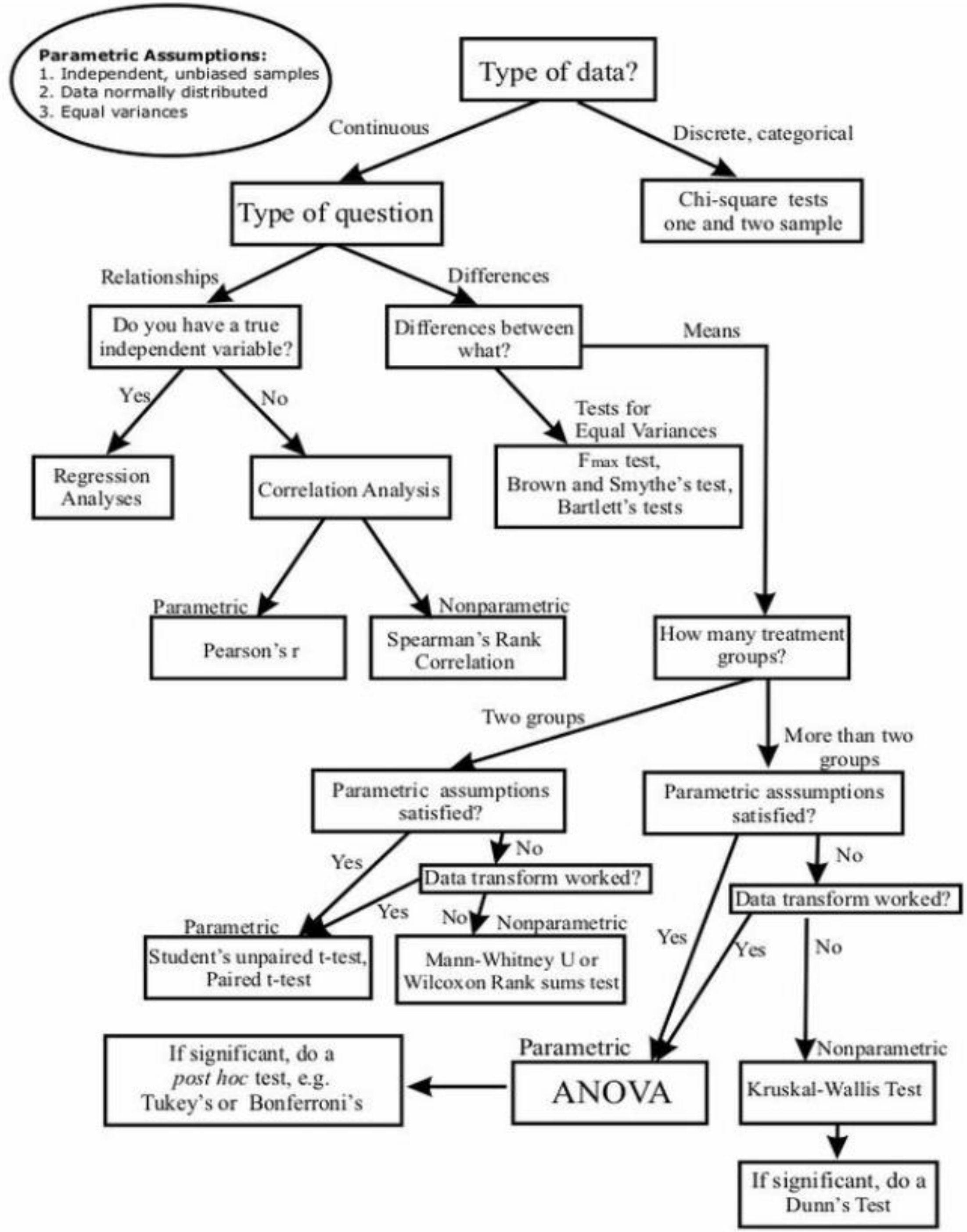
Diseño de experimentos - prueba de hipótesis

http://www.academia.utp.ac.pa/humberto-alvarez/diseno-de-experimentos-y-regresion

Inferencia estadística

- Conjunto de métodos y técnicas que permiten inducir, a partir de la información empírica proporcionada por una muestra, cual es el comportamiento de una determinada población con un riesgo de error medible en términos de probabilidad.
- Los métodos paramétricos de la inferencia estadística se pueden dividir, básicamente, en dos: métodos de estimación de parámetros y métodos de contraste de hipótesis.
- Ambos métodos se basan en el conocimiento teórico de la distribución de probabilidad del estadístico muestral que se utiliza como estimador de un parámetro.

Análisis estadísticos típicos



Prueba de hipótesis

- Tienen como objetivo comprobar si determinado supuesto referido a un parámetro poblacional, o a parámetros análogos de dos o más poblaciones, es compatible con la evidencia empírica contenida en la muestra.
- Los supuestos que se establecen respecto a los parámetros se llaman hipótesis paramétricas.
- El contraste se basa en establecer un criterio de decisión.
- Depende de la naturaleza de la población, de la distribución de probabilidad del estimador de dicho parámetro y del control que se desea fijar a priori sobre la probabilidad de rechazar la hipótesis contrastada en el caso de ser ésta cierta.

Planteamiento de una hipótesis estadística

- En general siempre se definen dos tipos de hipótesis:
- Hipótesis nula H_0,e Hipótesis alternativa H_1.
- El nombre de hipótesis nula se deriva del hecho de se plantea como una igualdad, lo cual facilita el tener una distribución de probabilidad de referencia específica.
- En general, la estrategia a seguir para probar una hipótesis es suponer que la hipótesis nula es verdadera, y que en caso de ser rechazada por la evidencia que aportan los datos, se estará H_1.
- La afirmación que se desea probar se aceptará como cierta, sólo en caso de rechazar la hipótesis nula.

¿De dónde nace H_0?

