ANÁLISIS ESTADÍSTICO

PRUEBA DE HIPOTESIS

Jorge Fallas
jfallas56@gmail.com

2010

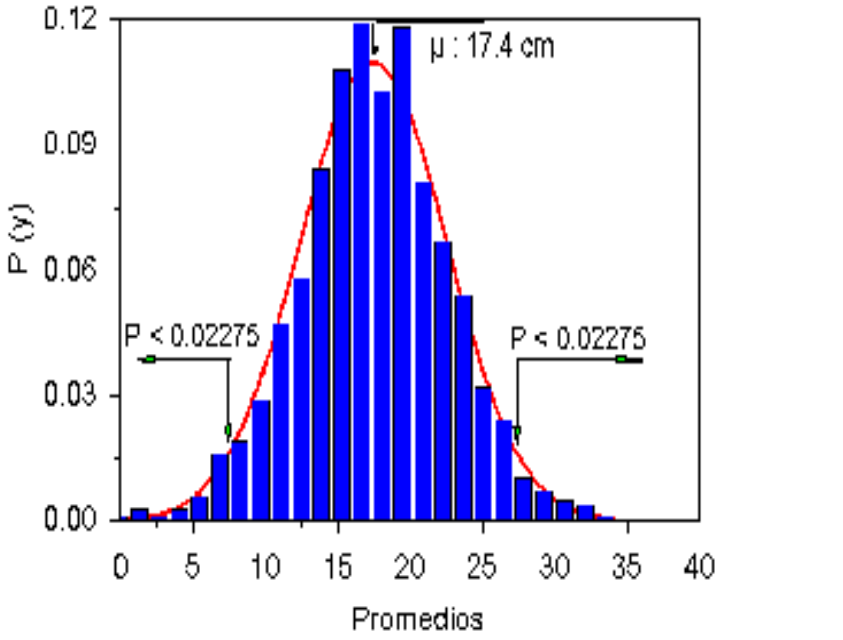
Temario

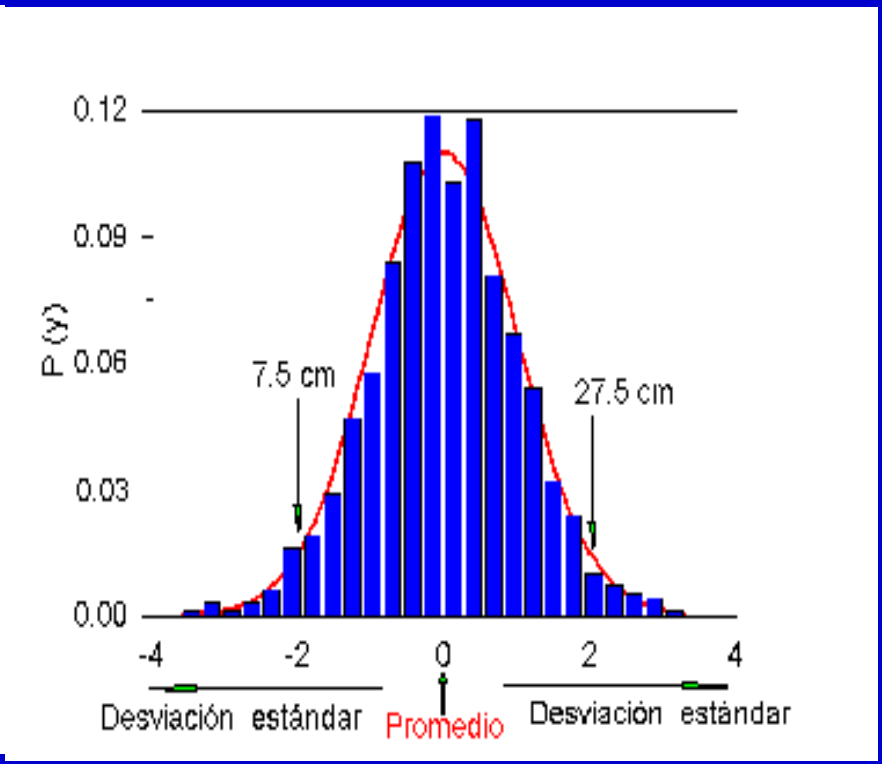
- Datos experimentales y distribuciones de referencia
- Una media poblacional
- Hipótesis nula, alternativa y nivel de significancia
- Prueba de hipótesis direccionada y no direccionada
- Comparación de dos medias
  - medias independientes
  - Medias pareados o dependientes
- Prueba de normalidad
- Prueba de igualdad de varianzas
- Herramientas
  - XlSTats
  - Instat
  - Remuestreo

