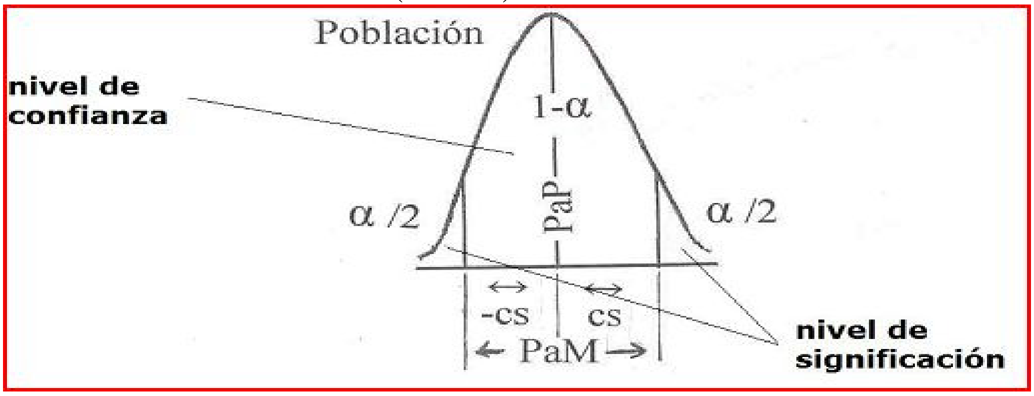
Tema 13 : Intervalos de probabilidad y confianza. Hipótesis y decisiones estadísticas.

---Intervalo de probabilidad (IP)

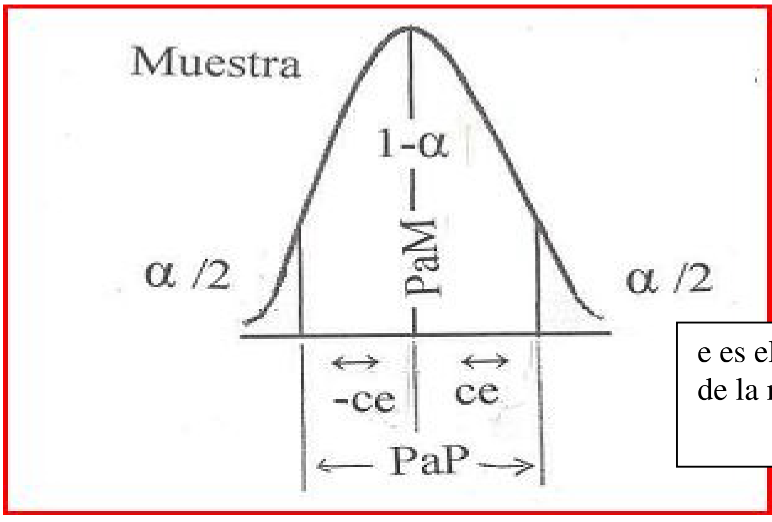
Permite predecir el comportamiento de las muestras.
Si de una población se sacan infinitas muestras y se calcula en ellas un parámetro (media, %, etc.), los resultados varían siguiendo una DN y la media de todos ellos coincide con el parámetro de la población (PaP o PP).
La probabilidad de que el parámetro de una muestra (PaM o PM) esté dentro de un determinado intervalo de valores es 1-α y la probabalidad de estar fuera de ese intervalo es α. A 1-α se le llama **nivel de confianza** y a α **nivel de significación**. La suma de ambos niveles vale 1 (ó 100%).



