigma^2{}_1 / \sigma^2 = 1}$ (las varianzas de ambas poblaciones es la misma)
Ha: $\boldsymbol{\sigma^2{}_1 / \sigma^2 \neq 1}$ (las varianzas de ambas poblaciones son diferentes)

Esta prueba de hipótesis se realiza por la prueba de hipótesis sobre la igualdad de medias requiere que usted elija entre medias con varianzas iguales o medias con varianzas diferentes.

El estadístico de prueba es:

Fmax = S^2x / S^2y

En donde, S^2x es la varianza mayor y S^2y es la varianza menor; el cual tiene una distribución F con n-1 (numerador) y m-1 (denominador) grados de libertad si la hipótesis nula de igualdad de varianzas es cierta. De lo contrario, tiene una distribución F no central.

La hipótesis nula (H_0) se rechaza si F es demasiado grande (cola superior) o demasiado pequeño (cola inferior). Esta prueba es extremadamente sensible a datos no normales (es poco robusta) y por lo tanto con frecuencia se prefieren las pruebas de Levene, Bartlett o Brown-Forsythe; aunque se le considera robusta para valores de alfa inferiores a 0,05 y diseños balanceados.

Si $F_{calculado} < F_{Crítico}$ se rechaza H_0 o si "p" calculado es menor que "p" crítico.

Prueba F para igualdad de varianzas

Variable	Grupo(1)	Grupo(2)	n(1)	n(2)	Var(1)	Var(2)	F	p	prueba
d cm	{Norte}	{Sur}	50	50	34.45	49.85	0.69	0.1995	Bilateral

El valor de p calculado (0,199) es superior al valor de p crítico (0,05) y por lo tanto no se rechaza H_0 y se concluye que las varianzas son iguales.

Implicación: Se puede utilizar una prueba de hipótesis de medias independientes con varianzas iguales. Recuerde que esta prueba asume que los datos provienen de una distribución normal (lo cual no es cierto para los datos del norte).

c. Prueba de hipótesis de medias independientes con varianzas iguales

El estadístico de prueba es "T":

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{s_1^2/N_1 + s_2^2/N_2}}$$ cuando las varianzas de las muestras son diferentes.

Donde:
N_1: tamaño de muestra1
N_2: tamaño de muestra2
$\bar{Y}_1$: media de la muestra 1
$\bar{Y}_2$: media de la muestra 2
S^2_1: varianza de muestra 1
S^2_2: varianza de muestra 2

Cuando las varianzas de las muestras son iguales, la expresión se reduce a:

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{s_p\sqrt{1/N_1 + 1/N_2}}$$

Donde, la varianza ponderada por el tamaño de la muestra es:

$$s_p^2 = \frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}$$

La hipótesis nula se rechaza si: $|T| > t_{1\text{-}\alpha/2,v}$

Donde $t_{1\text{-}\alpha/2,v}$ es el valor crítico de la distribución t Estudiante con un nivel de significancia α (lo cual equivale a un nivel de confianza de 1- α) y v grados de libertad, dados por:

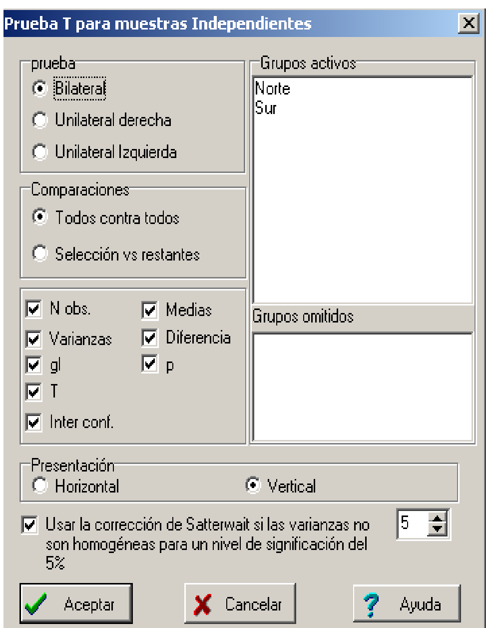
$$v = \frac{(s_1^2/N_1 + s_2^2/N_2)^2}{(s_1^2/N_1)^2/(N_1 - 1) + (s_2^2/N_2)^2/(N_2 - 1)}$$

Cuando las varianzas de las muestras son iguales, la expresión se reduce a:

$$v = N_1 + N_2 - 2$$

En este caso se utilizan los valores estandarizados de la distribución t como referencia o patrón para juzgar las diferencias observadas en los datos muestrales (para nuestro caso entre la media de las dos parcelas).

Para realizar esta prueba en InfoStat seleccione: Estadísticas, Inferencias basadas en dos muestras, Prueba t y configura la ventana de diálogo como se muestra a continuación:



Prueba: Bilateral o de dos colas.
Comparaciones: todos contra todos (en este caso Norte y Sur)
N obs: tamaño de muestra
Varianzas: varianza de cada muestra.
gl: grados de libertad
T: valor del estadístico t
Inter conf: intervalo de confianza
Medias
Diferencias: Media Norte – media sur
p: valor de p calculado
Usar la corrección de Welch–Satterthwaite (1946) para calcular los gl cuando las varianzas no son iguales para un alfa de 0.05.

Prueba T para muestras Independientes

Variable:d_cm - Clasific:Zona - prueba:Bilateral

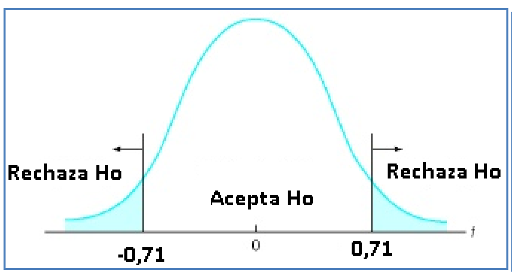
	Grupo 1	Grupo 2
	Norte	Sur
n	50	50
Media	17.21	18.14
Varianza	34.45	34.45
Media(1)-Media(2)	-0.92	
LI(95)	-3.50	
LS(95)	1.65	
pHomVar	0.1995	
T	-0.71	
gl	98	
p-valor	0.4793	

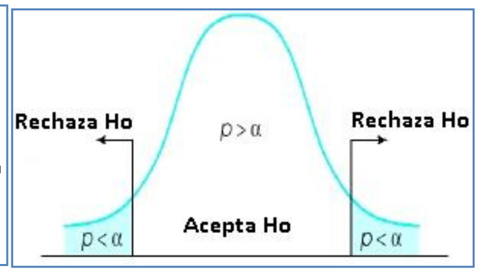
n: tamaño de muestra
Media de la muestra
Varianza de la muestra

Media (1) – Media (2): diferencia entre las medias muestrales.
LI (95): IC límite inferior
LS (95): IC límite superior
pHomVar: valor de p para prueba F de homogeneidad de varianzas
gl: grados de libertad $(n_1-1) + (n_2-1)$
p-valor: valor de p (percentil de la distribución t)

Nota: El software determina la probabilidad correspondiente al valor -0,71 dados que los datos sigan una distribución t de Estudiante.

4. Decisión





Dado un nivel de significancia de 0,05; el resultado de la prueba de hipótesis sobre la igualdad de medias con varianzas iguales indica que las mismas son estadísticamente iguales (p= 0,479) o sea no es posible rechazar H_0.

Recuerde: usted debe fijar su nivel de significancia alfa (α) antes de realizar la prueba de hipótesis.

5. Conclusión

Aun nivel de significancia de 5%, las muestras no muestran suficiente evidencia como para rechazar Ho y por lo tanto se concluye la especie crece igual en el Norte y en el Sur del país.

Observaciones:

- Observe que se utilizó un alfa de 0.05 para realizar la prueba (confianza de 95%).
- El paquete estadístico también le permite realizar una prueba de medias con varianzas diferentes.
- El paquete estadístico le brinda el valor de t crítico (valor a partir del cual se considera que las diferencias son significativas), el valor de t calculado (este valor corresponde al estadístico de prueba "t") así como la probabilidad asociado a dicho valor (valor de p). Si p es menor que el nivel de *significancia seleccionado* se declara la prueba como significativa o sea se rechaza Ho.

- Dado que en este caso p= 0,4793 se declara como no significativa la diferencia en diámetro medio entre las dos parcelas. Esto nos lleva a concluir que no existe evidencia estadística para argumentar que la especie se comporta en forma diferente en el Norte y en el Sur del país.

Uso de XLStatistics

Active Excel, cargue el complemento XLStatistics y lea los datos desde el archivo diámetros_zn_zn.xls. Seleccione las columnas d (cm) y zona y haga un clic sobre 1Num1Cat.