- Experiencia, pruebas pasadas o conocimiento del proceso. Interés: averiguar si ha cambiado el parámetro
- Alguna teoría o modelo sobre el funcionamiento del proceso. Interés: Verificar la validez de dicha teoría
- Especificaciones de diseño, obligaciones contractuales, normas a cumplir o solicitudes del cliente. Interés: probar el cumplimiento o incumplimiento de las especificaciones.

Hipótesis de contraste

- Una hipótesis estadística es un proposición sobre los parámetros de una población o sobre la distribución de probabilidad de una variable aleatoria
- Una prueba de hipótesis es una herramienta de análisis de datos que puede en general formar parte de un experimento comparativo más completo.
- En todo contraste intervienen dos hipótesis:
  - La hipótesis nula (H_0) es aquella que recoge el supuesto de que el parámetro toma un valor determinado y es la que soporta la carga de la prueba.
  - La proposición contraria a la hipótesis nula recibe el nombre de hipótesis alternativa (H_1) y suele presentar un cierto grado de indefinición.
  - Si la hipótesis alternativa se formula simplemente como 'la hipótesis nula no es cierta', el contraste es bilateral o a dos colas;
  - Por el contrario cuando se indica el sentido de la diferencia, el contraste es unilateral o a una sola cola.

Rechazar o no una hipótesis

- La decisión de rechazar la hipótesis nula, que en principio se considera cierta, está en función de que sea o no compatible con la evidencia empírica contenida en la muestra.
- El contraste clásico permite controlar a priori la probabilidad de cometer el error de rechazar la hipótesis nula siendo ésta cierta.
- Dicha probabilidad se llama nivel de significación del contraste (α) y se acostumbra a fijarse en el 1%, 5% o 10%.

Rechazar o no una hipótesis

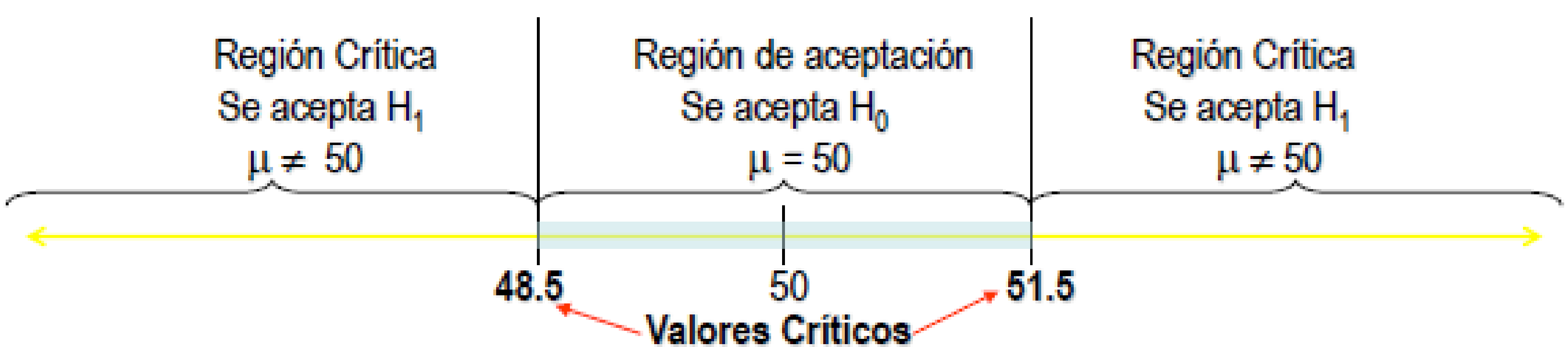
- Cuando se realiza un y no se fija el nivel de significación deseado, el programa calcula el valor-p o significación asintótica.
- Esta es la probabilidad de que el estadístico de prueba tome un valor igual o superior al muestral bajo el supuesto de que la hipótesis nula es cierta.
- Por tanto, si el valor-p es menor o igual que el nivel de significación deseado se rechazará Ho.
- Un valor-p próximo a cero indica que se rechazará la Ho para cualquier nivel de significación.

Ejemplo:

- Se tiene interés en la rapidez de combustión de un agente propulsor para los sistemas de salida de emergencia en aeronaves. (esta rapidez es una variable aleatoria con alguna distribución de probabilidad).
- Especialmente interesa la rapidez de combustión promedio (que es un parámetro (μ) de dicha distribución). De manera más específica, interesa decidir si esta rapidez promedio es o no 50 cm/seg.
- El planteamiento formal de la situación se realiza en términos de una Hipótesis Nula (que es la proposición que se quiere poner a prueba) y una Hipótesis Alternativa, la cual se aceptará si se rechaza la hipótesis nula:
  - Hipótesis Nula: H_0: μ= 50 cm/seg
  - Hipótesis Alternativa: H_1: $\mu \neq$50 cm/seg
- En el ejemplo se tiene una Hipótesis Alternativa Bilateral, ya que se verifica para valores de μ a ambos lados de 50 cm/seg.
- ¿Cómo se puede convertir la hipótesis en unilateral?

¿Se acepta o no H_0?