α la fijamos nosotros y habitualmente se manejan tres puntos de referencia: 0,05 (ó 5%), 0,01(ó 1%) y 0,001 (ó 1‰) Por tanto los correspondientes puntos de referencia del nivel de confianza son: 0,95 (95%) ; 0,99 (99%) ; 0,999 (99,9%).
A esos tres valores de α le corresponden en la DN los siguientes valores de c: 1,96 ; 2,58 y 3,30 , respectivamente

--Intervalo de confianza (IC)

Se obtiene a partir de una muestra en la que calculamos un parámetro y , aplicando la fórmula correspondiente, también un intervalo, en el que estará el verdadero valor del parámetro en la población al nivel de confianza que se elija.



e es el error estándar de la muestra

---Las **PRUEBAS DE HIPOTESIS**, típicas de la Estadística Inferencial se dividen en cuatro grandes clases:

1. **Pruebas de estimación**. A partir del parámetro de la muestra hacemos una estimación de ese parámetro en la población calculando el intervalo de confianza.
2. **Pruebas de conformidad**, que permiten verificar si el parámetro calculado en una muestra puede proceder de una población determinada. Puede proceder si ese parámetro está dentro del intervalo de probabilidad de la población. Estas pruebas contestan a las preguntas: ¿Puede proceder...de...?, ¿Es conforme...con...?

Pruebas de contraste de variables:

3**.--Pruebas de relación o dependencia**. Permiten verificar si dos o más variables están relacionadas o son independientes. Contestan a las preguntas: ¿Hay una relación entre las variables? , ¿los valores de Y dependen de los de X?,...

4.--**Pruebas de comparación**, que permiten saber si las diferencias observadas entre dos o más muestras se deben al azar, en cuyo caso no existen diferencias de importancia estadística; son muestras de la misma población y están dentro de su intervalo de probabilidad. Contestan a la pregunta: Los datos de las muestras que comparamos son más o menos iguales o difieren significativamente?

> Cuando los datos son independientes, relación y comparación son lo mismo, simples variantes de enfoque del mismo problema, y se resuelven utilizando las mismas fórmulas. En cambio, si los datos son apareados, las dos pruebas son esencialmente distintas y se resuelven con fórmulas distintas. ¡Hay que hacerse las preguntas correspondientes para elegir el camino adecuado!.

---Metódica de las pruebas de hipótesis

1. Se formula la hipótesis estadística
2. Se aplica la prueba o test estadístico que corresponda
3. En función de los resultados se toma una decisión estadística.

*** * *** La **HIPOTESIS ESTADISTICA** inicial es la **Hipótesis nula (H_0) de igualdad o no relación** entre las variables contrastadas. Dice que las diferencias de los parámetros de las variables no son diferencias importantes, que son debidas a las fluctuaciones del azar. O que no hay relación entre ellos. Todos proceden de la misma población, están dentro de su intervalo de probabilidad, también llamado zona de no rechazo de H_0. Ya sabemos que un valor cualquiera tiene una probabilidad $1-\alpha$ (el nivel de confianza) de estar en esa zona.
Si el resultado de la prueba, y sólo entonces, conduce al rechazo de H_0 , aparece y se acepta la **Hipótesis alternativa (H_1) de no igualdad o relación** entre las variables contrastadas. Las diferencias observadas no se explican por el azar, las muestras proceden de poblaciones distintas, ya que quedan fuera del IP, en la llamada zona de rechazo de H_0 , cuya p es el nivel de significación α.

> No hay que confundir la hipótesis del trabajo con la hipótesis estadística. Supongamos que hacemos un estudio esperando que un nuevo método terapéutico sea superior al clásico. Esta será la hipótesis del estudio. La hipótesis estadística será H_0 , o sea, que no hay diferencias de importancia estadística entre ambos métodos. Si la prueba estadística conduce al rechazo de H_0 , entonces se acepta H_1, que dirá que sí que hay diferencias significativas.

H_1 es habitualmente doble (pruebas bilaterales): las diferencias pueden estar a un lado u otro ; la relación puede ser positiva o negativa**. ¡Siempre que se acepte H_1 hay que indicar el sentido!**. En ocasiones, poco frecuentes en la práctica, puede interesar sólo uno de los sentidos (pruebas unilaterales).