1. *Cálculo de estadísticos descriptivos*

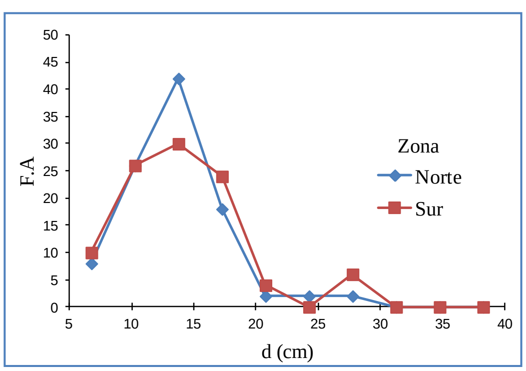
Category Labels and Numerical Summaries for d_cm

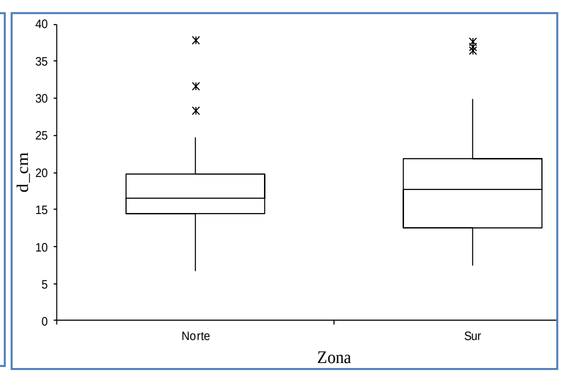
Zona	All	Norte	Sur
Number	100	50	50
Mean	17.68	17.21	18.14
St Dev	6.48	5.87	7.06
Skew	1.05	1.03	1.01
Min	6.8	6.8	7.5
Q_1	13.7	14.425	12.525
Median	17.25	16.5	17.65
Q_3	21.33	19.80	21.88
Max	37.9	37.9	37.7

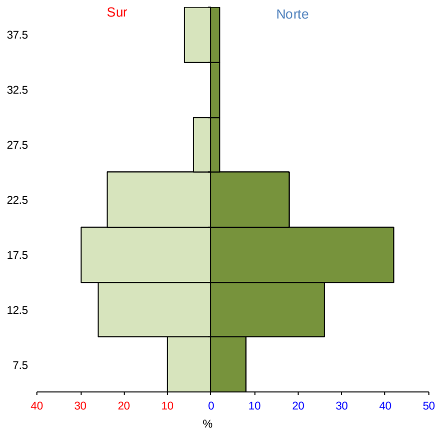
La media y mediana son muy similares en ambas parcelas aunque la del norte es un poco menos variable (S: 5,9 cm) que la del sur (S: 7,1 cm). La forma de la distribución es ligeramente asimétrica hacia la derecha y platicurtica, en especial para el set de datos del norte.

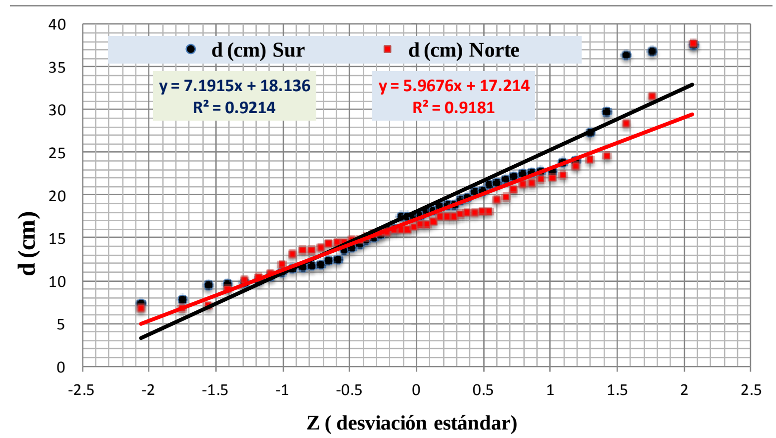
2. *Análisis gráfico*

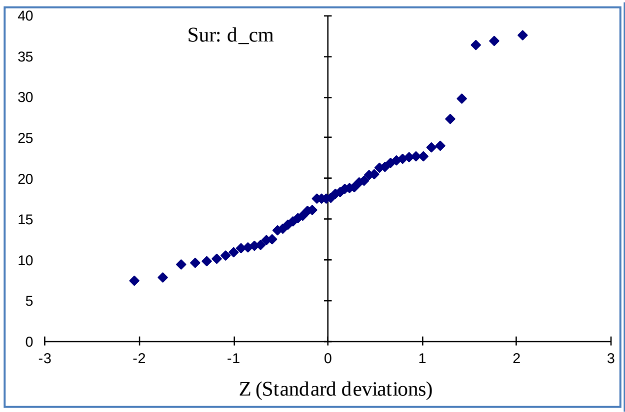
Con fines didácticos se han incluido gráficos diferentes a los utilizados con Infostat. Los gráficos de polígonos de frecuencia y pirámides permiten apreciar las similitudes y diferencias de la distribución diamétrica en las zonas Sur y Norte. La gráfica de cajas muestra que ambas series son muy similares en cuanto a tendencia central pero no en cuanto a variabilidad. Los datos de la Zona Norte son menos variables y tienden a una distribución normal aunque un tanto leptocúrtica (punteaguda) en tanto que los valores de la Zona Sur parecen no normales ya que tienden a ser más aplanados (distribución platicúrtica) y con una cola más larga hacia la derecha (asimetría positiva). Las gráficas de probabilidad normal permiten suponer que los datos del sur ajusten ligeramente mejor a una distribución normal que los del norte.

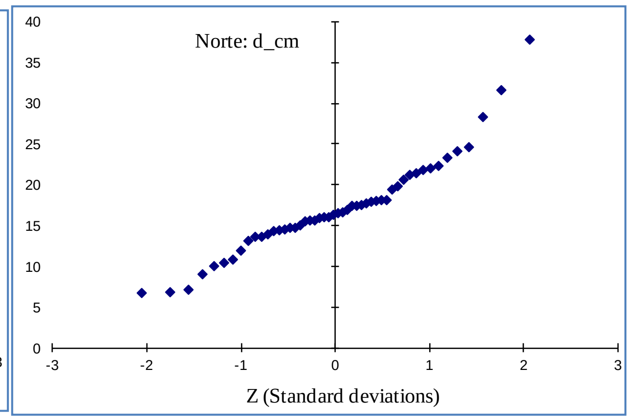












3. *Prueba de hipótesis*

3.1 Plantear la hipótesis nula y alternativa.

H_0: $\mu_1 = \mu_2$
H_A: $\mu_1 \neq \mu_2$

3.2 Seleccionar el estadístico de prueba y definir el nivel de significancia.
El estadístico de prueba es t con un nivel de significancia de 5% (alfa: 0,05).

a. Prueba de normalidad de Chi-cuadrado

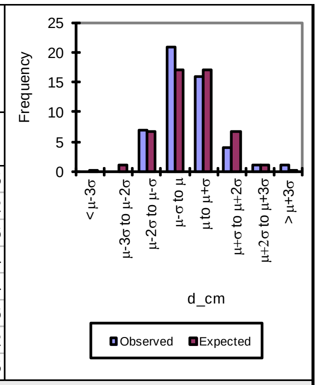
H_0: los datos provienen de una distribución normal.
H_A: los datos provienen de una distribución no normal.

Zona Norte
Goodness-Of-Fit Test for Normality of d_cm

Data

Mean (μ)	17.214
St Dev (σ)	5.869221

	Proportions		Frequencies	
Range	Observed	Expected	Observed	Expected
< μ-3σ	0	0.00135	0	0.067495
μ-3σ to μ-2σ	0	0.0214	0	1.070012
μ-2σ to μ-σ	0.14	0.135905	7	6.795256
μ-σ to μ	0.42	0.341345	21	17.06724
μ to μ+σ	0.32	0.341345	16	17.06724
μ+σ to μ+2σ	0.08	0.135905	4	6.795256
μ+2σ to μ+3σ	0.02	0.0214	1	1.070012
> μ+3σ	0.02	0.00135	1	0.067495



Hypothesis Test

H_0: Population is normally distributed with the stated Mean and St Dev

H_1: Population is not normally distributed with the stated Mean and St Dev

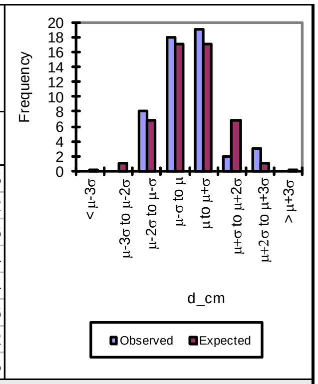
Chisquare	16.15448
DF	7
p-value =	0.023742

Zona Sur
Goodness-Of-Fit Test for Normality of d_cm

Data

Mean (μ)	18.136
St Dev (σ)	7.060201

	Proportions		Frequencies	
Range	Observed	Expected	Observed	Expected
< μ-3σ	0	0.00135	0	0.067495
μ-3σ to μ-2σ	0	0.0214	0	1.070012
μ-2σ to μ-σ	0.16	0.135905	8	6.795256
μ-σ to μ	0.36	0.341345	18	17.06724
μ to μ+σ	0.38	0.341345	19	17.06724
μ+σ to μ+2σ	0.04	0.135905	2	6.795256
μ+2σ to μ+3σ	0.06	0.0214	3	1.070012
> μ+3σ	0	0.00135	0	0.067495



Hypothesis Test

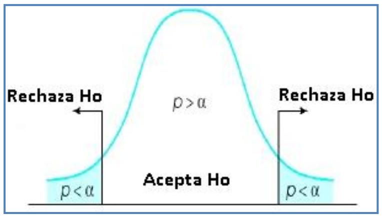
H_0: Population is normally distributed with the stated Mean and St Dev

H_1: Population is not normally distributed with the stated Mean and St Dev

Chisquare	8.553481
DF	7
p-value =	0.286322

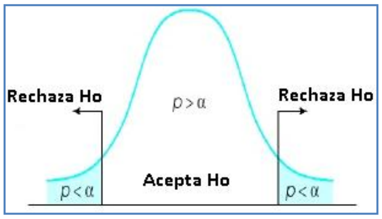
Decisión:

A. Zona Norte:



Para un nivel de significancia de 0,05 (p crítico); la prueba de chi-cuadrado indica que el set de datos es no normal ya que el valor de p calculado (0,0237) es menor que el valor del p crítico (se rechaza H_0).