Introducción

- ¿Cómo determinar si un valor muestral es igual, diferente, mayor ó menor que un valor poblacional?
- Estimación puntal y estimación por intervalos
- Conceptos de nivel de confianza y nivel de significancia
- Existen métodos estadísticos que nos permiten realizar inferencias a partir de muestras
- Deducción: Investigador(a) no posee datos sobre la totalidad de la población y lo que hace es generalizar la conclusión obtenida a partir de una muestra a aquellos individuos no medidos o estudiados
- Opciones de análisis
  - Pruebas paramétricas
  - Pruebas no paramétricas
  - Remuestreo (Ej. Resampling Stats)
    - http://www.uvm.edu/~dhowell/StatPages/Resampling/Resampling.html

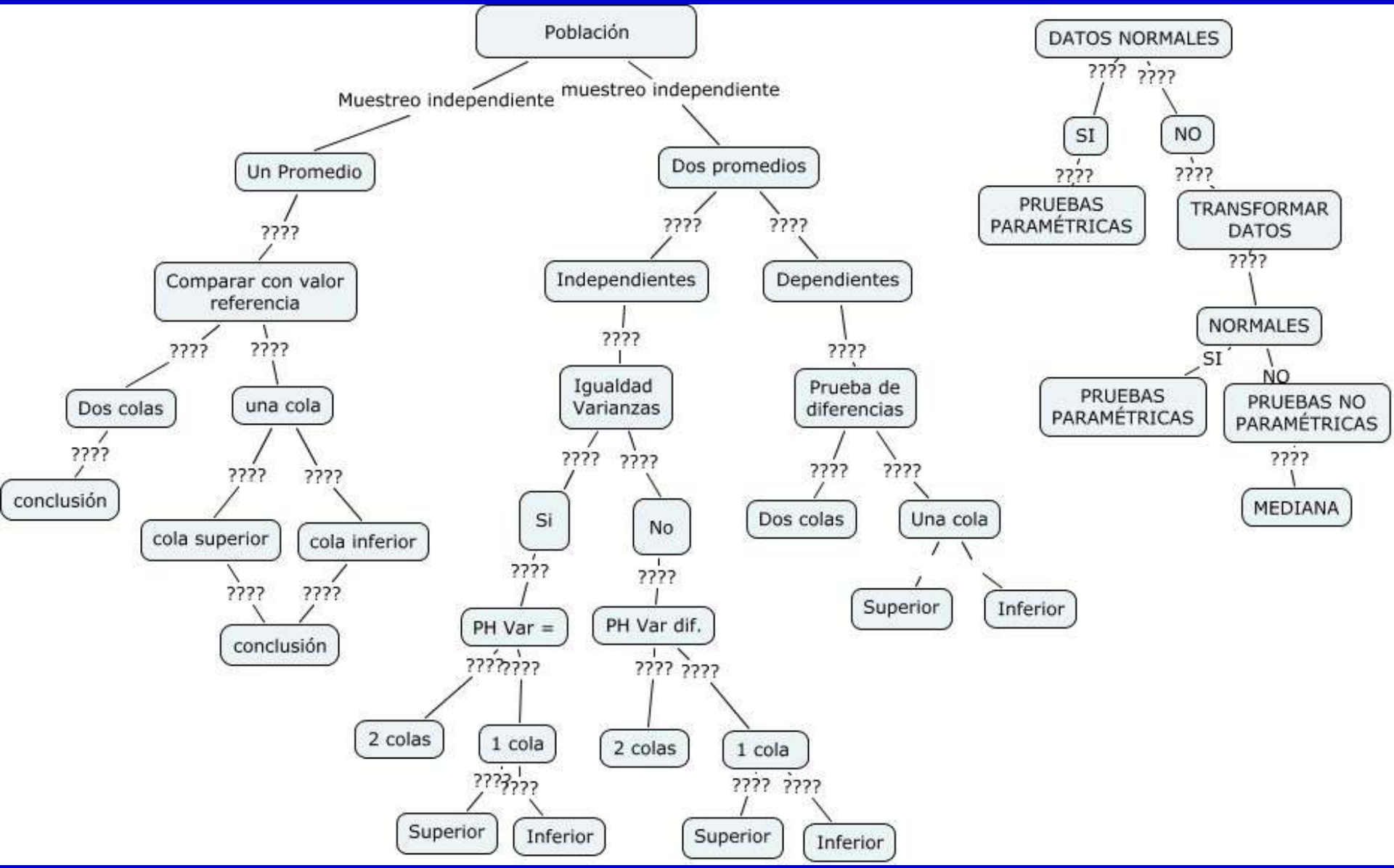
Datos experimentales y distribuciones teóricas