- Aceptación de H_0.- Un valor de la media muestral $\bar{x}$ "muy cercano" a 50 cm/seg es una evidencia que apoya a la hipótesis nula
- Sin embargo es necesario introducir un criterio para decidir que tanto es muy cercano, para el ejemplo este criterio pudiera ser: $48.5 \leq \bar{x} \leq 51.5$, si esto ocurre se acepta H_0
- De lo contrario, es decir, si $\bar{x} < 48.5$ o $\bar{x} > 51.5$, se acepta H_1



El valor de P

- Las pruebas de hipótesis se denominan de significancia pues los valores de p por sobre o debajo del límite se denominan como significativos, altamente significativos o no significativos.
- La p no es necesariamente efectiva como guía de la decisión científica, pues el valor de p no tiene que ver con la magnitud de la diferencia que se estudia
- La p no es confiable porque depende absolutamente del tamaño de la muestra y esto la hace especialmente peligrosa en trabajos epidemiológicos.
- Es necesario tener claro que se rechaza la hipótesis nula porque es poco probable que sea verdadera con los datos obtenidos.
- En ningún caso se está probando la hipótesis alternativa de que sí hay efecto. Sólo se dice que la probabilidad de que el efecto observado no exista realmente es tan baja que se acepta que lo más probable es que efectivamente exista.
- A la inversa, el rechazo de la hipótesis nula jamás implica la comprobación de la igualdad.
- La distinción podría parecer sutil semánticamente pero es muy importante en la interpretación de los resultados.

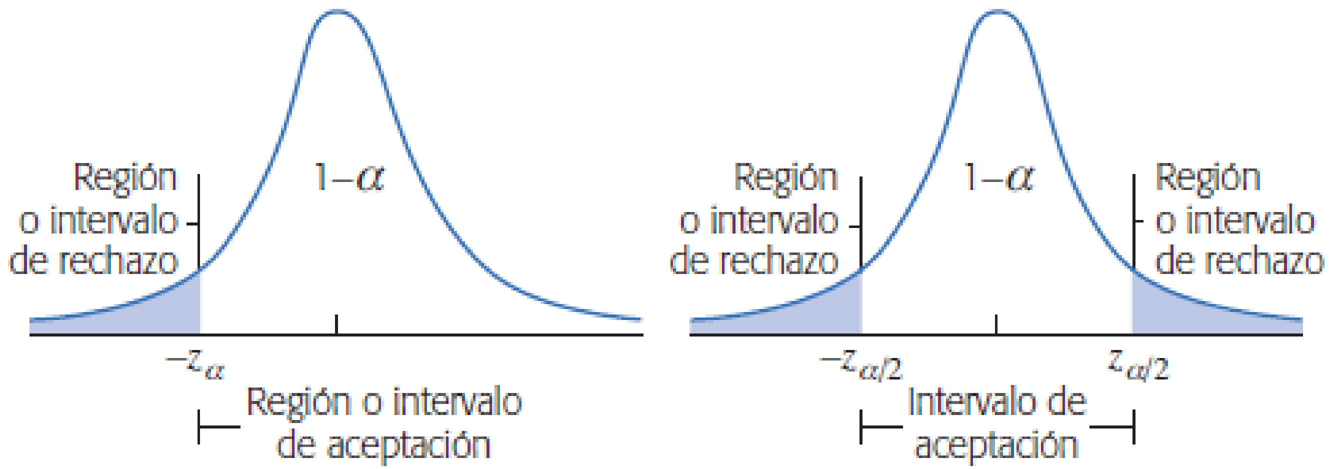
Proceso de contraste:

- Paso 1:
  - Expresar el interrogante de la investigación como hipótesis estadística:
    - H_0: "no hay diferencia"
    - H_1: "Hay diferencia"
- Paso 2:
  - Decidir sobre la prueba estadística adecuada según la población y el tipo de variable.
  - Elegir la muestra
- Paso 3:
  - Seleccionar el grado de significancia α (probabilidad de rechazar H_0 siendo esta correcta o error Tipo I).
- Paso 4:
  - Realizar los cálculos y exponer conclusiones.

Criterios equivalente de rechazo o aceptación

- Existen tres criterios equivalentes para decidir cuándo rechazar la hipótesis nula y, en consecuencia, aceptar la hipótesis alternativa.
- **Estadístico de prueba frente a valor crítico**: rechazar H0 si el estadístico de prueba cae en la región de rechazo que está delimitada por el valor crítico.
- **Significancia observada frente a significancia predefinida**: la *significancia predefinida* se denota por α es el riesgo máximo que se está dispuesto a correr por rechazar H_0 indebidamente (error tipo I). La *significancia observada* o *calculada*, también conocida como "valor-*p*", es el área bajo la distribución de referencia más allá del valor del estadístico de prueba.
- **Intervalo de confianza**: se rechaza H_0 si el valor del parámetro declarado en la hipótesis nula se encuentra fuera del intervalo de confianza para el mismo parámetro.

- Si p es menor que α, rechazar H_0 y aceptar la alternativa; en caso contrario, se acepta la hipótesis nula.
- El conjunto de valores que resultarían en el rechazo de H_0 – calculados conociendo la prueba usada, α y el número de observaciones – se conoce con el nombre de región crítica.
- En otras palabras, se rechaza la H_0 si el estadístico cae en la región crítica.
- Elegir la prueba estadística apropiada de acuerdo al diseño experimental, el tipo de datos y el número de grupos que se comparan.
- La cifra que resulta de usar la prueba en los datos recolectados se conoce como el estadístico del test en cuestión: z; estadístico t o de Student, la r de Pearson, F del análisis de varianza, entre otros.

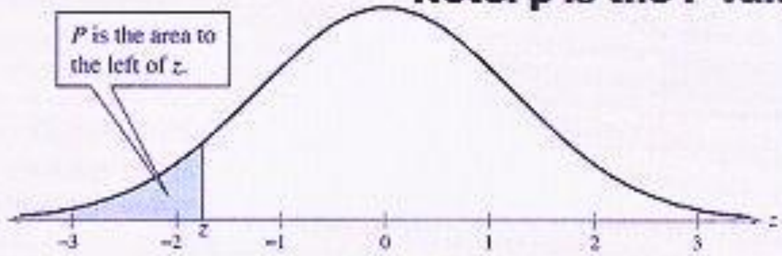


una cola   dos colas

Tomado de Gutiérrez Pulido y De La Vara Salazar (2008)