B. Zona Sur



Para un nivel de significancia de 0,05 (p crítico); la prueba de chi-cuadrado indica que el set de datos es normal ya que el valor de p calculado (0,2863) es mayor que el valor del p crítico (no se rechaza H_0).

Conclusión:

Para un nivel de significancia de 5% (α=0,05) la prueba de normalidad de chi-cuadrado indica que el set de datos del norte *es no normal* ya que el "p" calculado (0,024) es menor que el valor de "p" critico (0,05); en tanto que el set de datos del sur es *normal* dado que el p calculado (0,286) es mayor que el p crítico (0,05). Sin embargo si usted observa las tablas de frecuencias observará que existen dos (zona sur) y tres clases (zona norte) que poseen menos que cinco observaciones cada una, lo que viola uno de los requerimientos de esta prueba.

> En este caso el resultado de la prueba de chi-cuadrado es opuesto al de la prueba de Shapiro-Wilks para un mismo nivel de significancia.

Dada esta discrepancia se optó por realizar la prueba de Lilliefors disponible en línea.

Prueba de Lilliefors en línea (http://in-silico.net/tools/statistics/lillieforstest)

H_0: los datos provienen de una distribución normal.
H_A: los datos provienen de una distribución no normal.

Sur		**Norte**	
p-value	0.0401	p-value	0.0214
critical value	0.1245	critical value	0.1245
statistic	0.1276	statistic	0.1359

La prueba de Lilliefors indica que para un nivel de significancia de 5% (α=0,05) el valor de "p" calculado en ambos casos es menor (0,04 sur y 0,02 norte) y por lo tanto los datos provienen de una distribución no normal.

b. Prueba de hipótesis para igualdad de varianzas

Dado que se trata de dos poblaciones se debe realizar primero una prueba de igualdad de varianzas.

Ho: $\sigma^2{}_1 / \sigma^2 = 1$ (las varianzas de ambas poblaciones es la misma)
Ha: $\sigma^2{}_1 / \sigma^2 \neq 1$ (las varianzas de ambas poblaciones son diferentes)

Hartley (Fmax)

El resultado de la prueba de Hartley o Fmax es el siguiente:

F-Test for Variance

Sample Data					
n_1	50	n_2	50	$H_0 : \sigma_1^2 = \sigma_2^2$	
s_1^2	34.44776	s_2^2	49.84643	$H_1 : \sigma_1^2 \neq \sigma_2^2$	
				F	0.691078
				p-value =	0.199469

El valor de p calculado (0,199) es superior al valor de p crítico (0,05) y por lo tanto no se rechaza H_0 y se concluye que las varianzas son iguales.

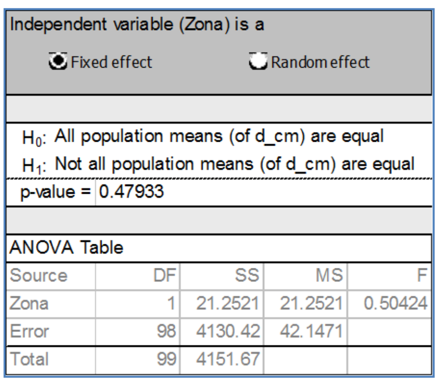
Implicación: Se puede utilizar una prueba de hipótesis de medias independientes con varianzas iguales. Recuerde que esta prueba asume que los datos provienen de una distribución normal (lo cual no es cierto para los datos del norte).

3.3. Prueba de hipótesis para dos medias independiente con varianzas iguales

El estadístico de prueba es "t", el cual tiene un distribución t de Estudiante con ($n_1 + n_2$)-2 grados de libertad.

XLSTatistics realiza una prueba de varianza (**Test-n cateogries**), la cual se utiliza para analizar la diferencia entre tres o más muestras. En esta prueba se utiliza el estadístico F.

H_0: Las medias son iguales.
H_A: No todas las medias son iguales.

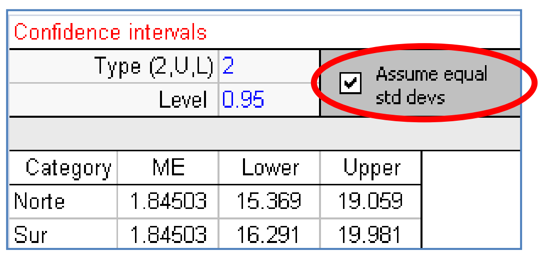

Independent variable (Zona) is a

Fixed effect Random effect

H_0: All population means (of d_cm) are equal
H_1: Not all population means (of d_cm) are equal
p-value = 0.47933

ANOVA Table

Source	DF	SS	MS	F
Zona	1	21.2521	21.2521	0.50424
Error	98	4130.42	42.1471	
Total	99	4151.67		


Confidence intervals

Type (2,U,L) 2
Level 0.95
Assume equal std devs

Category	ME	Lower	Upper
Norte	1.84503	15.369	19.059
Sur	1.84503	16.291	19.981

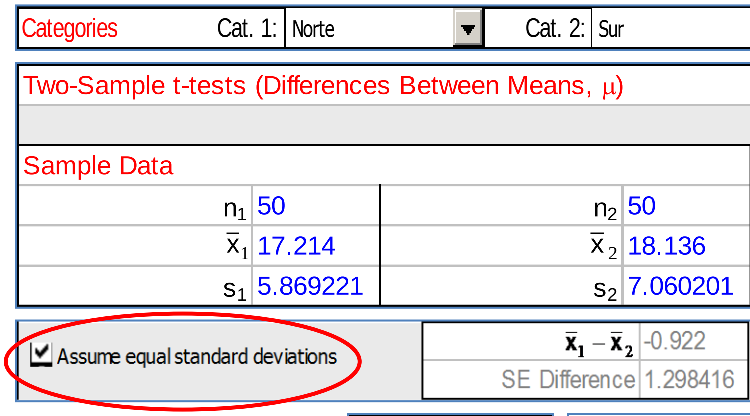
Debe marcar esta casilla porque las varianzas son iguales.

El valor de p calculado (0,479) es superior al valor de p crítico (0,05) y por lo tanto no se rechaza H_0 y se concluye que las medias son iguales; o sea ambas pertenecen a la misma población.

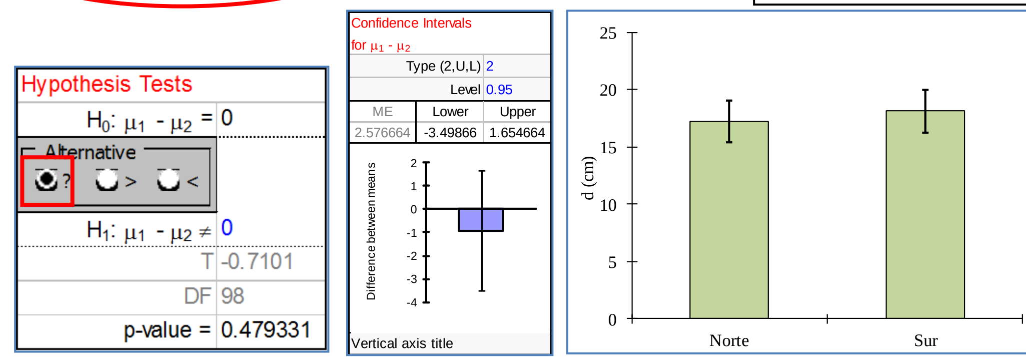
Test-2 Categories

La prueba de dos categorías (en este caso Norte y Sur) es la que se utiliza para someter a prueba la hipótesis sobre la igualdad de dos medias independientes.

H_0: Las dos medias iguales ($\mu_1 = \mu_2$).
H_A: Las dos medias no son iguales ($\mu_1 \neq \mu_2$).

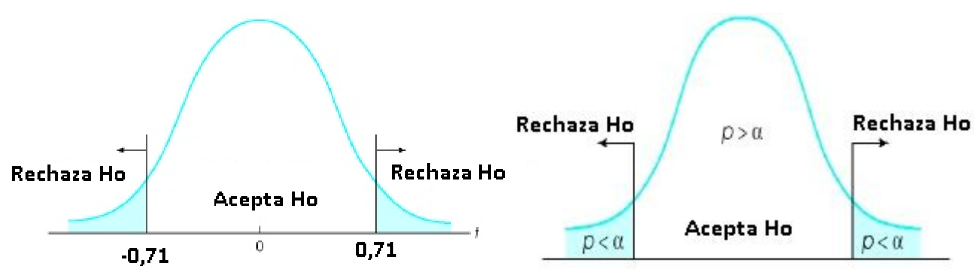


Debe marcar esta casilla porque las varianzas son iguales.



