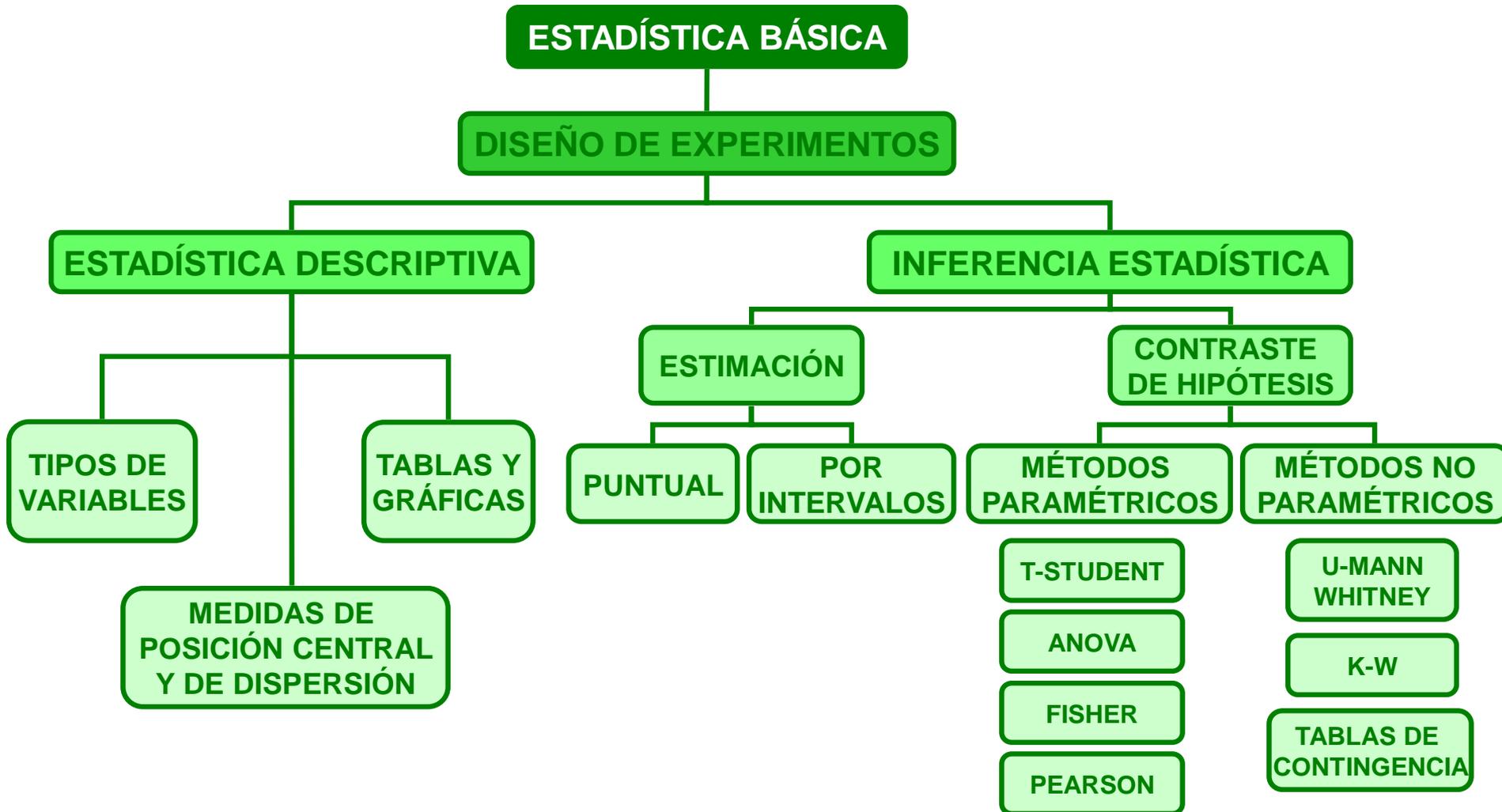


CURSO DE ESTADÍSTICA BÁSICA

ESQUEMA DEL CURSO



SESIÓN 1

DISEÑO DE ESTUDIOS EN INVESTIGACIÓN MÉDICA

1.1 Estudios observacionales

1.2 Estudios experimentales

SESIÓN 2

DESCRIPTIVA

2.1 Escalas de medida

2.2 Variables

2.3 Resumen de datos con números

2.4 Resumen de datos con gráficos

1

2

3

4

5

6

SESIÓN 3

ESTIMACIÓN

3.1 Concepto de estimación

3.2 Error estándar

3.3 Intervalo de confianza

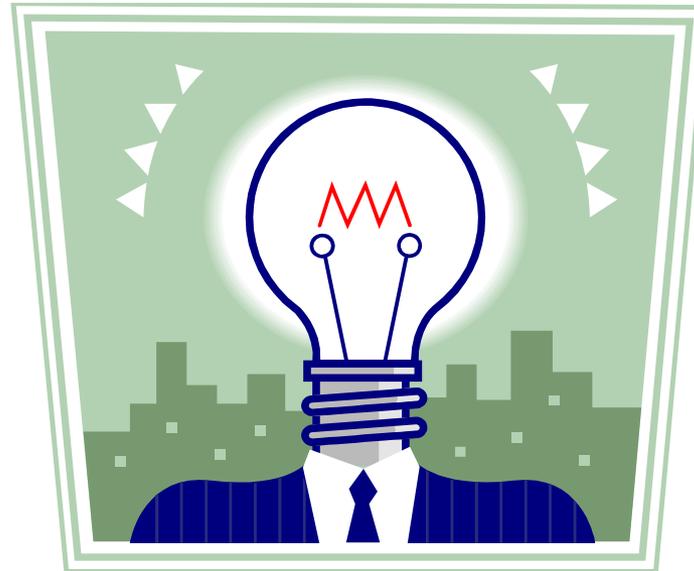
INFERENCIA ESTADÍSTICA I

SESIÓN 4

- 4.1 Contraste de hipótesis
- 4.2 Métodos Paramétricos
- 4.3 Transformaciones de datos
- 4.4 Métodos no Paramétricos

***“Si oigo algo lo olvido.
Si lo veo lo entiendo.
Si lo hago lo aprendo”.***

Confucio (551-478 A.C)



SESIÓN 3

ESTIMACIÓN

3.1 Concepto de estimación

3.2 Error estándar

3.3 Intervalo de confianza