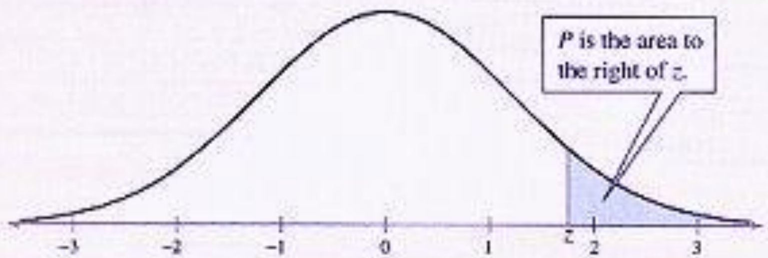
1. If the alternative hypothesis H_a contains the less-than inequality symbol ($<$), the hypothesis test is a **left-tailed test.**

Note: p is the P-value



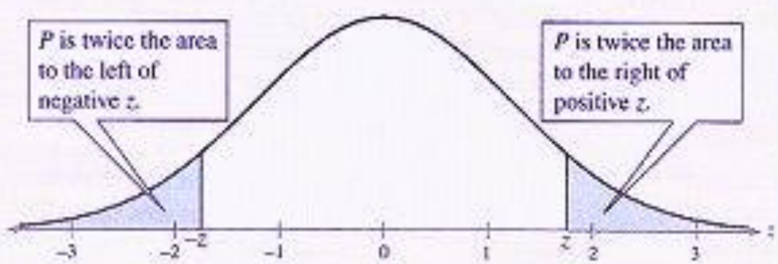
Left-Tailed Test

2. If the alternative hypothesis H_a contains the greater-than inequality symbol ($>$), the hypothesis test is a **right-tailed test.**



Right-Tailed Test

3. If the alternative hypothesis H_a contains the not-equal-to symbol ($\neq$), the hypothesis test is a **two-tailed test.** In a two-tailed test, each tail has an area of $\frac{1}{2}P$.



Two-Tailed Test

Estimación de parámetros

- Consiste en asignar un valor concreto al parámetro o parámetros que caracterizan la distribución de probabilidad de la población.
- Un error de estimación es la diferencia entre la estimación y el verdadero valor del parámetro.
- Para valorar el grado de precisión asociado con una estimación puntual se parte de dicha estimación para construir un intervalo de confianza.
- La amplitud de éste constituye una medida del grado de precisión con el que se estima el parámetro.

Estimación puntual y por intervalo

- Para definir completamente una distribución de probabilidad hay que definir sus parámetros.
- Estimación puntual:
  - Un estimador puntual es un estadístico que estima el valor específico de un parámetro.
  - Los parámetros más comunes son:
    - La media μ del proceso (población).
    - La varianza σ^2 o la desviación estándar σ del proceso.
    - La proporción
    - *p* de artículos defectuosos.
  - Los estimadores típicos son:
    - La media muestral $\hat{\mu} = \bar{X}$
    - La varianza muestral $\hat{\sigma}^2 = S^2$
    - La proporción de defectuosos en la muestra, $\hat{p} = \frac{x}{n}$, donde x es la cantidad de defectuosos y n el número total de la muestra.

- Estimación por intervalo:
  - Toma en cuenta la variación o error aleatorio.
  - Se hace a través de un intervalo de confianza que indica el rango donde puede estar un parámetro dado.
  - Un intervalo al (1-α)% para un parámetro desconocido θ, estará definido por dos estadísticos L y U de tal manera de que la probabilidad de que θ esté dentro del intervalo L – U está dada por:

$$P(L \leq \theta \leq U) = 1 - \alpha$$

  - La longitud del intervalo de confianza es una medida de la precisión de la estimación, por lo que se desea que este sea pequeño y con un alto nivel de confianza.

Intervalo de confianza para una media:

- Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de tamaño n, de una población normalmente distribuida con media μ y varianza σ^2, el intervalo de confianza con un nivel $1-\alpha$, estará dado por:

$$P\left(-t_{\alpha/2} \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq t_{\alpha/2}\right) = 1-\alpha$$

ó

$$P\left(\bar{X} - t_{\alpha/2}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha$$

Donde t sigue una distribución T de Student, con n-1 grados de libertad, tal que:

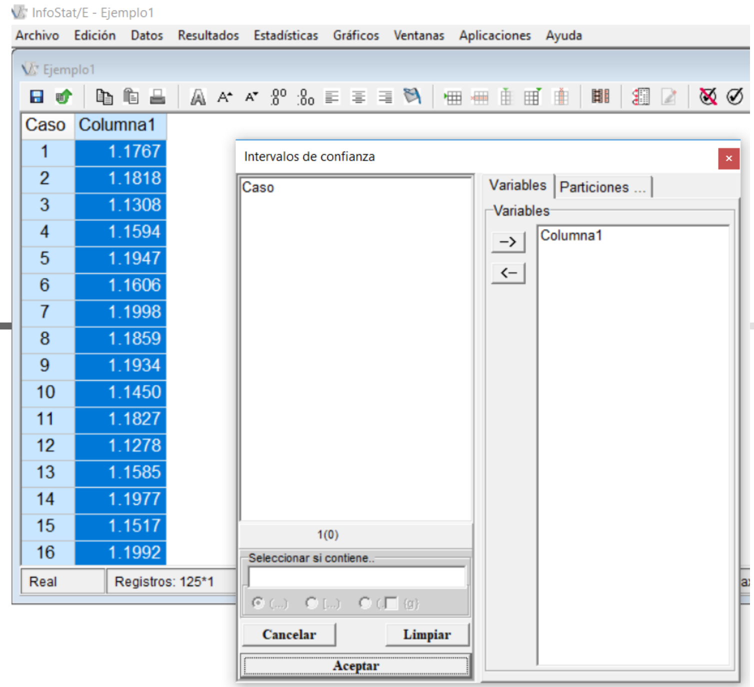
$$t = \frac{\bar{X}-\mu}{S/\sqrt{n}}$$

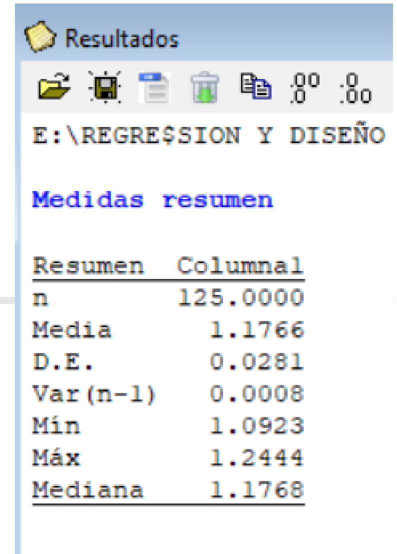
El error estándar está dado por $\dfrac{S}{\sqrt{n}}$

Ejemplo:

- Se tiene un proceso donde el producto debe tener un espesor de 1.20mm con una tolerancia de ±0.1. Así el producto deberá tener un rango de grosor aceptado de entre 1.10 y 1.30 mm.
- Para evaluar si se estaba dentro del rango de diseño, se hace un muestreo de 125 partes y se obtiene la información mostrada a continuación.
- Tamaño de la muestra n 125
- Media muestral 1.1773 mm
- Desviación estándar muestral 0.0244

Utilizando Infostat




Resultados

E:\REGRE$SION Y DISEÑO

Medidas resumen

Resumen	Columna1
n	125.0000
Media	1.1766
D.E.	0.0281
Var(n-1)	0.0008
Mín	1.0923
Máx	1.2444
Mediana	1.1768

Intervalos de confianza

Bilateral

Estimación paramétrica

Variable	Parámetro	Estimación	E.E.	n	LI(95%))	LS(95%))
Columna1	Media	1.176558	0.002513	125	1.171583	1.181533

Intervalos de confianza

Bilateral
Estimación paramétrica

Variable	Parámetro	Estimación	E.E.	n	LI(90%))	LS(90%))
Espesor en mm	Media	1.17753	0.00202	125	1.17417	1.18088

Intervalos de confianza

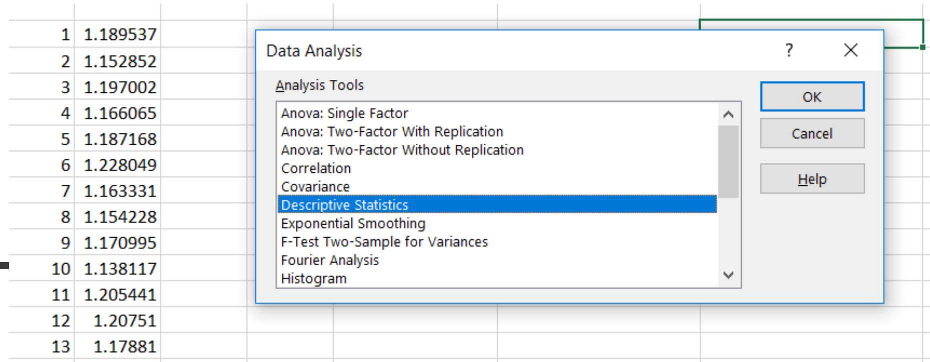
Bilateral
Estimación paramétrica

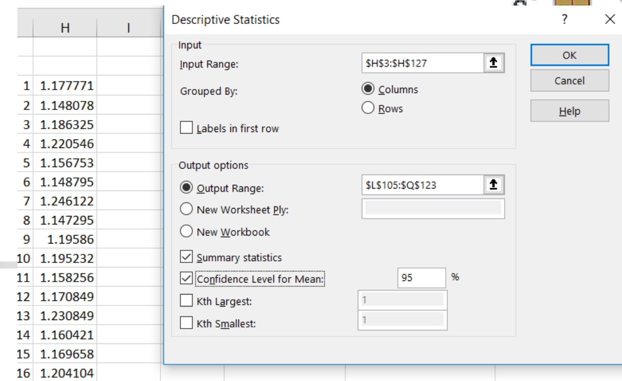
Variable	Parámetro	Estimación	E.E.	n	LI(95%))	LS(95%))
Espesor en mm	Media	1.1775	0.0020	125	1.1735	1.1815

Intervalos de confianza

Bilateral
Estimación paramétrica

Variable	Parámetro	Estimación	E.E.	n	LI(99%))	LS(99%))
Espesor en mm	Media	1.1775	0.0020	125	1.1722	1.1828





Excel

	Column1
Mean	1.176558005
Standard Error	0.002513371
Median	1.176755027
Mode	#N/A
Standard Deviatior	0.028100341
Sample Variance	0.000789629
Kurtosis	0.295489558
Skewness	-0.0565701
Range	0.152104401
Minimum	1.092282932
Maximum	1.244387333
Sum	147.0697506
Count	125
Confidence Level(9	0.004974665

Intervalo para la varianza

- Es posible definir intervalos de confianza para la varianza σ^2.
- La distribución de referencia es una ji-cuadrada (chi-cuadrada) con $n - 1$ grados de libertad o $\chi^2_{\alpha/2,\, n-1}$, tal que:

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2,\, n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,\, n-1}}$$

Prueba t para la media

- Cuando se estudia el comportamiento de un proceso o un fenómeno suelen interesar su media y varianza (o desviación estándar).
- En particular, al estudiar la media μ, es de interés preguntarse si ésta es igual, mayor o menor a cierto valor μ_0, donde μ_0 es un número conocido.
- En general, para estos casos, tanto el valor de la media μ y el valor de la desviación estándar σ son desconocidas.
- Sea X una variable aleatoria de media μ y desviación estándar σ, ambas desconocidas, se quiere probar la hipótesis:

$$H_0 : \mu = \mu_0$$

$$H_A : \mu \neq \mu_0$$

- Para probar esta hipótesis se toma una muestra aleatoria de tamaño n de los posibles valores de la variable X y se calcula el estadístico de prueba:

$$t_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

donde S es la desviación estándar de la muestra y $\bar{X}$ la media muestral.