Marque la casilla diferente y en la celda de H_1 digite 0.

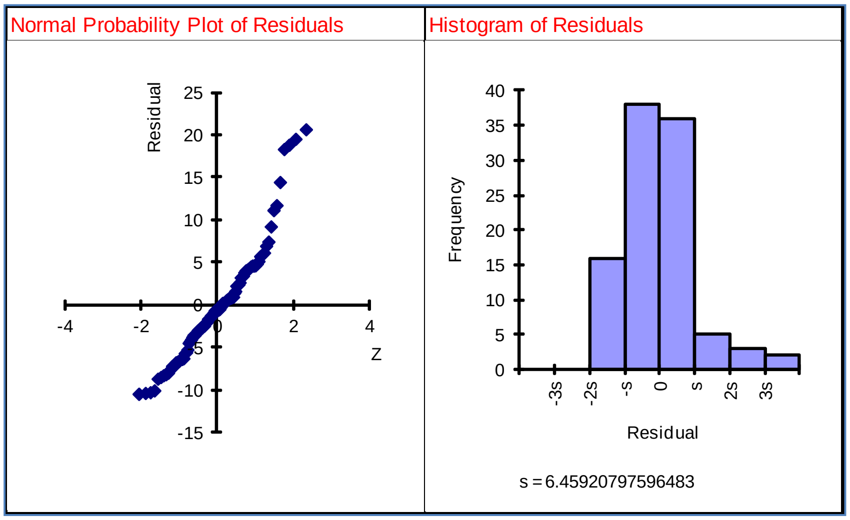
Decisión: El valor de p calculado (0,479) es superior al valor de p crítico (0,05) y por lo tanto no se rechaza H_0.



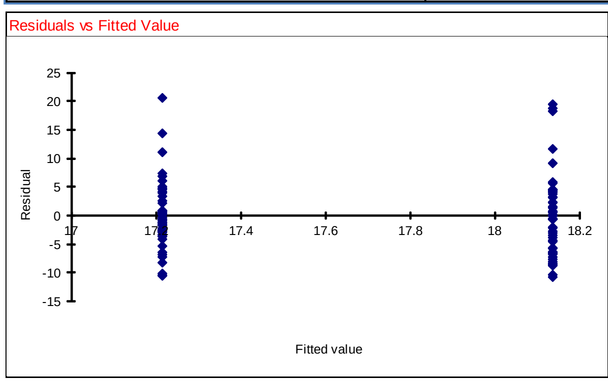
Conclusión: A un nivel de significancia de 0,05 la media del diámetro para la parcela del norte y del sur es igual; o sea, ambas pertenecen a la misma población. Observe que este caso el valor de "p" para el estadístico "t" es igual al valor de "p" para el estadístico F. El intervalo de confianza para la diferencia entre medias (μ_1-μ_2) es: -3,5 cm a 1,65 cm e incluye el valor cero (0), lo que también nos indica que la diferencia entre las medias es cero (0).

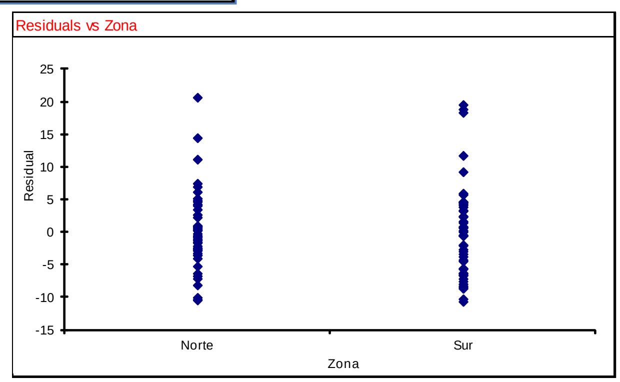
Análisis de residuos

Al realizar una prueba de hipótesis de dos o más grupos XLSTatistics le brinda los siguientes gráficos de residuos. Estos gráficos permiten evaluar los supuestos de la prueba de hipótesis: en este caso normalidad e igualdad de varianzas entre grupos (norte y sur).



Tanto el gráfico de normalidad como el histograma de residuos indican que no se cumplió con el supuesto de normalidad. Algo que ya sabíamos por los análisis previos.





Los gráficos de residuos versus valores ajustados y versus “zona” permiten evaluar el supuesto de igualdad de varianzas. Como puede observarse, el análisis sí cumplió con este supuesto (también lo demostramos con la prueba de hipótesis Fmax).

4.3.2. Datos no normales o simétricos: ¿qué hacer?

Cuando los datos no cumplen con el supuesto de normalidad, usted tiene las siguientes opciones.

1. Transformar los datos. Para variables positivas se pueden utilizar las transformaciones logarítmica y raíz cuadrada y para datos sin ceros el inverso (1/x). La transformación de Box-Cox es una familia de transformaciones definida como: $T(Y) = (Y^{\lambda} - 1)/\lambda$, donde Y es la variable respuesta y lambda (λ) es el parámetro de la transformación. Para lambda = 0 la transformación es igual a utilizar el logaritmo natural de los datos.

2. Utilizar un equivalente no parámetro de la prueba T. Por ejemplo, para dos muestras independientes con distribuciones asimétricas, la prueba de U de Mann-Whitney es una excelente opción ya que puede tener de tres a cuatro veces más potencia que la prueba t. Sin embargo cuando las muestras cumplen con el supuesto de normalidad la prueba U de Mann-Whitney tiene una potencia relativa de 95% con respecto a la prueba t de Estudiante. Esto significa que se consigue la misma potencia con una muestra de 100 elementos cuando se utiliza U de Mann-Whitney que con 95 elementos cuando se utiliza la t de Estudiante.

3. Utilizar remuestreo y dejar que sus datos definan el valor de p calculado.

Prueba de U de Man-Witney (equivalente de prueba t para dos muestras independientes)

A continuación se ilustra cómo utilizar InfoStat y XLStatiscs para realizar una prueba U de Man-Witney, la cual prueba por igualdad de medianas y no requiere que los datos sean normales.

InfoStat

Hipótesis nula y alternativa

H_0: La mediana de la muestra 1 es igual a la mediana de la muestra 2.
H_A: La mediana de la muestra 1 es diferente a la mediana de la muestra 2.
Nivel de significancia: 0,05

Prueba de Wilcoxon para muestras independientes

Clasific	Variable	Grupo 1	Grupo 2	n(1)	n(2)	Media(1)	Media(2)	DE(1)	DE(2)	Mediana(1)	Mediana(2)	W	p(2 colas)
Zona	d cm	Norte	Sur	50	50	17.21	18.14	5.87	7.06	16.50	17.65	2437.50	0.5463

Decisión: El valor de p calculado (0,5463) es superior al valor de p crítico (0,05) y por lo tanto no se rechaza H_0.

Conclusión: A un nivel de significancia de 0,05 la media del diámetro para la parcela del norte y del sur es igual; o sea, ambas pertenecen a la misma población.

XLSTatistics

Hipótesis nula y alternativa

H_0: La mediana de la muestra 1 es igual a la mediana de la muestra 2.
H_A: La mediana de la muestra 1 es diferente a la mediana de la muestra 2.
Nivel de significancia: 0,05

Mann-Whitney Test (Differences Between Medians) (Diferencias entre medianas)

Mann-Whitney Test	
Sample $Median_1$	16.5
Sample $Median_2$	17.65
H_0: $Median_1$ - $Median_2$ =	0
Alternative: ● ? ○ > ○ <	
H_1: $Median_1$ - $Median_2$ ≠	0
U	1162.5
Z	0.603209
p-value =	0.54637

Decisión: El valor de p calculado (0,5464) es superior al valor de p crítico (0,05) y por lo tanto no se rechaza H_0.

Conclusión: A un nivel de significancia de 0,05 la media del diámetro para la parcela del norte y del sur es igual; o sea, ambas pertenecen a la misma población.

Prueba de aleatorización para dos grupos

XLSTatistics le ofrece la opción de utilizar el método de aleatorización para calcular la diferencia entre medias de dos grupos así como el respectivo valor de "p" calculado. La ventaja de este método es que no dependen del supuesto de normalidad.

Hipótesis nula y alternativa

H_0: La mediana de la muestra 1 es igual a la mediana de la muestra 2.
H_A: La mediana de la muestra 1 es diferente a la mediana de la muestra 2.
Nivel de significancia: 0,05

Randomised 2-Group/Category Test

Number randomised samples generated	1100
	Reset
	Another sample
Difference between sample means	
Observed	Randomised
0.922	0.722
H_0: Group/Category membership has no effect on d_cm	
H_1: Group/Category membership effects d_cm	
p-value (est.) =	0.466363636

Decisión: El valor de p calculado (0,4664) es superior al valor de p crítico (0,05) y por lo tanto no se rechaza H_0.

Conclusión: A un nivel de significancia de 0,05 la media del diámetro para la parcela del norte y del sur es igual; o sea, ambas pertenecen a la misma población.

4.3.3. Prueba de hipótesis de una cola o unilateral

Hasta el momento hemos realizado pruebas de dos colas o bilaterales. A continuación se brida un ejemplo de una prueba de una cola o unilateral. Los conceptos y terminología expuestos para la prueba de dos colas también aplican a la prueba de una cola. Lo único que cambia es la hipótesis alternativa y la zona de rechazo de Ho. Veamos un ejemplo.