3.4 Tamaño de muestra

3.1 CONCEPTO DE ESTIMACIÓN

3.2 ERROR ESTÁNDAR

3.3 INTERVALO DE CONFIANZA

3.4 TAMAÑO DE MUESTRA

DEFINICIÓN DE ESTIMACIÓN

Proceso de utilizar información de una muestra para extraer conclusiones acerca de toda la población

Se utiliza la información para estimar un valor

3.1 CONCEPTO DE ESTIMACIÓN

3.2 ERROR ESTÁNDAR

3.3 INTERVALO DE CONFIANZA

3.4 TAMAÑO DE MUESTRA

PROPIEDADES DE LA ESTIMACIÓN

- No tener sesgos
- Poca variabilidad de una muestra a otra

3.1 CONCEPTO DE ESTIMACIÓN

3.2 ERROR ESTÁNDAR

3.3 INTERVALO DE CONFIANZA

3.4 TAMAÑO DE MUESTRA

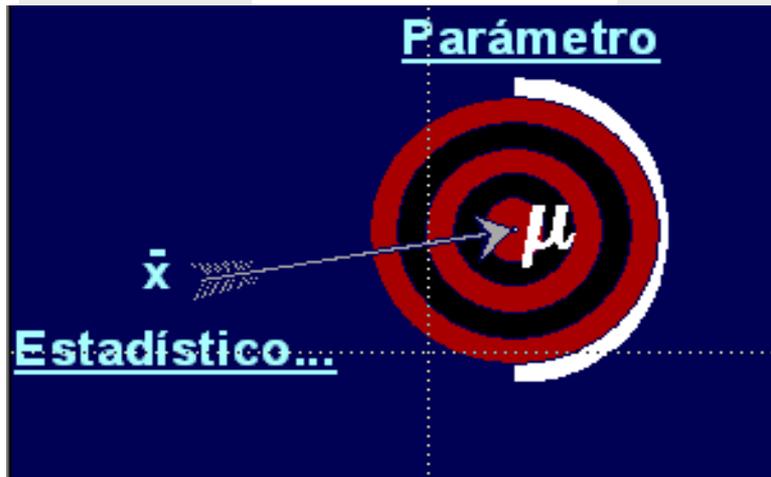
TIPOS DE ESTIMACIÓN

- **PUNTUAL:** Se obtiene un único número al que se le puede asignar un punto de la recta
- **POR INTERVALOS:** Se obtienen dos puntos que representan un límite inferior y superior (li , ls)