*** * *** Las **pruebas estadísticas** se irán viendo en temas sucesivos.

** * La **decisión estadística** se toma en general siguiendo estos pasos:

1) **se aplica la prueba estadística** correspondiente, obteniendo un resultado, que para unificar el lenguaje llamaremos **z** , nombre arbitrario (podría llamarse de cualquier otra forma) que evita las confusiones que origina el hábito muy extendido de llamar a los resultados de las pruebas con el nombre de la distribución de referencia con que se valora los resultados (t de Student, χ^2, etc.). La prueba estadística se elige en función de la variable (CL o CT), de la naturaleza de los datos (independientes o apareados), del tamaño de la muestra y del cumplimiento de determinadas condiciones de aplicación.
2) **se busca el valor de referencia** (c de la DN, t , , χ^2 , F...) correspondiente al nivel de significación propuesto o en su defecto a 0,05.
3) **se compara z(en valor absoluto, con el valor de referencia (Ref.)** :
    a. **si $|z| <$ Ref. : no se puede rechazar H_0**. No se han encontrado diferencias estadísticamente significativas entre los grupos contrastados o no hay relación entre ellos, son independientes. Realmente es más correcto decir que no se puede rechazar H_0, que decir, cosa que se hace con frecuencia, que se acepta H_0 o que H_0 es verdadera. Nos quedamos con ella porque no podemos rechazarla. Es como una absolución por falta de pruebas. Se indica por n.s. (no significativo) ó $p>0,05$.
    b. si $|z| \geq$ Ref. : se rechaza H_0 y se acepta H_1 a ese nivel de significación. Hay diferencias o una relación con significación estadística. El sentido de las diferencias o de la relación, que siempre se debe dar, se deduce de los datos y parámetros. Se simboliza por $p< \alpha$ (el que corresponda).

En las pruebas de estimación y de conformidad, si no se dice otra cosa, sólo se toma el nivel de significación de 0,05. En las pruebas de contraste, si se supera un nivel hay que probar con el siguiente. El último superado es el definitivo. Es como en el salto de altura.

---Tres puntualizaciones:

--una significación estadística sólo permite establecer una relación de causalidad si se trata de un estudio experimental

--Una diferencia estadísticamente importante no quiere decir de forma automática que lo encontrado tenga importancia práctica. Eso lo dirán las circunstancias.

--si hay significación estadística, hay que buscar siempre la posible existencia de **factores de confusión**. Así un estudio puede sugerir que los alcohólicos tienen un riesgo alto de padecer cáncer de pulmón, pero resulta que casi todos los alcohólicos eran fumadores. En estos casos hay que estratificar en subgrupos del presunto ·confundidor”

---Errores de las decisiones estadísticas

-Como un α% de los PaM caen en la zona de rechazo, aunque H_0 sea verdadera, todo rechazo de H_0 conlleva un riesgo de error, el **ERROR TIPO I**, que es el que se comete cuando se acepta H_1 siendo H_0 verdadera. Podría decirse que es un FALSO POSITIVO. Su riesgo es α. Este riesgo lo fijamos nosotros y es por tanto conocido. por consenso el máximo riesgo que se admite es de 0,05 (ó 5%). Si no se dice otra cosa se acepta ese valor de α.. El error tipo I puede ser disminuido aumentando el tamaño muestral.

-El **ERROR TIPO II** es el que se comete al no rechazar H_0 , siendo H_1 verdadera. Equivale a un falso negativo. El riesgo de cometerlo se llama β (beta) y no lo conocemos exactamente, aunque hay métodos para estimarlo, que no veremos aquí. El problema es que si queremos disminuirlo, aumentamos α , y viceversa. Las fórmulas para el tamaño muestral tienen en cuenta esta circunstancia y , asumiendo una β entre 0,05 y 0,1 .En todo caso β disminuye también aumentando el tamaño de la muestra. A 1-β se la llama potencia de una prueba estadística.

Las decisiones estadísticas no "demuestran" nada. Sólo apoyan de una forma razonable una decisión o hecho concreto. Aceptar H_1 equivale a decir con un pequeño riesgo de error (α) que H_0 es falsa. No rechazar H_0 no quiere decir que sea verdadera, sólo que no ha podido ser rechazada (el riesgo β acecha...)