EJEMPLO

Suponga que en los suelos fértiles y bien drenados de la Zona Norte el diámetro medio del bosque es de 15 cm a la altura del pecho en quince años (μ= 15 cm). Un corredor de bienes raíces le ofrece una finca a un potencial comprador indicándole que el bosque crece mejor que en los mejores sitios de la Zona Norte. El inversionista desea saber si dicha afirmación es correcta y solicita un estudio técnico. El profesional responsable selecciona al azar una parcela con cien árboles en un parche de bosque de quince años ubicado en la finca y obtiene un diámetro medio de 17,93 cm con una desviación estándar de 6,84 cm.

Dado el planteamiento del problema, la pregunta que interesa responder es ¿muestran los datos de la parcela suficiente evidencia como para afirmar que el crecimiento del bosque en la finca es mayor que el crecimiento observado en los mejores sitios de la Zona Norte del país? Y por esta razón se plantea una prueba de una cola.

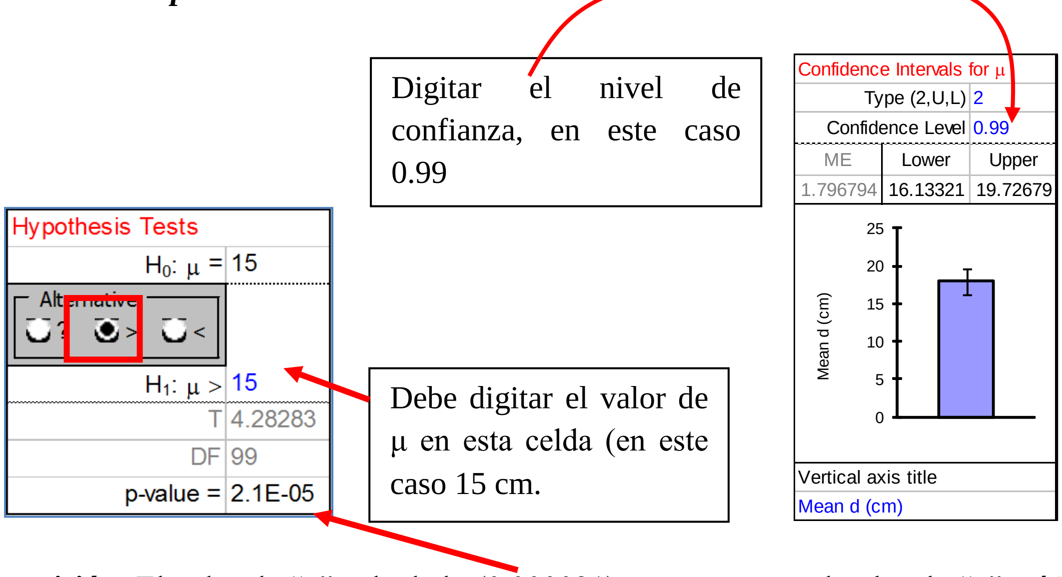
H_O: μ_0 = 15 cm (La media de la población es igual a 15cm)
$H_{A:}$ μ_0 > 15 cm (La media de la población es mayor que 15cm)
Nivel de significancia es α= 0,01 (1%).

Estadístico de prueba

El estadístico de prueba es t:

$t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$ donde S = desviación estándar, $\bar{X}$ = media muestral, n= tamaño de muestra y μ_0 la media poblacional. El estadístico "t" tiene una distribución "t" de Estudiante con n-1 grados de libertad. Los requisitos de la prueba son muestra independiente y datos con una distribución normal.

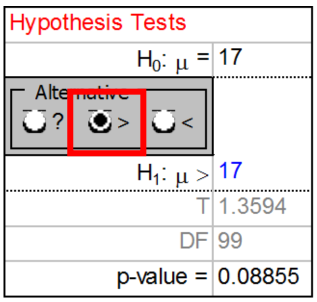
***Prueba de hipótesis utilizando* XLStatistics**



Decisión: El valor de "p" calculado (0,000021) es menor que el valor de "p" crítico (0,01) y por lo tanto se rechaza Ho.

Nota: Otra forma de responder a lo planteado en la hipótesis nula es observar el ámbito del intervalo de confianza (16,13 cm < µ <19,73 cm), el cual si no contiene el valor de µ (15 cm) indicando que Ho debe rechazarse.

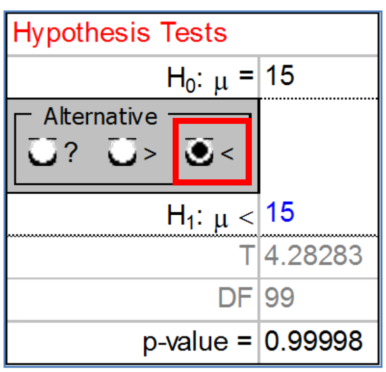
Conclusión: El diámetro medio del parche de bosque es *estadísticamente mayor* que el diámetro esperado para un bosque de quince años en suelos fértiles y bien drenados de la Zona Norte; o sea el crecimiento diamétrico del bosque es superior al de los mejores sitios de la zona. Su recomendación al comprador sería: compre la finca.



NOTA: Si el crecimiento en diámetro en los mejores sitios fuese 17 cm en 15 años, Ho no se habría rechazado como puede apreciarse en la imagen de la izquierda. Pues el valor de p calculado es 0,088; el cual es mayor que el p critico 0,01.

Por otro lado, si usted hubiese planteado de manera equivoca la hipótesis y hubiese elegido probar por la cola inferior; el resultado habría sido:

H_O: μ_0 = 15 cm (La media de la población es igual a 15cm)
$H_{A:}$ μ_0 < 15 cm (La media de la población es menor que 15cm)
Nivel de significancia es α= 0,01 (1%).



Si usted hubiese elegido probar por la cola inferior o sea preguntarse si el diámetro de la parcela es inferior al esperado para un bosque de 15 años en los mejores sitios de la Zona Norte, el resultado hubiese sido que no puede rechazar Ho; pues el valor de p calculado 0,999; el cual es mayor que el p critico 0,01.

Este ejemplo ilustra la importancia de plantear correctamente su hipótesis nula y alternativa.

4.4. Medias pareadas o dependientes

Los datos pueden proceder de la medición antes (pre-test) y después de aplicar un tratamiento a un grupo de sujetos (post-test) (e.g. persona, animal, proceso) o de un proceso de emparejamiento de muestras utilizando una variable control que tenga sentido para el caso en estudio como por ejemplo que las muestras procedan de la misma familia o por grupo de edad. Al primer caso se le conoce también como "mediciones repetidas" en tanto que al segundo como "muestras pareadas o emparejadas". La técnica de medidas repetidas compara diferencias al interior de sujetos y no entre sujetos por cuanto tendrá en general más poder que una prueba no pareada.

En el diseño pre y post tratamiento se mide la variable respuesta del grupo de interés antes y después de someterlo a un determinado tratamiento. En este caso la principal limitación es la ausencia de un grupo control independiente. En este diseño se asume que el efecto en la variable respuesta es atribuible al tratamiento aplicado al sujeto experimental y no a otros posibles factores explicativos.

Las diferencias de las medias se evalúan utilizando una prueba *t de estudiante pareada*. Un ejemplo de este tipo de diseño es la evaluación del cambio en la tasa de sedimentos de un cauce antes y después de un aprovechamiento forestal ó el cambio en los niveles de demanda bioquímica de oxígeno (DBO) como resultado de la aplicación de un programa de reducción de contaminantes orgánicos. Para diferencias normales, el estadístico de prueba es:

$t = \frac{\overline{X}_D - \mu_0}{s_D/\sqrt{n}}$, el cual tiene una distribución de t de Estudiante con n-1 grados de libertad. El subíndice "D" indica que la prueba se realiza para las diferencias y no para los valores originales, n es el tamaño de muestras (número de diferencias) y μ_0 el valor contra el cual se desea probar las diferencias. Por ejemplo, si se desea probar que no existe efecto de tratamiento μ_0 es igual a cero; en tanto que si se desea probar que el tratamiento tiene un efecto en la variable respuesta debe utilizarse un valor mayor o menor que cero.

Situación	Argumento	Planteamiento de contraste
Cambio en la tasa de sedimentos de un cauce antes y después de un aprovechamiento forestal.	Aprovechamiento forestal aumento la tasa de sedimentos.	H_O: $\mu_D = 0$ H_A: $\mu_D > 0$
Cambio en los niveles de demanda bioquímica de oxígeno (DBO) como resultado de la aplicación de un programa de reducción de contaminantes orgánicos	El tratamiento redujo la demanda bioquímica de oxígeno (DBO).	H_O: $\mu_D = 0$ H_A: $\mu_D < 0$

Desde el punto de vista estadístico, el grupo control es aquel que es comparable al grupo experimental pero que no recibe ningún tratamiento. Para la estadística moderna solo existe una forma de asegurar la comparabilidad entre el grupo control y el experimental: *la asignación aleatoria de sujetos a los grupos control y experimental*. De esta manera se espera que en promedio cualquier diferencia entre los grupos se compense y por tanto no sería necesario hacer ninguna suposición sobre sus diferencias o determinar su grado de similaridad previo al experimento. En los dos ejemplos mencionados previamente esto no es posible ya que el investigador(a) no tiene la libertad de asignar al azar un segmento del río a cada uno de los tratamientos.