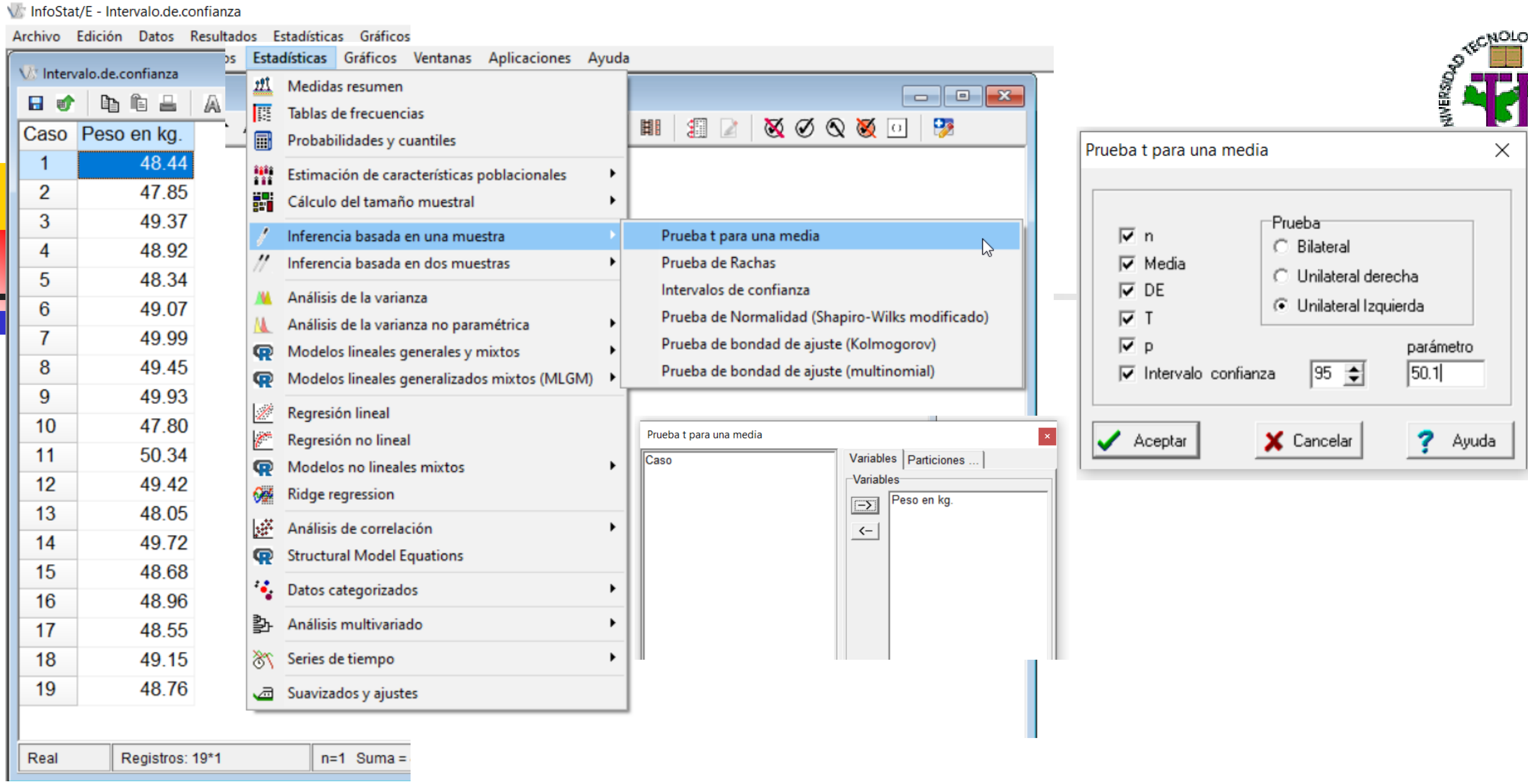
- Bajo el supuesto de que H_0 es verdadera, este estadístico se distribuye T de Student con $n - 1$ grados de libertad.
- Se rechaza H_0 si el valor absoluto del estadístico de prueba es mayor que el valor crítico de la distribución, es decir, se rechaza H_0 si $|t_0| > \alpha/2$.
- Cuando $\mu \leq \mu_0$ o $\mu \geq \mu$, la hipótesis alternativa es unilateral y se rechazará H_0 cuando $t_0 \leq -t_\alpha$, o cuando $t_0 \geq t_\alpha$, respectivamente.

Ejemplo: peso de sacos de arroz:

- Una empresa que compra arroz tiene la duda de que el peso de los sacos es el que realmente le están vendiendo. De acuerdo con el molino, los sacos tienen una media de 50.1 kg, con una desviación estándar de 0.75 kg.
- A fin de comprobar esto, selecciona una muestra de 19 sacos y los pesa, obteniendo la siguiente información que se muestra.
- Plantea las siguientes hipótesis a contrastar:

- H_0: $\mu=50.1$
- H_1: $\mu<50.1$

48.44
47.85
49.37
48.92
48.34
49.07
49.99
49.45
49.93
47.80
50.34
49.42
48.05
49.72
48.68
48.96
48.55
49.15
48.76



No se acepta H_0

Prueba t para una media

Valor de la media bajo la hipótesis nula: 50.1

Variable	n	Media	DE	LS(95)	T	p(Unilateral I)
Peso en kg.	19	48.99	0.73	49.28	-6.67	<0.0001

Hipótesis para dos medias independientes: comparación para dos tratamientos.