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 17.4cm y una varianza de 25 cm2. Cualquier parcela 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.

Comparación de medias



Hipótesis nula, alternativa y nivel de significancia

- 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.
- Hipótesis alternativa (Ha) decisión alternativa que se ofrece cuando se rechaza Ho.
- La Ho es sometida a prueba en lugar de la hipótesis alternativa (Ha) porque la muestra provee los datos que permiten estimar los parámetros de su distribución muestral; en tanto que Ha no ofrece esta posibilidad
- Ej. Ho:$\mu = 8.7$ cm, Ha: $\mu \neq 8.7$ cm
- El *nivel de significancia* se designa con la letra griega α e indica cuan rara deber ser la diferencia entre la media muestral y la poblacional como para rechazar la hipótesis nula dado que ésta sea correcta.
- Alfa 0.1 *diferencia significativa* y alfa 0.001*diferencia altamente significativa*
- Concepto de significancia estadística Vs. significancia practica

Prueba de hipótesis direccionada y no direccionada

- PH no direccionada estamos interesados en determinar si el valor muestral es diferente al valor poblacional; sin embargo no interesa saber sí el valor es mayor o menor que el parámetro.
- Este tipo de pruebas se denomina también de *dos colas*; ya que Ho se rechazará si el valor muestral se ubica *por encima o por debajo del valor del parámetro.*
- PH direccionada estamos interesados en determinar si el valor muestral es mayor ó menor que el valor poblacional
- Ejemplo norma o estándar
- Recuerde usted debe plantear su hipótesis nula y alternativa antes de realizar la colecta de sus datos. No debe permitir que los datos dicten el tipo de hipótesis a realizar.

Una población

¿Es la media poblacional (μ) igual a un valor determinado?

Estadístico de prueba, t de Estudiante

$$to = \frac{\overline{X} - \mu}{S/n^{0,5}}$$

en donde:

S= desviación estándar, $\overline{X}$: promedio muestral
n: tamaño de muestra

Se desconocen los valores de μ y σ

SUPUESTOS PRUEBA PARAMÉTRICA:
Muestreo aleatorio
Población normal

PH una población: XLStats

• Seleccione del menú principal la opción

1	1 Numerical	1Num

Data
X
12
15.5

Inserte los datos en la columna

•Seleccione la hoja tests

Information | Data & Description | Summaries | Tests | Extra Tools

Configure la ventana de diálogo como se muestra a continuación:

Tests on the Mean (μ) (t-tests)

Sample Data

Sample Size	50
Mean	14.228
Standard Deviation	1.26023
SE Mean	0.178223

Hypothesis Tests

H_0: μ =	15
Alternative	≠ > <
H_1: μ ≠	15
T	-4.3316
DF	49
p-value =	7.3E-05

Confidence Intervals for μ

Type (2,U,L)	2	
Confidence Level	0.95	
ME	Lower	Upper
0.358153	13.86985	14.58615

Esta prueba se conoce como de ***dos colas*** ya que no nos interesa saber si la media es mayor ó menor que el valor de referencia. La probabilidad de rechazar Ho se divide por igual en ambas colas de la distribución. En términos cotidianos esto implica que estamos dispuestos a aceptar que la diferencia sea positiva ó negativa. Cuando se realiza una prueba de *una cola* nos interesa saber si la media es mayor que el estándar (diferencia positiva) ó si por el contrario es menor (diferencia negativa).

Observe que Ho es μ = 15 cm y que Ho es μ ≠ 15 cm. Esto indica que se hará una prueba de dos colas (Ho ≠ 15.0).

La decisión de realizar una prueba de una ó dos colas debe tomarse antes de colectar los datos.