Para ilustrar este tipo de prueba de hipótesis se utilizarán los datos del archivo peso_cola_ blanca.xlsx, el cual consigna el resultado hipotético de una dieta ingerida por 20 venados cola blanca en Cóbano, Península de Nicoya. La pregunta que se desea responder es la siguiente: ¿es la nueva dieta superior a la actual?

Prueba paramétrica

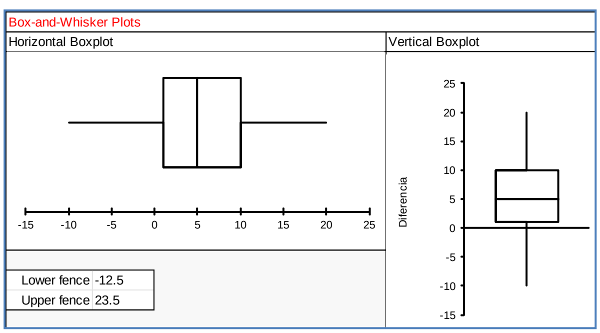
Hipotesis nula y alternativa

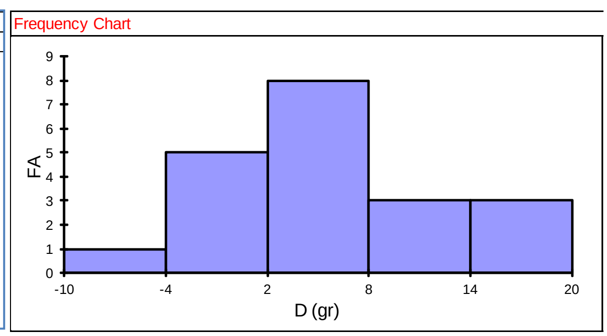
H_0: $\mu_D = 0$ (no diferencia en las dietas)
H^A: $\mu_D >$ (la nueva dieta es mejor que la actual)
Nivel de significancia: 0,01 (nivel de confianza 99%)
Procedimientoutilizando XlSTatistics
Lea el archivo peso_cola_ blanca.xlsx y seleccione la columna “diferencia” y seleccione 1Num. Recuerde que la prueba de hipótesis se realiza con las diferencias de peso y no con los valores originales.

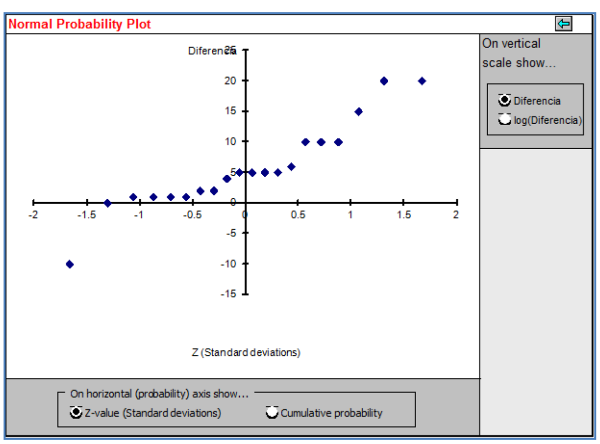
Estadísticos descriptivos para las diferencias

Numerical Summaries for Diferencia					
Number	20	Kurtosis	0.98537	Min	-10
Mean	5.65	10 % Tr mean	6.473684	Q_1	1
St Dev	7.073114	StdErr Mean	1.581596	Median	5
Coeff of Var	1.251879			Q_3	10
Skew	0.405872			Max	20

Análisis gráfico







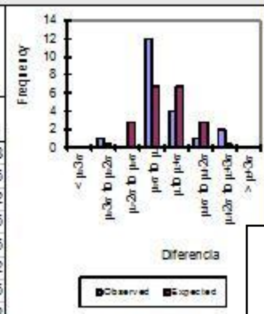
Prueba de normalidad

Goodness-Of-Fit Test for Normality of Diferencia

Data

Mean (μ)	5.65
St Dev (σ)	7.073114

	Proportions		Frequencies	
Range	Observed	Expected	Observed	Expected
< μ-3σ	0	0.00135	0	0.026998
μ-3σ to μ-2σ	0.05	0.0214	1	0.428005
μ-2σ to μ-σ	0	0.135905	0	2.718102
μ-σ to μ	0.6	0.341345	12	6.826895
μ to μ+σ	0.2	0.341345	4	6.826895
μ+σ to μ+2σ	0.05	0.135905	1	2.718102
μ+2σ to μ+3σ	0.1	0.0214	2	0.428005
> μ+3σ	0	0.00135	0	0.026998



Hypothesis Test

H_0: Population is normally distributed with the stated Mean and St Dev

H_1: Population is not normally distributed with the stated Mean and St Dev

Chisquare	15.48874
DF	7
p-value =	0.030242

Para un nivel de significancia 0,01 las diferencias en peso tienen una distribución normal. El valor de p calculado (0,030) es mayor que el valor de p crítico (0,01) y por lo tanto no se debe rechazar Ho.

Prueba de hipótesis

Hypothesis Tests

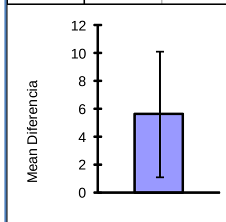
H_0: μ =	0
Alternative: ? > <	
H_1: μ >	0
T	3.57234
DF	19
p-value =	0.00102

Confidence level: Digite el nivel de confianza; en este caso 0.99.

Debe digitar el valor de μ en esta celda (en este caso 0 gr.

Confidence Intervals for μ

	Type (2,U,L)	2
	Confidence Level	0.99
ME	Lower	Upper
4.524844	1.125156	10.17484



Vertical axis title

Mean Diferencia

Decisión: El valor de p calculado (0,001) es menor que el valor de p crítico (0,01) y por lo tanto se rechaza Ho.

Nota: Otra forma de responder a lo planteado en la hipótesis nula es observar el ámbito del intervalo de confianza (1,15 gr < μ < 10,17 gr), el cual si no contiene el valor de μ (0 gr) indicando que Ho debe rechazarse.

Conclusión

Dado que el valor de p calculado (0.001**)** es menor que el p crítico (0,01) se concluye que la media de las diferencias es estadísticamente diferente de cero a un nivel de confianza de 99%, y por lo tanto se puede afirmar que la dieta tiene un efecto positivo en el peso de los venados.

Resumen:

- El promedio de las diferencias es 5.65 g con una desviación estándar de 7,07gr y un coeficiente de variación de 125%.
- Se observó una ganancia en el peso medio del grupo.
- La prueba de hipótesis se realizó utilizando un alfa de 0,001 (confianza de 99%).
- El paquete estadístico le brinda el valor de t crítico (valor a partir del cual se considera que las diferencias son significativas), el valor de t calculado así como la probabilidad asociado a dicho valor (valor de p). Si “p” calculado es menor que el nivel de significancia seleccionado se declara la prueba como significativa (o sea se rechaza Ho).
- Dado que en este caso p= 0.001 se declara como significativa la diferencia media en peso. Esto nos lleva a concluir que existe evidencia estadística para argumentar que el tratamiento tuvo un efecto en los sujetos experimentales. Note que la conclusión se refiere a la diferencia media y no a diferencias individuales.
- La conclusión anterior presupone que no existe ninguna otra posible razón, a parte de la dieta, que explique el cambio de peso en los venados. Como ejercicio se sugiere que usted liste todas las posibles variables que pueden influenciar los resultados obtenidos. Clasifique las variables como: 1) de efecto probable, 2) no probable y 3) aquellas que afectan a todos los venados por igual y las que podrían afectarlos en forma individual. ¿Cómo podría usted diseñar un experimento para asegurar que dichas variables no afecten sus resultados?

4.5. Bibliografía

Fry, J. C. 1996. Biological data analysis. A practical approach. Oxford University Press.418p.

Gómez-Gómez Manuel , Danglot-Banck Cecilia, Vega-Franco Leopoldo. 2003. Sinopsis de pruebas estadísticas no paramétricas. Cuándo usarlas.Revista Mexicana de Pediatría.Vo. 70, No. 2:91-99. Disponible en http://www.medigraphic.com/pdfs/pediat/sp-2003/sp032i.pdf. Visitado 28 junio 2012.

Hesterberg Tim. 2008. It's Time To Retire the "n >= 30" Rule. Proceedings of the American Statistical Association, Statistical Computing Section (CD-ROM). Disponible en http://home.comcast.net/~timhesterberg/articles/JSM08-n30.pdf. Visitado 20 junio 2012.

Hesterberg Tim, Moore David S., Monaghan Shaun, Clipson Ashley, Epstein Rachel and Craig Bruce A. 2007. Bootstrap Methods and Permutation Tests, Chapter 16 for Introduction to the Practice of Statistics, 6th edition, by David S. Moore, George P. McCabe and Bruce A. Craig, W. H. Freeman, N.Y. Disponible en http://bcs.whfreeman.com/ips6e/content/cat_040/pdf/ips6e_chapter16.pdf. Visitado 20 junio 2012.