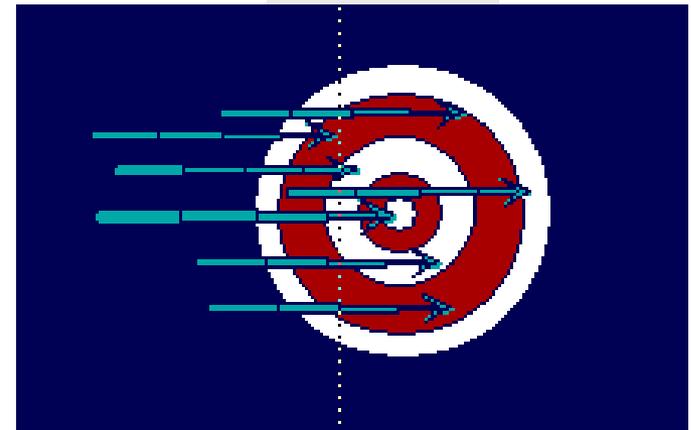
3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA

CONCEPTO DE ESTIMACIÓN

Ser un estimador adecuado no significa ...



significa ...

... manejo de la incertidumbre y
de la imprecisión

3.1 CONCEPTO DE ESTIMACIÓN

3.2 ERROR ESTÁNDAR

3.3 INTERVALO DE CONFIANZA

3.4 TAMAÑO DE MUESTRA

CONCEPTO DE ESTIMACIÓN

Un estimador puntual...

Difiere del verdadero valor

Por lo tanto...

Es deseable acompañar la estimación de alguna medida posible de error

3.1 CONCEPTO DE
ESTIMACIÓN

3.2 ERROR
ESTÁNDAR

3.3 INTERVALO DE
CONFIANZA

3.4 TAMAÑO DE
MUESTRA

DEFINICIÓN DE ERROR ESTÁNDAR

Diferencia entre el valor probable y los valores reales de la variable dependiente

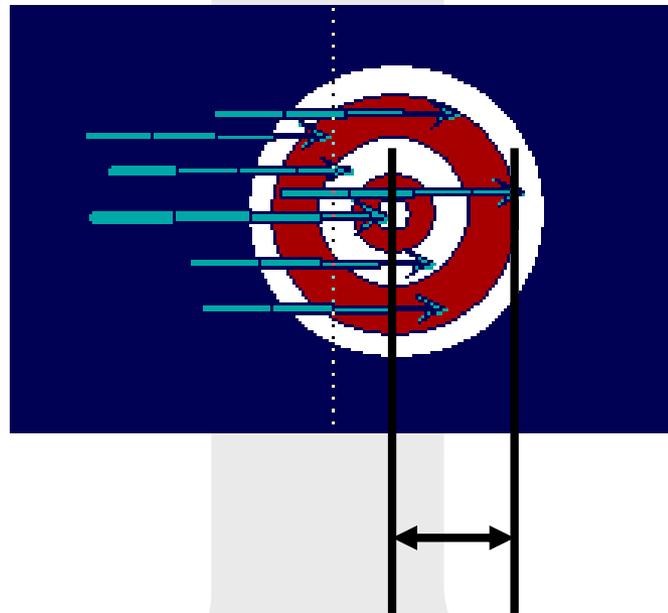
3.1 CONCEPTO DE
ESTIMACIÓN

3.2 ERROR
ESTÁNDAR

3.3 INTERVALO DE
CONFIANZA

3.4 TAMAÑO DE
MUESTRA

EL ERROR ESTÁNDAR ES...



3.1 CONCEPTO DE
ESTIMACIÓN

3.2 ERROR
ESTÁNDAR

3.3 INTERVALO DE
CONFIANZA

3.4 TAMAÑO DE
MUESTRA

TIPOS DE ERROR ESTÁNDAR

El error estándar puede ser de dos tipos

- ALEATORIO
- SISTEMÁTICO

3.1 CONCEPTO DE
ESTIMACIÓN

3.2 ERROR
ESTÁNDAR

3.3 INTERVALO DE
CONFIANZA

3.4 TAMAÑO DE
MUESTRA

ERROR ALEATORIO

Error inevitable que se produce por eventos únicos imposibles de controlar durante el proceso de medición

En un estudio de investigación, el error aleatorio viene determinado por el hecho de tomar sólo una muestra de una población para realizar inferencias

Puede disminuirse aumentando el tamaño de la muestra

3.1 CONCEPTO DE
ESTIMACIÓN

3.2 ERROR
ESTÁNDAR

3.3 INTERVALO DE
CONFIANZA

3.4 TAMAÑO DE
MUESTRA

ERROR SISTEMÁTICO

Error que se produce de igual modo en todas las mediciones que se realizan de una magnitud