- Se presenta al comparar la media de dos procesos o dos tratamientos independientes.
- Esta suposición se justifica por la manera en que se obtienen los datos; es decir, a la muestra a la que se le aplica el tratamiento 1 es independiente de la muestra para el tratamiento 2, y los datos se obtienen en orden completamente *al azar.*
- Suponga que interesa comparar dos tratamientos A y B, que realizan la misma acción. Para ello se obtendrá una muestra aleatoria de observaciones de cada Tratamiento.
- Suponga que los datos a observar en A son Y_{Al}, Y_{A2}, ..., Y_{An} y los datos de B son Y_{Bl}, Y_{B2}, ..., Y_{Bn}.

	Tratamientos	
Prueba	A	B
1	Y_{al}	Y_{Bl}
2	Y_{A2}	Y_{B2}
...	...	...
n	Y_{An}	Y_{Bn}

- Sean dos tratamientos con medias y varianzas μ_x; σ^2_x y μ_y; σ^2_y, respectivamente.
- Es posible plantear las siguientes hipótesis:

$$H_0 : \mu_x = \mu_y$$

$$H_A : \mu_x \neq \mu_y$$

- Reescribiendo las mismas:

$$H_0 : \mu_x - \mu_y = 0$$

$$H_A : \mu_x - \mu_y \neq 0$$

- Si cada proceso sigue una distribución normal y son independientes entre ellos, el estadístico de prueba adecuado para probar la hipótesis de igualdad de medias está dado por,

$$t_0 = \frac{\bar{X} - \bar{Y}}{S_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}}$$

- Se rechaza H_0 si $|t_0| > t_{\alpha/2}$

Se tienen dos tratamientos y se quiere analizar si ambos son iguales en términos de peso, en mg, obtenido. Se tiene la siguiente información

X	Y
4.71	4.31
5.14	4.74
4.61	4.27
3.96	3.75
4.69	4.33
4.02	3.81
4.89	4.53
4.56	4.41
4.91	4.62
4.63	4.34
4.87	4.78
4.92	4.71
4.99	4.75

¿Existen diferencias entre los tratamientos?

Caso	Tratamiento	Peso
1	X	4.71
2	X	5.14
3	X	4.61
4	X	3.96
5	X	4.69
6	X	4.02
7	X	4.89
8	X	4.56
9	X	4.91
10	X	4.63
11	X	4.87
12	X	4.92
13	X	4.99
14	Y	4.31
15	Y	4.74
16	Y	4.27
17	Y	3.75
18	Y	4.33
19	Y	3.81
20	Y	4.53
21	Y	4.41
22	Y	4.62
23	Y	4.34
24	Y	4.78
25	Y	4.71
26	Y	4.75

Prueba T para muestras Independientes

Clasific	Variable	Grupo 1	Grupo 2	n(1)	n(2)	Media(1)	Media(2)	Media(1)-Media(2)	LI(95)	LS(95)	pHomVar	T	p-valor	prueba
Tratamiento	Peso	{X}	{Y}	13	13	4.68	4.41	0.27	-4.5E-03	0.55	0.8771	2.03	0.0535	Bilateral

Comparando muestras pareadas

- *Orden completamente al azar* es aquel en el que las unidades se asignan de manera aleatoria a los tratamientos y las pruebas experimentales se hacen en orden aleatorio
- Muestras pareadas son aquellas en las que los datos de ambos tratamientos se obtienen por pares, de manera que éstos tienen algo en común y no son independientes.
- Ejemplos:
  - A los mismos pacientes se les aplican dos medicamentos (tratamientos) para el dolor en distintas ocasiones; los tratamientos a comparar son los dos medicamentos.
  - A las mismas piezas se les hace una prueba de dureza con distintos instrumentos; aquí se quieren comparar los instrumentos.

La hipótesis planteada

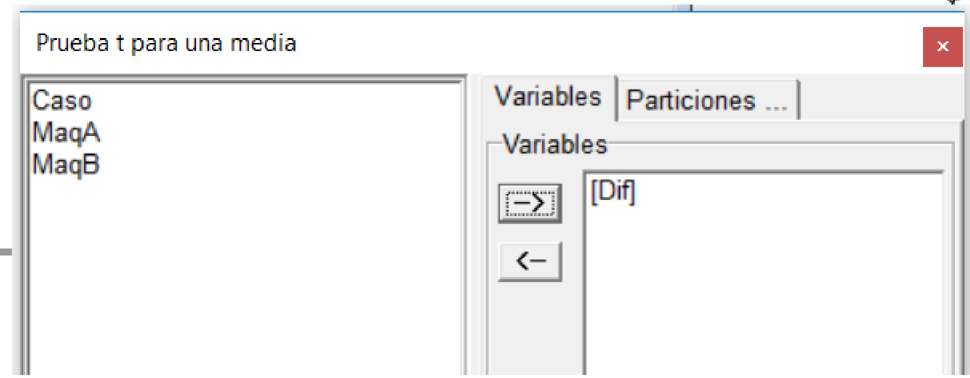
- La comparación entre ambos tratamientos puede hacerse bajo la siguiente hipótesis:
  - H_0: $\mu_1=\mu_2$
  - H_1: $\mu_1\neq\mu_2$
- Debido a su dependencia es posible definir $\mu_D= \mu_1- \mu_2$, y
  - H_0: $\mu_D = 0$
  - H_1: $\mu_D\neq 0$
- De esta manera, el problema de
- comparar las medias de dos poblaciones se convierte en el problema de comparar la media de una población con una constante cuando $\mu_0=0$.