Hesterberg, Tim C. 2004. Unbiasing the Bootstrap-Bootknife Sampling vs. Smoothing, Proceedings of the Section on Statistics and the Environment, American Statistical Association, 2924-2930. Disponible en http://home.comcast.net/~timhesterberg/articles/JSM04-bootknife.pdf. Visitado 20 junio 2012.

Hesterberg, Tim C. 2002. Performance Evaluation using Fast Permutation Tests. Proceedings of the Tenth International Conference on Telecommunication Systems, 465-474. Disponible en http://home.comcast.net/~timhesterberg/articles/Telecom02-permutation.pdf. Visitado 20 junio 2012

Hesterberg, Tim C. 2001. Bootstrap Tilting Diagnostics. Proceedings of the Statistical Computing Section (CD-ROM), American Statistical Association Disponible en http://home.comcast.net/~timhesterberg/articles/JSM01-diagnostics.pdf. Visitado 20 junio 2012.

Hesterberg, Tim C. 1999. Bootstrap Tilting Confidence Intervals and Hypothesis Tests. Computing Science and Statistics, 31, 389--393, Interface Foundation of North America, Fairfax Station, VA. Disponible en http://home.comcast.net/~timhesterberg/articles/Interface99-tiltingCI.pdf. Visitado 20 junio 2012.

Sawilowsky, S. 2002. "Fermat, Schubert, Einstein, and Behrens–Fisher:The Probable Difference Between Two Means When $\sigma^2_1 \neq \sigma^2_2$. Journal of Modern Applied Statistical Methods. 1(2): 461–472.

Shapiro, S. S. 1965. An analysis of variance test for normality (complete samples). Biometrika 52 (3-4): pp. 591–611. Dsiponible en http://sci2s.ugr.es/keel/pdf/algorithm/articulo/shapiro1965.pdf. Visitado 29 junio 2012. Visitado 28 junio 2012.

Snedecor, G.W. and Cochran, W.G. 1980. Statistical methods. Seventh Ed. Iowa, The Iowa State University Press. 507p.

Steel, R.G.D. y Torrie, J.H. 1980. Principles and procedures of McGraw-Hill. 629p.

4.6. Ejercicios

1. ¿Cuál es la diferencia entre un intervalo de confianza y una prueba de hipótesis?

2. ¿Qué relación existe entre el nivel de significancia, el nivel confianza y la potencia de una prueba de hipótesis?

3. Al realizar una prueba de hipótesis: ¿Cuándo debe utilizar una prueba paramétrica y cuando una no paramétrica?

4. ¿Qué se entiende por hipótesis nula e hipótesis alternativa?

5. Una investigadora desea determinar el poder de una prueba hipótesis t de Estudiante de rechazar Ho dado que sea falsa para una muestra con una media de 2224 mm (H1) y una población con los siguientes parámetros:

- Desviación estándar (SD sigma): 315 mm
- Media poblacional (μ): 1909 mm (Ho)

6. Utilizando los datos del archivo longitud_hojas_cm.xlsx realice lo siguiente:

   A. Plantee y realice una prueba de hipótesis para probar que la media de la población es igual a 16 cm.
   B. Realice otra prueba para probar que la media de la población es mayor que 13 cm.

7. Al establecer una parcela en el campo se aconseja que los datos obtenidos estén libres del efecto de borde. La teoría del efecto de borde indica que los árboles que se encuentran en el borde de la parcela tendrán mejores condiciones para crecer y por tanto su desempeño será mejor que los árboles en el centro de la parcela. Utilizando los datos del archivo efecto_borde.xlsx corresponden a una parcela de 10*10 árboles agrupada por datos de borde (primera fila y columna) y datos libres del efecto de borde (resto de los árboles), realice lo siguiente:

   A. Plantee y realice la prueba de hipótesis respectiva para determinar si los datos indican que existe un efecto de borde.
   B. ¿Cuál sería la conclusión si decidimos que las dos primeras filas y columnas son árboles de borde. Realice la respectiva prueba de hipótesis.

8. Al establecer una parcela en el campo se aconseja que los datos obtenidos estén libres del efecto de borde. La teoría del efecto de borde indica que los árboles que se encuentran en el borde de la parcela tendrán mejores condiciones para crecer y por tanto su desempeño será mejor que los árboles en el centro de la parcela. Los datos del archivo efecto_borde.xlsx corresponden a una parcela de 10*10 árboles agrupada por datos de borde (primera fila y columna) y datos libres del efecto de borde (resto de los árboles). Plantee y realice la prueba de hipótesis respectiva para determinar si los datos indican que existe un efecto de borde. ¿Cuál sería la conclusión si decidimos que las dos primeras filas y columnas son árboles de borde.

9. Utilizando los datos del archivo ppt_mm.xlsx responda a la siguiente pregunta: ¿Es la ciudad de Coronado más lluviosa que las inmediaciones del aeropuerto Juan Santa María? Basado en los resultados de la prueba de hipótesis podría usted argumentar que Coronado es más lluvioso que la ciudad de Alajuela y que la ciudad de Heredia?

10. Utilizando los datos del archivo ppt_5_estaciones.xlsx responda a las siguientes preguntas:

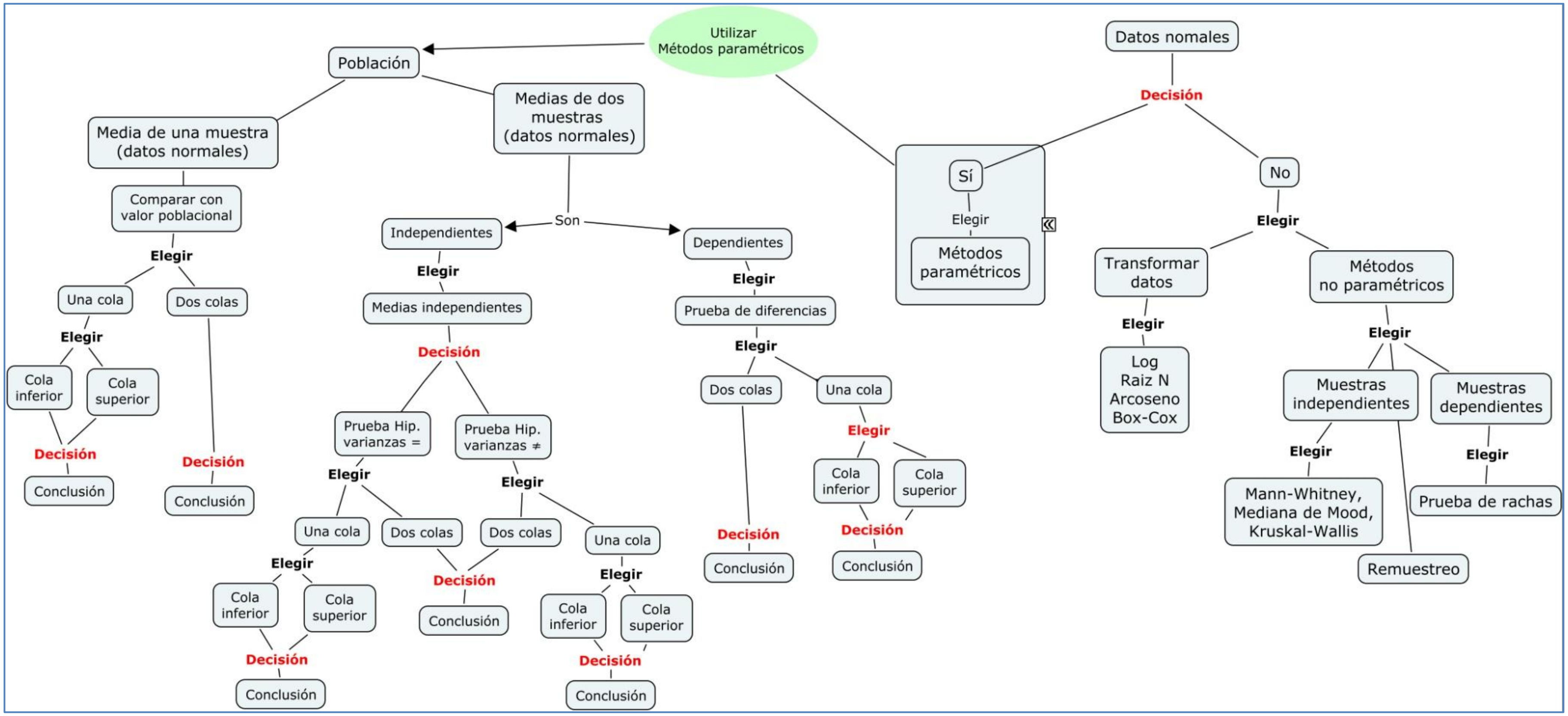
A. Basado en una inspección visual de los datos ¿Cuál es la ciudad más lluviosa? ¿Por qué?
B. Basado en un análisis gráfico ¿cuáles ciudades reciben en promedio la misma cantidad de lluvia anual?
C. ¿Cómo agruparía usted las estaciones según su precipitación media anual? ¿Por qué?
D. ¿Tendría sentido utilizar la media de la precipitación para describir las condiciones climáticas de la zona que cubre las estaciones?