El nivel de significancia (α) es 0.05

El valor de "p" calculado es 0.000073.

Conclusión: rechazar Ho.

Resultados

Prueba de hipótesis para diámetros (cm)
Sección 1: cálculo de estadísticos muestrales
Sample mean = 14.2735
Sample median = 14.2

Sección 2: Plantear prueba de hipótesis
t-test
La decisión de aplicar una prueba de una ó dos colas debe realizarse antes de colectar los datos
Hipótesis nula: media = 15.0 (la media de la población es igual a 15cm)
Hipótesis alternativa: media diferente mu <> 15.0

Sección 3: Cálculo de estadístico de prueba
Estadístico "t" = -4.13085 P-Value = 0.000143724

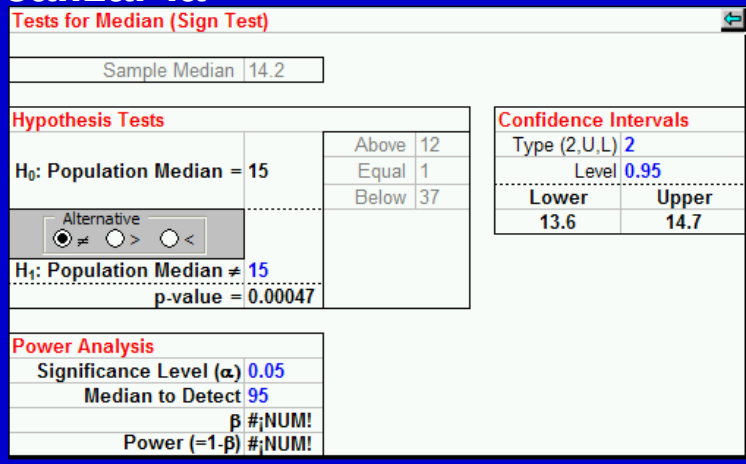
Sección 4: Conclusión estadística
Rechazar Ho a un nivel de significancia de 0.05

Conclusión: Dado que el valor de P calculado (0.00014) es menor que el valor del alfa seleccionado (0.05), se rechaza la hipótesis nula con un nivel de confianza de 95% y se concluye que el valor muestral es diferente al valor poblacional.

Equivalentes no paramétricos: **Sign test**

Sign test : Esta prueba utiliza el número de observaciones que se encuentran sobre y bajo la *mediana* para realizar la prueba de hipótesis
Ha: mediana = 15.0 cm
Ho: mediana diferente de 15.0 cm
Número de observaciones bajo la mediana : 37
Número de observaciones sobre la mediana : 12

Valor estimado de "p" 0.00047



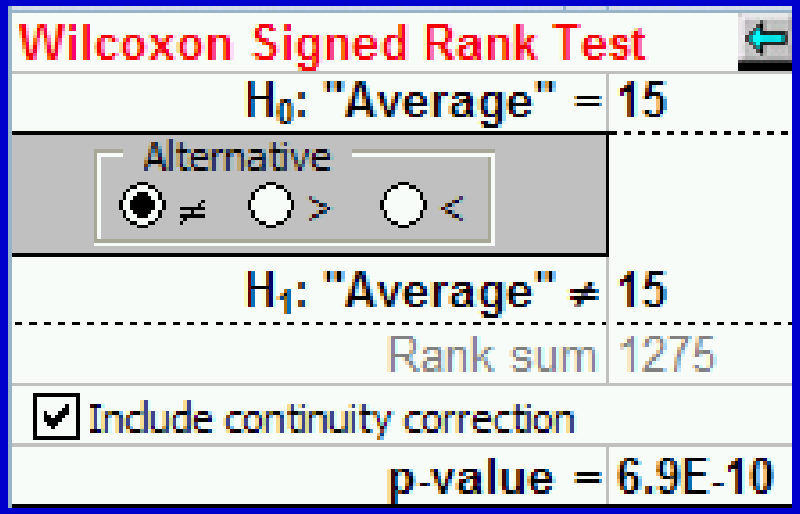
Conclusión: Dado que el valor de P calculado (0.00047) es menor que el valor del alfa seleccionado (0.05), se rechaza la hipótesis nula con un nivel de confianza de 95% y por tanto la mediana muestral es diferente de la mediana de la población de referencia.