---Grado de significación

Se expresa por el mismo número que α, pero el concepto es ligeramente distinto. Es la probabilidad de que un resultado alcance un determinado valor cuando H_0 es verdadera. Cuantifica también la p de cometer un error tipo I. Su símbolo es p. Y se expresa como veíamos para α: $p < 0{,}05$ ó $p< 0{,}01$...

---Pruebas paramétricas y no paramétricas.

-**Las pruebas paramétricas** utilizan en sus cálculos parámetros, como media, varianza, frecuencia, porcentaje, etc.. Estas pruebas tienen unas condiciones de aplicación, que se especifican en cada prueba. Las mas frecuentes son: normalidad de la población origen, igualdad de varianzas, y tamaño adecuado. En la práctica, si la muestra es grande ($\geq$ 30) cumple siempre. Por tanto es en las muestras pequeñas donde hay que comprobar las condiciones de aplicación. Si no las cumplen, no pueden utilizarse esas pruebas y hay que recurrir a las pruebas no paramétricas, que no tienen condiciones de aplicación y se pueden utilizar siempre. Algunas pruebas son muy robustas (como el ANOVA) y la no observancia de las condiciones de aplicación no altera sustancialmente la decisión estadística, por lo que casi nunca se tienen en cuenta.

-**Las pruebas no paramétricas** se basan en la comparación de los datos aislados y en su ordenación según el criterio propio de cada test.. A igualdad de tamaño muestral son menos eficientes que las prueba paramétricas, por lo que siempre que sea posible se deben usar éstas.

Recordatorio

ZONA DE ACEPTACION Y ZONA DE RECHAZO

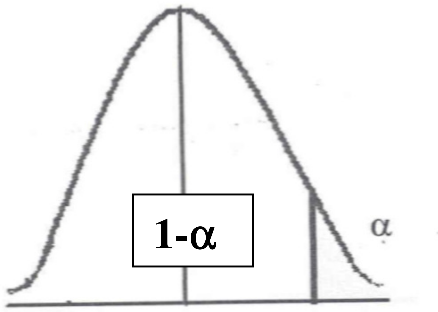
las pruebas estadísticas prueban la hipótesis nula H_0, que puede rechazarse o no rechazarse

la zona de no rechazo corresponde a **1-α**

la zona de rechazo correponde a **α** , y puede ser única, en un solo lado de la campana (pruebas unilaterales) o doble, en ambos lados de la campana (pruebas bilaterales).

Al rechazar H_0 se acepta H_1 y en las pruebas bilaterales (casi todas) hay que dar el sentido del rechazo

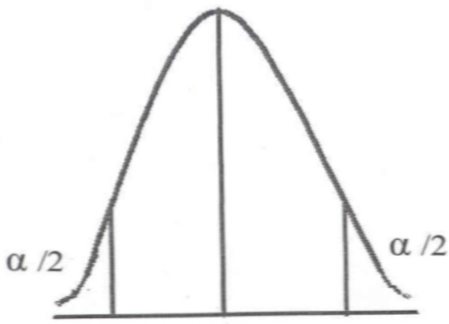
El no rechazo de H_0 no prueba que H_0 sea verdadera, sólo que no puede rechazarse (algo así como una absolución por falta de pruebas, que no afirma que el acusado sea inocente, sino que no hay pruebas para considerarlo culpable)



Prueba unilateral

H_1 es única

A > B



Prueba bilateral

H_1 es doble:

A > B

B < A

En principio las pruebas son bilaterales, y si no se dice otra cosa hay que entender que la prueba es bilateral. Interesa cualquier tipo de diferencias o relaciones.

Decisión estadística	**p acertar**	**p no acertar**	**¿conocido?**	**< riesgo si**
no rechazo H_o	**1 - β = potencia**	**β**	**no**	**> N < β**
rechazo de H_0 y aceptación H_1	**1 - α = nivel de confianza**	**α = nivel de significación**	**sí 0,05 ó menos**	**> N < α**