11. Suponga que usted debe diseñar un estudio para evaluar el efecto de un nuevo sistema de transplante de plántulas de almendro en un vivero de la Zona Norte. Plantee un diseño estadístico para dicho estudio. Liste todas las variables que usted considera que deben monitorearse. Clasifique las variables como críticas, muy importantes, poco importantes y no importantes. Usted cuenta con un presupuesto de US$1.000 para realizar el estudio ¿cómo asignaría dicho dinero al monitoreo de las variables?.

Anexo 1: Guía para el análisis de datos

1. Lea el material de referencia (teoría); asesórese con otros(as) colegas.
2. Lea la pregunta o las instrucciones y responda a lo siguiente:
   - ¿Cuáles son las variables a analizar y cuál es su nivel de medición?
   - ¿Cuáles es el contexto de los datos? ¿Población?
   - ¿Cuál es el producto esperado o solicitado?
   - ¿Qué se desea resaltar del set de datos? ¿Para qué y por qué analiza usted el set de datos?
   - Liste los métodos de análisis estadístico que puede utilizar.
   - ¿Cuáles son los supuestos de cada método de análisis estadístico? ¿Cómo los puedo probar?
   - En caso de no cumplir con los supuestos; ¿cuáles son las alternativas de análisis?
   - ¿Cuál software puede utilizar para realizar el análisis?
   - Una vez realizado en análisis estadístico ¿Cuales son los argumentos estadísticos y disciplinarios (e.g. biológicos, agronómicos, forestales) o de otra índole que le permiten explicar las conclusiones obtenidas? Ej. tamaño de muestra, muestreo sesgado, efecto de confusión (variables no medidas), nivel de significancia utilizado, condiciones ambientales no normales (i.e. El Niño, La Niña)
   - ¿Cuál sería su recomendación final (acción)?

Anexo 2: Prueba de hipótesis: flujograma



Anexo 3: Guía para prueba de hipótesis

1. *Análisis gráfico*
   El objetivo del análisis gráfico es detectar patrones o tendencias en el set de datos. Por ejemplo, se puede analizar la tendencia central, la variabilidad y la forma de la distribución que caracteriza al set de datos. Los gráficos de Box-Whisker y de barra de errores (desviación estándar, error estándar, intervalo de confianza) son apropiados para visualizar el comportamiento de dos o más sets de datos. Cuando se desea evaluar la normalidad de los datos puede utilizarse el diagrama de probabilidad normal.

2. *Cálculo de estadísticos descriptivos*
   Los estadísticos descriptivos resumen lo relevante de los datos en términos de tendencia central, variabilidad y forma de la distribución. Normalmente se calcula el promedio, la desviación estándar, el coeficiente de variación y los coeficientes estandarizados de curtosis y asimetría.

3. *Prueba de normalidad*
   Dado que las pruebas de hipótesis requirieren normalidad en los datos se debe someter a prueba la siguiente hipótesis:
   Ho: Los datos son normales
   Ha: Los datos no son normales
   Recuerde elegir su alfa antes de realizar la prueba de hipótesis.
   Decisión: rechazar Ho si el valor de P calculado es menor que el valor de P critico (alfa)

Nota: En caso de rechazar Ho debe transformar los datos y realizar nuevamente la prueba de hipótesis. Para variables positivas se pueden utilizar las transformaciones logarítmica y raíz cuadrada y para datos sin ceros el inverso (1/x). La transformación de Box-Cox es una familia de transformaciones definida como: , donde Y es la variable respuesta y lambda (λ) es el parámetro de la transformación. Para lambda = 0 la transformación es igual a utilizar el logaritmo natural de los datos. En caso de no lograr normalidad utilizar una prueba no paramétrica ó alguna técnica de remuestreo.

4. *Prueba de hipótesis (una muestra independiente)*
   Una vez que usted se ha familiarizado con el set de datos y que ha probado por el supuesto de normalidad puede proceder a realizar la prueba de hipótesis. El proceso involucra los siguientes pasos:
   A. Plantear prueba de hipótesis a realizar
      La prueba puede ser de dos colas
      Ho: La media es igual a un valor dado
      Ha: La media no es igual a un valor dado
      Recuerde elegir su alfa antes de realizar la prueba de hipótesis

La prueba puede ser de una cola
Ho: La media es igual a un valor dado
Ha: La media es mayor que a un valor dado (cola superior)
Ha: La media es menor que a un valor dado (cola inferior)

Solo puede plantear una hipótesis alternativa

Recuerde elegir su alfa antes de realizar la prueba de hipótesis

B. Definir nivel de significancia (alfa)
C. Efectuar la prueba de hipótesis
D. Tomar una decisión
E. Conclusión estadística y practica

5. *Prueba de hipótesis (dos muestras independientes)*

Una vez que usted se ha familiarizado con el set de datos y que ha probado por el supuesto de normalidad puede proceder a realizar la prueba de hipótesis. El proceso involucra los siguientes pasos:

A. Plantear las hipótesis nula y alternativa

a. **Prueba de igualdad de varianzas**: Cuando realice una prueba de hipótesis de dos muestras independientes debe realizar primero la prueba de *igualdad de varianzas*
Ho: Las varianzas son iguales
Ha: Las varianzas son diferentes

Recuerde elegir su alfa antes de realizar la prueba de hipótesis
Decisión: rechazar Ho si el valor de P calculado es menor que el valor de P critico (alfa)

b. Prueba de medias

La prueba puede ser de dos colas
Ho: Las medias son iguales
Ha: Las medias son diferentes

Recuerde elegir su alfa antes de realizar la prueba de hipótesis
Al realizar la prueba de hipótesis debe elegir con varianzas iguales ó diferentes (acorde con conclusión de punto a)

La prueba puede ser de una cola
Ho: Las medias son iguales
Ha: Una de las medias es mayor (cola superior)
Ha: Una de las medias es menor (cola inferior)

Solo puede plantear una hipótesis alternativa

Recuerde elegir su alfa antes de realizar la prueba de hipótesis
Al realizar la prueba de hipótesis debe elegir con varianzas iguales ó diferentes (acorde con conclusión de punto a)

B. Definir nivel de significancia (alfa)
C. Efectuar la prueba de hipótesis
D. Tomar una decisión
E. Conclusión estadística y practica

6. *Prueba de hipótesis (dos muestras dependientes o pareadas)*
Una vez que usted se ha familiarizado con el set de datos y que ha probado por el supuesto de normalidad puede proceder a realizar la prueba de hipótesis. El proceso involucra los siguientes pasos:

a. **Prueba de medias pareadas**
La prueba puede ser de dos colas
Ho: La media de las diferencias es igual a cero
Ha: La media de las diferencias es diferente de cero

Recuerde elegir su alfa antes de realizar la prueba de hipótesis

La prueba puede ser de una cola
Ho: La media de las diferencias es igual a cero
Ha: La media de las diferencias es mayor que cero (cola superior)
Ha: La media de las diferencias es menor que cero (cola inferior)

Solo puede plantear una hipótesis alternativa

Recuerde elegir su alfa antes de realizar la prueba de hipótesis

b. Definir nivel de significancia (alfa)
c. Efectuar la prueba de hipótesis
d. Tomar una decisión
e. Conclusión estadística y practica

Anexo 4: Fórmulas

Nombre	**Fórmula**	**Supuestos**
Prueba Z. Una muestra.	$z = \dfrac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$	Población Normal ($n \geq 30$) σ conocida
Prueba Z. Dos muestras	$z = \dfrac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	Población Normal obs. independientes σ1 y σ2 conocidas
Prueba t. Una muestra.	$t = \dfrac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$, $gl = n - 1$	Población Normal ($n \geq 30$) σ desconocida
Prueba t varianzas iguales. Dos muestras independientes	$t = \dfrac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$, $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$, gl = $n_1 + n_2 - 2$	Población Normal ó *n1* + *n2* > 40) observaciones independientes σ1 = σ2 (σ1 y σ2 desconocidas)
Prueba t varianzas no iguales. Dos muestras independientes	$t = \dfrac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$, $df = \dfrac{(n_1 - 1)(n_2 - 1)}{(n_2 - 1)c^2 + (n_1 - 1)(1 - c^2)}$, $c = \dfrac{\frac{s_1^2}{n_1}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ **o** $df = \min\{n_1, n_2\}$	Población Normal o *n1* + *n2* > 40) observaciones independientes σ1 ≠ σ2 **y** (σ1 y σ2 desconocidas)
Prueba t pareada	$t = \dfrac{\overline{d} - d_0}{s_d}$, $df = n - 1$	Población Normal de diferencias **o** *n* > 30) **y** σ desconocidas
Prueba Z una muestra (proporciones)	$z = \dfrac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$	$np > 10$ **y** $n(1 - p) > 10$
Prueba Z dos muestras (proporciones) varianzas iguales	$z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1 - \hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$ $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$	n1p > 5 Y *n1*(1 − *p1*) > 5 **y** *n2p2* > 5 **y** *n2*(1 − *p2*) > 5 **y** observaciones independientes
Prueba Z dos muestras (proporciones) varianzas diferentes	$z = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$	n1p > 5 Y *n1*(1 − *p1*) > 5 **y** *n2p2* > 5 **y** *n2*(1 − *p2*) > 5 **y** observaciones independientes