Observe que se llega a la misma conclusión que con la prueba paramétrica.

Equivalentes no paramétricos: **Wilcoxon Signed rank test**

Signed rank test: Esta prueba utiliza los rangos medios de los valores que se encuentran sobre y bajo la *mediana* para realizar la prueba de hipótesis

Ha: mediana = 15.0 cm
Ha: mediana diferente de 15.0 cm
Valor estimado de "p" 0.000000000069

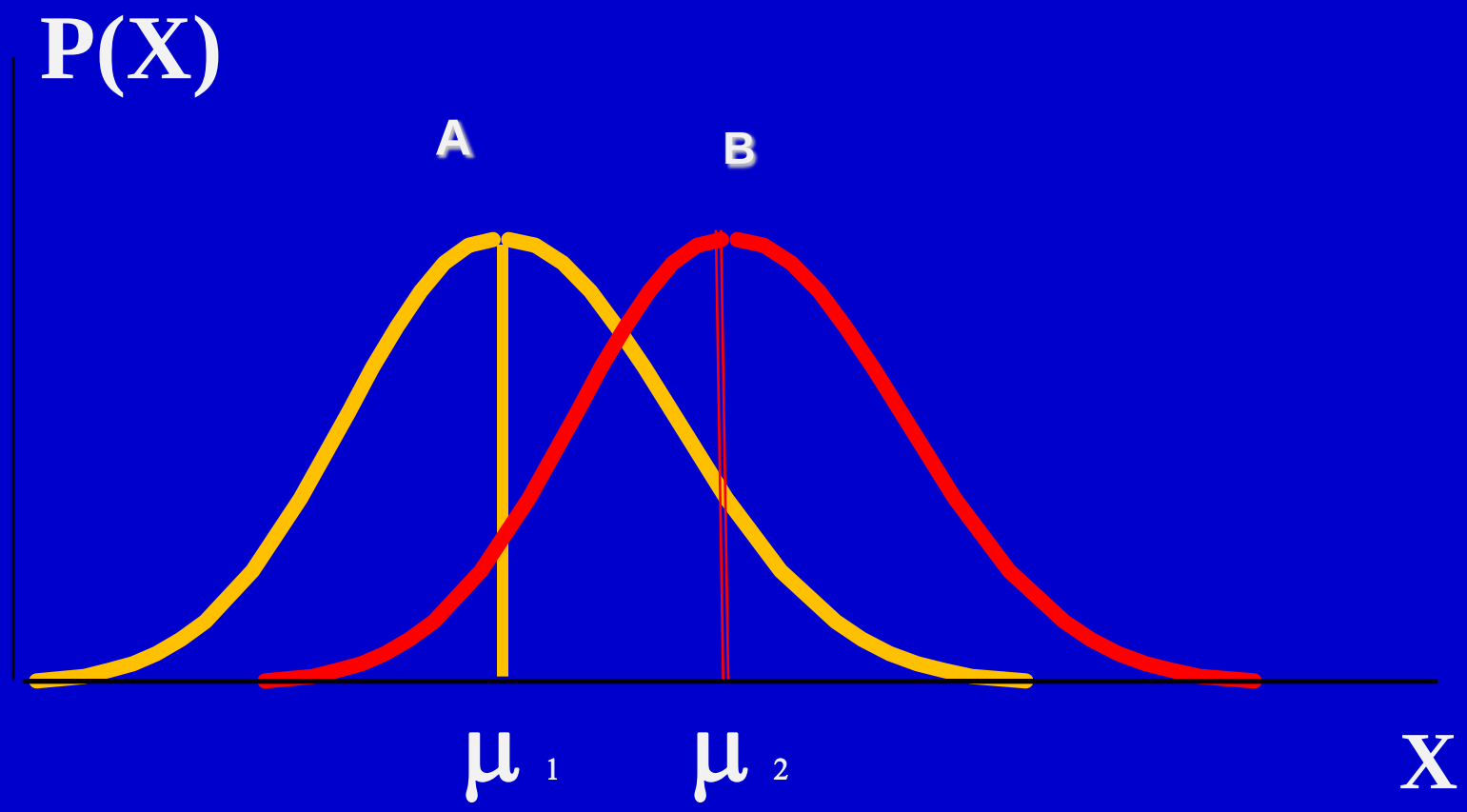


Conclusión: Dado que el valor de P calculado (0.000000000069) es menor que el valor del alfa seleccionado (0.05), se rechaza la hipótesis nula con un nivel de confianza de 95% y por tanto la mediana muestral es diferente de la mediana de la población de referencia.