Puede estar originado en un defecto del instrumento, en una particularidad del operador o del proceso de medición, etc

Se contrapone al concepto de error aleatorio

3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA

ESTIMACIÓN POR INTERVALOS

Asociado a cada
estimación
siempre hay

Un intervalo

Una medida de confianza

3.1 CONCEPTO DE
ESTIMACIÓN

3.2 ERROR
ESTÁNDAR

3.3 INTERVALO DE
CONFIANZA

3.4 TAMAÑO DE
MUESTRA

DEFINICIÓN DE INTERVALO DE CONFIANZA

Espacio que tiene una cierta probabilidad de contener el verdadero valor del parámetro desconocido

3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA

MEDIDA DE CONFIANZA

Coeficiente de
confianza

$$= 1 - \alpha$$

Nivel de
confianza

$$= 100 * (1 - \alpha) \%$$

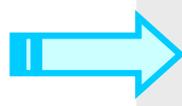
3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA

MEDIDA DE CONFIANZA

Elegiremos probabilidades cercanas a la unidad

Lo decidimos nosotros:

Probabilidad
del 95%



$$1 - \alpha = 0.95$$

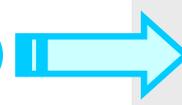


$$\alpha = 0.05$$

Probabilidad
del 90%



$$1 - \alpha = 0.90$$



$$\alpha = 0.10$$

Probabilidad
del 99%



$$1 - \alpha = 0.99$$



$$\alpha = 0.01$$

3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR**3.3 INTERVALO DE
CONFIANZA**3.4 TAMAÑO DE
MUESTRA

LOS INTERVALOS DE CONFIANZA...

- Se utilizan como indicadores de la variabilidad de las estimaciones.
- Cuánto más **“estrecho”** sea, mejor.

3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA

LOS INTERVALOS DE CONFIANZA...

- Se pueden crear para cualquier parámetro de la población.

EJEMPLOS

- Media: tiempo medio de recuperación
- Proporción: de niños que sufren apendicitis
- Desviación estándar: del error de medida de un aparato médico