Ejemplo:

- Se quieren probar dos básculas sobre la misma muestra para ver si los mismas son similares. Se tienen los siguientes datos:

Espécimen	Báscula 1	Báscula 2	Diferencia
1	11.23	11.27	–0.04
2	14.36	14.41	–0.05
3	8.33	8.35	–0.02
4	10.50	10.52	–0.02
5	23.42	23.41	–0.01
6	9.15	9.17	–0.02
7	13.47	13.52	–0.05
8	6.47	6.46	0.01
9	12.40	12.45	–0.05
10	19.38	19.35	0.03
Medias:	12.87	12.89	–0.022

Caso	MaqA	MaqB	[Dif]
1	11.23	11.27	-0.04
2	14.36	14.41	-0.05
3	8.33	8.35	-0.02
4	10.50	10.52	-0.02
5	23.42	23.41	0.01
6	9.15	9.17	-0.02
7	13.47	13.52	-0.05
8	6.47	6.46	0.01
9	12.40	12.45	-0.05
10	19.38	19.35	0.03



Prueba t para una media

Valor de la media bajo la hipótesis nula: 0

Variable	n	Media	DE	LI(95)	LS(95)	T	p(Bilateral)
[Dif]	10	-0.02	0.03	-0.04	5.1E-04	-2.21	0.0548

Valor de la media bajo la hipótesis nula: 0

Variable	n	Media	DE	LI(95)	T	p(Unilateral D)
[Dif]	10	-0.02	0.03	-0.04	-2.21	0.9726

Valor de la media bajo la hipótesis nula: 0

Variable	n	Media	DE	LS(95)	T	p(Unilateral I)
[Dif]	10	-0.02	0.03	-3.4E-03	-2.21	0.0274

Prueba de hipótesis para un parámetro

Hipótesis	Estadístico de prueba	Criterio de rechazo
a) $H_0 : \mu = \mu_0$ $H_A : \mu \neq \mu_0$ $H_A : \mu > \mu_0$ $H_A : \mu < \mu_0$	$t_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$\lvert t_0 \rvert > t_{\alpha/2}$ $t_0 > t_\alpha$ $t_0 < -t_\alpha$
b) $H_0 : \sigma^2 = \sigma_0^2$ $H_A : \sigma^2 \neq \sigma_0^2$ $H_A : \sigma^2 > \sigma_0^2$ $H_A : \sigma < \sigma_0^2$	$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$	$\chi_0^2 > \chi^2_{\alpha/2,\, n-1}$ o $\chi_0^2 < \chi^2_{1-\alpha/2,\, n-1}$ $\chi_0^2 > \chi^2_{\alpha,\, n-1}$ $\chi_0^2 < \chi^2_{1-\alpha,\, n-1}$
c) $H_0 : p = p_0$ $H_A : p \neq p_0$ $H_A : p > p_0$ $H_A : p < p_0$	$z_0 = \frac{X - np_0}{\sqrt{np_0(1-p_0)}}$ X número de defectos	$\lvert z_0 \rvert > z_{\alpha/2}$ $z_0 > z_\alpha$ $z_0 < -z_\alpha$

Tomado de Gutiérrez Pulido y De La Vara Salazar (2008)

Prueba de hipótesis para dos parámetros

Tomado de Gutiérrez Pulido y De La Vara Salazar (2008)

Hipótesis	Estadístico de prueba	Criterio de rechazo
a) $H_0 : \mu_1 = \mu_2$ $H_A : \mu_1 \neq \mu_2$ $H_A : \mu_1 > \mu_2$ $H_A : \mu_1 < \mu_2$	$t_0 = \dfrac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ donde $S_p = \sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$	$\lvert t_0 \rvert > t_{\alpha/2,\, n_1+n_2-2}$ $t_0 > t_{\alpha,\, n_1+n_2-2}$ $t_0 < -t_{\alpha,\, n_1+n_2-2}$
b) $H_0 : \mu_1 = \mu_2$ $H_A : \mu_1 \neq \mu_2$ $H_A : \mu_1 > \mu_2$ $H_A : \mu_1 < \mu_2$	$t_0 = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}}}$ donde $v = \dfrac{\left(\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}\right)^2}{\dfrac{(S_1^2/n_1)^2}{n_1+1}+\dfrac{(S_2^2/n_2)^2}{n_2+1}} - 2$	$\lvert t_0 \rvert > t_{\alpha/2,\, v}$ $t_0 > t_{\alpha,\, v}$ $t_0 < -t_{\alpha,\, v}$
c) $H_0 : \sigma_1^2 = \sigma_2^2$ $H_A : \sigma_1^2 \neq \sigma_2^2$ $H_A : \sigma_1^2 > \sigma_2^2$ $H_A : \sigma_1^2 < \sigma_2^2$	$F_0 = \dfrac{S_1^2}{S_2^2}$	$F_0 > F_{\alpha/2,\, n_1-1,\, n_2-1}$ o $F_0 < F_{1-\alpha/2,\, n_1-1,\, n_2-1}$ $F_0 > F_{\alpha,\, n_1-1,\, n_2-1}$ $F_0 < F_{1-\alpha,\, n_1-1,\, n_2-1}$
d) $H_0 : p_1 = p_2$ $H_A : p_1 \neq p_2$ $H_A : p_1 > p_2$ $H_A : p_1 < p_2$	$z_0 = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$ donde $\hat{p} = \dfrac{X_1 + X_2}{n_1 + n_2}$	$\lvert z_0 \rvert > z_{\alpha/2}$ $z_0 > z_\alpha$ $z_0 < -z_\alpha$
e) $H_0 : \mu_1 = \mu_2$ $H_A : \mu_1 \neq \mu_2$ $H_A : \mu_1 > \mu_2$ $H_A : \mu_1 < \mu_2$	$t_0 = \dfrac{\bar{d}}{S_d / \sqrt{n}}$	$\lvert t_0 \rvert > t_{\alpha/2,\, n-1}$ $t_0 > t_{\alpha,\, n-1}$ $t_0 < -t_{\alpha,\, n-1}$