Observe que se llega a la misma conclusión que con la prueba paramétrica.

Comparación de dos medias

- *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.
- *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, la desviación estándar, el coeficiente de variación y los coeficientes de curtosis y asimetría.
- Prueba de normalidad de los datos: Las pruebas paramétricas asumen que los datos son normales.
- *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:
  - Plantear la hipótesis nula y alternativa
  - Seleccionar el estadístico de prueba y definir el nivel de significancia
  - Efectuar la prueba de hipótesis (realizar primero prueba de igualdad de varianzas)
  - Tomar una decisión
  - Proponer a una conclusión y explicación



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, más difícil será detectar una diferencia significativa entre las poblaciones.

PH dos poblaciones: Prueba “t” en XLStats

La hoja Tests - 2 Categories **del menú** 2 | 1 Numerical, 1 Categorical | 1Num1Cat **le permite someter a prueba la siguiente hipótesis:**

Ho: $\mu_1 = \mu_2$
Ha: $\mu_1 \neq \mu_2$
Varianzas: iguales (*Assume equal standard deviations*)

Recuerde que antes de realizar la prueba de hipótesis de dos medias independientes usted debe realizar la prueba de igualdad de varianzas (*F-Test for Variance*).

Ho: $\sigma^2_1 / \sigma^2 = 1$
Ha: $\sigma^2_1 / \sigma^2 \neq 1$

F-Test for Variance

Sample Data			
n_1	50	n_2	50
s_1^2	34.44776	s_2^2	49.84643

$H_0 : \sigma_1^2 = \sigma_2^2$
$H_1 : \sigma_1^2 \neq \sigma_2^2$

F	0.691078
p-value =	0.199469

El valor de p calculado es 0.1994 mayor que el valor de p critico (0.05) y por tanto no se rechaza Ho y se concluye que las varianzas son iguales.

Tests for comparing two categories

Categories — Cat. 1: A — Cat. 2: B

Two-Sample t-tests (Differences Between Means, μ)

Sample Data			
n_1	50	n_2	50
$\bar{x}_1$	17.214	$\bar{x}_2$	18.136
s_1	5.869221	s_2	7.060201

Assume equal standard deviations

$\bar{x}_1 - \bar{x}_2$	-0.922
SE Difference	1.298416

Hypothesis Tests

H_0: $\mu_1 - \mu_2 = 0$
Alternative: ≠ > <
H_1: $\mu_1 - \mu_2 \neq 0$

T	-0.7101
DF	98
p-value =	0.479331

Confidence Intervals for $\mu_1 - \mu_2$

Type (2,U,L) 2
Level 0.95

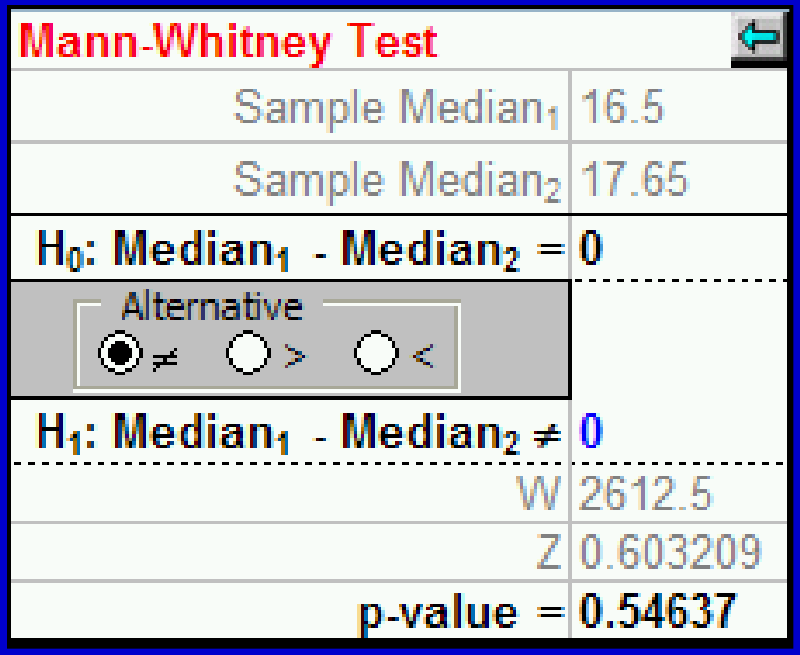
ME	Lower	Upper
2.576664	-3.49866	1.654664