3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA

FÓRMULA DEL INTERVALO DE CONFIANZA PARA LA MEDIA

Factor relacionado
con la confianza

Parámetro: Media Poblacional

Error Estándar

Estimo

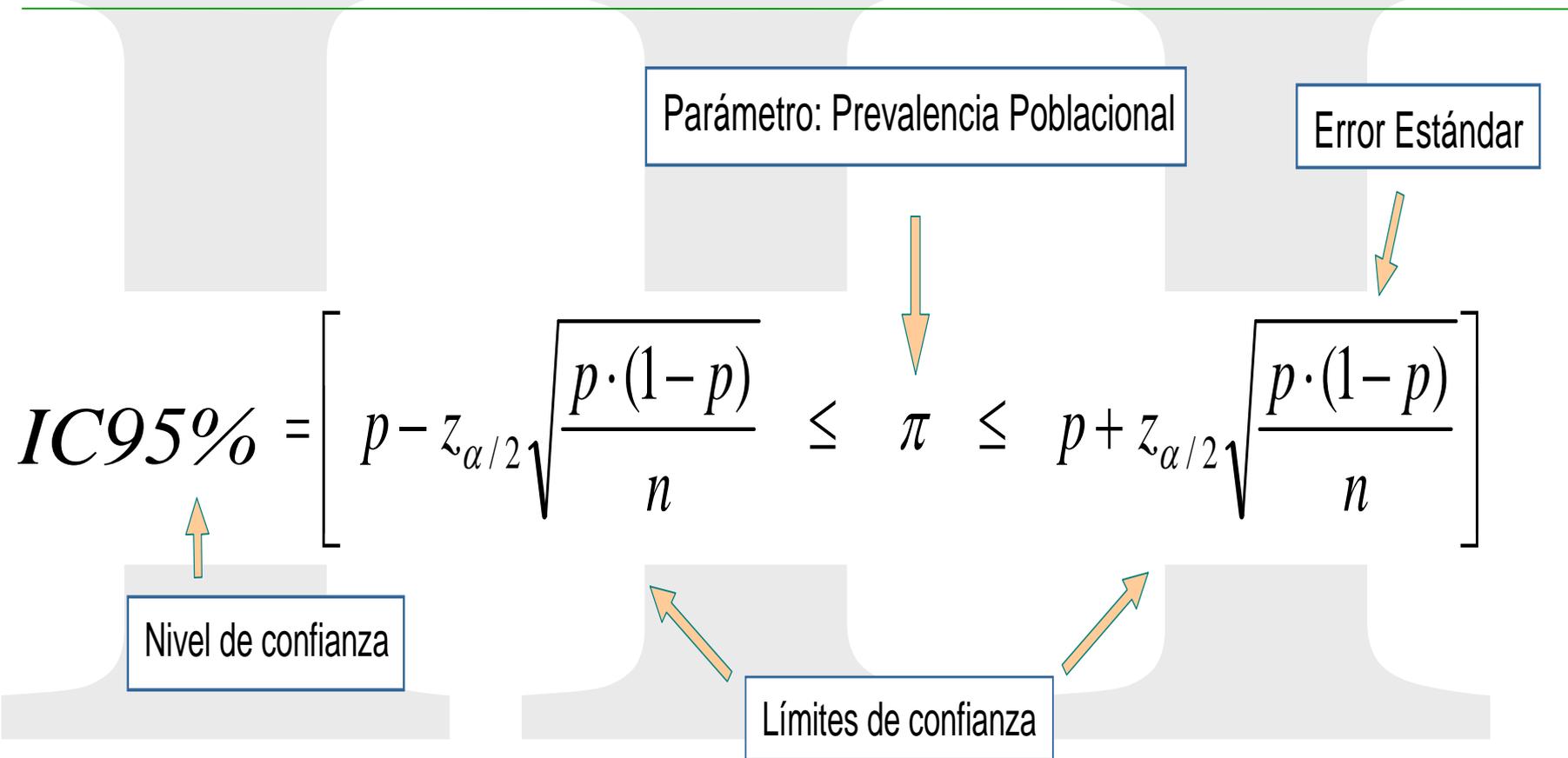
$$IC95\% = \left[\bar{x} - t_{n-1} \frac{S_{n-1}}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1} \frac{S_{n-1}}{\sqrt{n}} \right]$$

Nivel de
confianza

Límites de confianza

3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA

FÓRMULA DEL INTERVALO DE CONFIANZA PARA LA PROPORCIÓN



3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA

¿Cuántos individuos deberemos elegir para estimar una proporción?

Para contestar a esta pregunta debemos fijar a priori la precisión **D** deseada y a partir de ella estimar la proporción

FÓRMULA

$$N = \frac{Z_{1-\alpha/2}^2 \cdot P \cdot (1-P)}{D^2}$$

CONDICIONES DE
APLICACIÓN

La muestra debe ser extraída al azar de una población infinita (muy grande)

3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA¿Cómo se obtiene el valor de p ?

En la fórmula anterior, debemos sustituir p por un valor aproximado, obtenido según uno de los siguientes criterios:

 $p =$

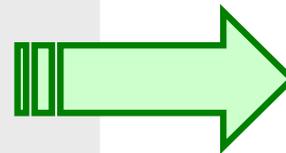
- **Estudios anteriores** → proporcionan una idea aproximada del valor de p
- Realizar un **sondeo previo** → se extrae una pequeña muestra y se sustituye p por la proporción p_0 observada en ella
- Sustituir el valor p por el **caso más desfavorable** → ocurre cuando $p=0,5$

3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA

EJEMPLO

“Se desea calcular el tamaño de muestra necesario para predecir, mediante un sondeo al azar, el resultado de una votación a nivel nacional con error inferior a 4%”

El resultado de un sondeo previo indica que la proporción de votos a favor es de 62%



Tomamos $p=0,62$

3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA

¿Cuántos individuos deberemos elegir para estimar una media?

Al igual que para el cálculo de la proporción debemos estimar a priori el error que estamos dispuestos a asumir.

FÓRMULA

$$n_{ca} = \frac{z_{\alpha/2}^2 \sigma^2}{e^2}$$

CONDICIONES DE
APLICACIÓN

La muestra debe ser extraída al azar de una población infinita (muy grande)

3.1 CONCEPTO DE
ESTIMACIÓN3.2 ERROR
ESTÁNDAR3.3 INTERVALO DE
CONFIANZA3.4 TAMAÑO DE
MUESTRA¿Cómo se obtiene el valor de σ ?