Mann-Whitney Test (Differences Between Medians)
F-Test for Variance
Randomised 2-Group Test
Power Analysis Sample Size Determination
Residuals Analysis

Conclusión: Dado que el valor de P calculado (0.47) es mayor que el valor del alfa seleccionado (0.05), no se rechaza la hipótesis nula con un nivel de confianza de 95% y por tanto las medias muestrales no son diferentes.

Comparación de Medianas: Prueba de Mann-Whitney

La prueba de Mann-Whitney es el equivalente no parámetrico de la prueba de "t" de Estudiante



Median of sample 1: 16.5 (Zona Norte)
Median of sample 2: 17.65 (Zona Sur)

Mann-Whitney (Wilcoxon) W test para comparar medianas

Hipótesis nula : mediana1 = mediana2
Hipótesis Alt. : mediana1 ≠ mediana2

W = 2612.5 P-value = 0.54637

Conclusión:
La prueba de Mann-Whitney compara las medianas de dos muestras. Dado que el valor de "p" calculado (0.548637) es mayor que el "p" critico (0.05) no se rechaza Ho y se concluye que las medianas de ambas parcelas son iguales a un nivel de confianza de 95%. En este caso las muestras no proveen suficiente evidencia estadística como para suponer que las parcelas presentan un crecimiento diferente.

Prueba de igualdad de varianzas

- Antes de aplicar la prueba “t” para dos medias independientes es necesario determinar sí las varianzas son iguales ó diferentes; este es otro ejemplo de una prueba de dos colas.
- En este caso se utiliza el estadístico F = Varianza 1/ Varianza 2; el cual tiene una distribución de F con n1 y n2 grados de libertad.

F-Test for Variance

Sample Data			
n_1	50	n_2	50
s_1^2	34.44776	s_2^2	49.84643

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

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

F	0.691078
p-value =	0.199469

El valor de p calculado es 0.1994 mayor que el valor de p critico (0.05) y por tanto no se rechaza Ho y se concluye que las varianzas son iguales.

Pruebas de normalidad

- W de Shapiro-Wilk
- Considerada una de las pruebas mas fidedignas o imparciales.
- D de Kolmogorov-Smirnov
- Prueba de bondad de ajuste de la distribución normal
- Lillifors
- Versión modificada de la D de Kolmogorov-Smirnov

- Ho: los datos provienen de una distribución normal
- Ha: los datos no provienen de una distribución normal
- Ver Gráfico de probabilidad normal

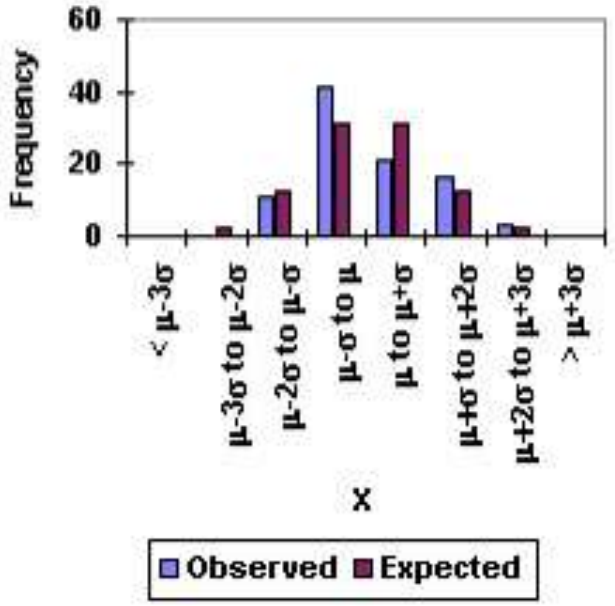
XLStats: prueba normalidad

Goodness-Of-Fit Test for Normality of X

Data

Mean (μ)	96.91304
St Dev (σ)	10.96613

	Proportions		Frequencies	
Range	Observed	Expected	Observed	Expected
< μ-3σ	0	0.00135	0	0.124191
μ-3σ to μ-2σ	0	0.0214	0	1.968822
μ-2σ to μ-σ	0.119565	0.135905	11	12.50327
μ-σ to μ	0.445652	0.341345	41	31.40372
μ to μ+σ	0.228261	0.341345	21	31.40372
μ+σ to μ+2σ	0.173913	0.135905	16	12.50327
μ+2σ to μ+3σ	0.032609	0.0214	3	1.968822
> μ+3σ	0	0.00135	0	0.124191



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	10.29499
DF	7
p-value =	0.172464

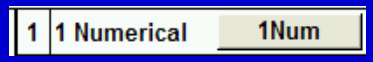
<u>Conclusión</u>: El valor de todos los P calculados es >0.1, por tanto con una confianza 90% se concluye que la variable proviene de una distribución normal

Medias pareadas o dependientes

- 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.
- Se asume que el efecto en la variable respuesta es atribuible al tratamiento aplicado al sujeto experimental y no a otros factores.
- Los datos se analizan utilizando una prueba *t de estudiante pareada*. Se somete a prueba la media de las diferencias y no las observaciones originales.
- Lo sujetos deben *asignarse al grupo control y al grupo testigo de manera aleatorio*. Esto no es posible en un estudio de pre u post tratamiento.

Medias pareadas

•Utilice del menú principal de XLStat la opción



Data
diferencia
5
5

Inserte las diferencias en la columna de datos

Recuerde para realizar esta prueba usted debe calcular primero las diferencias (antes-después)

•Seleccione la hoja tests

•Configure la ventana de diálogo como se muestra a continuación:
Ho: la media de las diferencia es igual a cero (no efecto de la dieta)
Ha: la media de las diferencia es diferente de cero (existe efecto de la dieta)
Nivel de significancia: 0.05
•Valor de "t": 3.57234
•p-value = 0.00203

	A	B	C	D	E	F	G
1	Tests on the Mean (μ) (t-tests)						
2							
3	Sample Data						
4			Sample Size		20		
5			Mean		5.65		
6			Standard Deviation		7.073114		
7			SE Mean		1.581596		
8							
9	Hypothesis Tests				Confidence Intervals for μ		
10			H_0: μ =		0	Type (2,U,L)	2
11	Alternative					Confidence Level	0.95
12	◉ ≠ ○ > ○ <				ME	Lower	Upper
13			H_1: μ ≠		0	3.310319	2.339681	8.960319
14			T		3.57234		
15			DF		19		
16			p-value =		0.00203		

Numerical Summaries for diferencia			
Number	20	Min	-10
Mean	5.65	Q_1	1
St Dev	7.073114	Median	5
Coeff of Var	1.251879	Q_3	10
Skew	0.405872	Max	20

Conclusión

Dado que el valor de P (0.00203) es menor que 0.05 se concluye que la media de las diferencias es estadísticamente diferente de cero a un nivel de confianza de 95%, y por tanto se puede afirmar que la dieta tiene un efecto en el peso de los venados. Sin embargo dado que se realizó una prueba de 2 colas no es posible afirmar que la diete aumentó ó redujo el peso de los venados. Para responder a esta última pregunta debemos realizar una prueba de 1 cola. Recuerde que la hipótesis a probar debe establecerse antes de realizar el experimento.

SÍNSTESIS

- 1. Lea el material de referencia (teoría); asesórese con otros (as) colegas
- 2. Lea 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? ¿Muestreo?
- ¿Cuáles son las variables a analizar y cuál es su nivel de medición?
- ¿Cuál es el producto esperado o solicitado?
- ¿Qué se desea resaltar del set de datos? ¿Para qué está usted analizando los datos?
- ¿Cuáles son los métodos de análisis estadístico que puedo utilizar? (Gráficos, tablas, pruebas de hipótesis, intervalos de confianza)
- ¿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 puedo utilizar para realizar el análisis?
- Una vez realizado en análisis estadístico ¿Cuáles son los argumentos estadísticos y biológicos o de otra índole que me permiten explicar las conclusiones obtenidas? Ej. tamaño de muestra, muestreo sesgado, confusión en variables explicativas, nivel de confianza utilizado, condiciones ambientales no normales (Ej. El Niño, La Niña)
- ¿Cual sería su recomendación final (acción)? ¿Otro estudio?