En la fórmula anterior, debemos sustituir σ por un valor aproximado, obtenido según uno de los siguientes criterios:

 $\sigma =$

- **Estudios anteriores** → proporcionan una idea aproximada del valor de σ
- Realizar un **sondeo previo** → se extrae una pequeña muestra y se sustituye p por la proporción σ_0 observada en ella
- Sustituir el valor p por el **caso más desfavorable** → ocurre cuando $\sigma=0,5$

INFERENCIA ESTADÍSTICA I

SESIÓN 4

- 4.1 Contraste de hipótesis
- 4.2 Métodos Paramétricos
- 4.3 Transformaciones de datos
- 4.4 Métodos no Paramétricos

DEFINICIÓN DE INFERENCIA ESTADÍSTICA

Métodos empleados para sacar conclusiones a partir de una muestra y extenderlas a una población

¡IMPORTANTE!

En la Inferencia Estadística, a diferencia que en la Estimación, primero se formula una hipótesis y después se comprueba su veracidad, aceptando o rechazando dicha hipótesis

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

MÉTODO DE CONTRASTE DE HIPÓTESIS

PASO 1

PASO 2

PASO 3

PASO 4

Expresar el interrogante de la investigación como una hipótesis estadística:

- H0: “No hay diferencia”
- H1: “Hay diferencia”

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

MÉTODO DE CONTRASTE DE HIPÓTESIS

PASO 1

PASO 2

PASO 3

PASO 4

Decidir sobre la prueba estadística adecuada

¿CÓMO? →

Según la población y el tipo de variables

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

MÉTODO DE CONTRASTE DE HIPÓTESIS

PASO 1

PASO 2

PASO 3

PASO 4

- **Seleccionar grado de significación para la prueba estadística.**
- **Grado de significación = alfa = probabilidad de rechazar de manera incorrecta H_0 cuando sea cierta (normalmente 0.05, 0.01, 0.001)**

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

MÉTODO DE CONTRASTE DE HIPÓTESIS

PASO 1

PASO 2

PASO 3

PASO 4

Realizar los cálculos y exponer conclusiones

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

ERRORES EN LAS PRUEBAS DE HIPÓTESIS

CONCLUSIÓN HIPÓTESIS	Existe diferencia H1	No existe diferencia H0
Existe diferencia	Poder 1-β	Falso positivo, α Error tipo I
No existe diferencia	Falso negativo, β Error tipo II	

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

ERRORES EN LAS PRUEBAS DE HIPÓTESIS

▪ β = Probabilidad de error tipo II

▪ Potencia = $1 - \beta$ || 

Los test serán más adecuados cuanto más potentes

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

POTENCIA DEL TEST

¿Cómo se puede aumentar la potencia de un test?



Aumentando el tamaño de la muestra

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

VALOR DE "P"



4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

MÉTODOS PARA EL CONTRASTE DE HIPÓTESIS

Si rechazo una hipótesis con un nivel de significación (α) de 0.001...

¿Se rechazaría a un nivel de 0.01?

SÍ

Porque si la rechazamos es porque $p < 0.001$, lo que implica que $p < 0.01$

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

MÉTODOS PARA EL CONTRASTE DE HIPÓTESIS

¿Y si fuera al revés: $p < 0,01$, puedo rechazar para $p < 0,001$...?

NO

En ese caso no puedo asegurar que la región de rechazo siga siendo la misma

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

MÉTODOS PARA EL CONTRASTE DE HIPÓTESIS

Si un test se rechaza con una potencia de 0.90...

¿se seguirá rechazando si rebajamos la potencia a 0.80?

NO

Estamos aumentando la probabilidad de error de tipo II (aceptar H_0 | H_0 falsa) y habría que rehacer el contraste

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

TIPOS DE DISEÑO



4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

T-Student

Es el Método Estadístico más utilizado

CONDICIONES

- La muestra se ajuste a un modelo lineal (una observación directa sigue un modelo lineal)
- Datos distribuidos normalmente e independientes
- Variable dependiente cualitativa dicotómica
- Variable independiente cuantitativa

T-Student

Datos distribuidos normalmente...

Es lo más frecuente

Bibliografía anterior

El Teorema Central del Límite nos asegura que cualquier muestra de un tamaño > 30 sigue una distribución normal

Test de contraste de normalidad

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

T-Student

Datos independientes...



No se puede asignar los dos tratamientos a la muestra



Condición indispensable: que las unidades de las variables sean homogéneas

T-Student

Los test de contraste de hipótesis pueden ser:

- 1. Unilaterales
- 2. Bilaterales

- Se usan cuando se sospecha que una media es mayor que otra
- En este caso el nivel de significación pasa a ser 2α

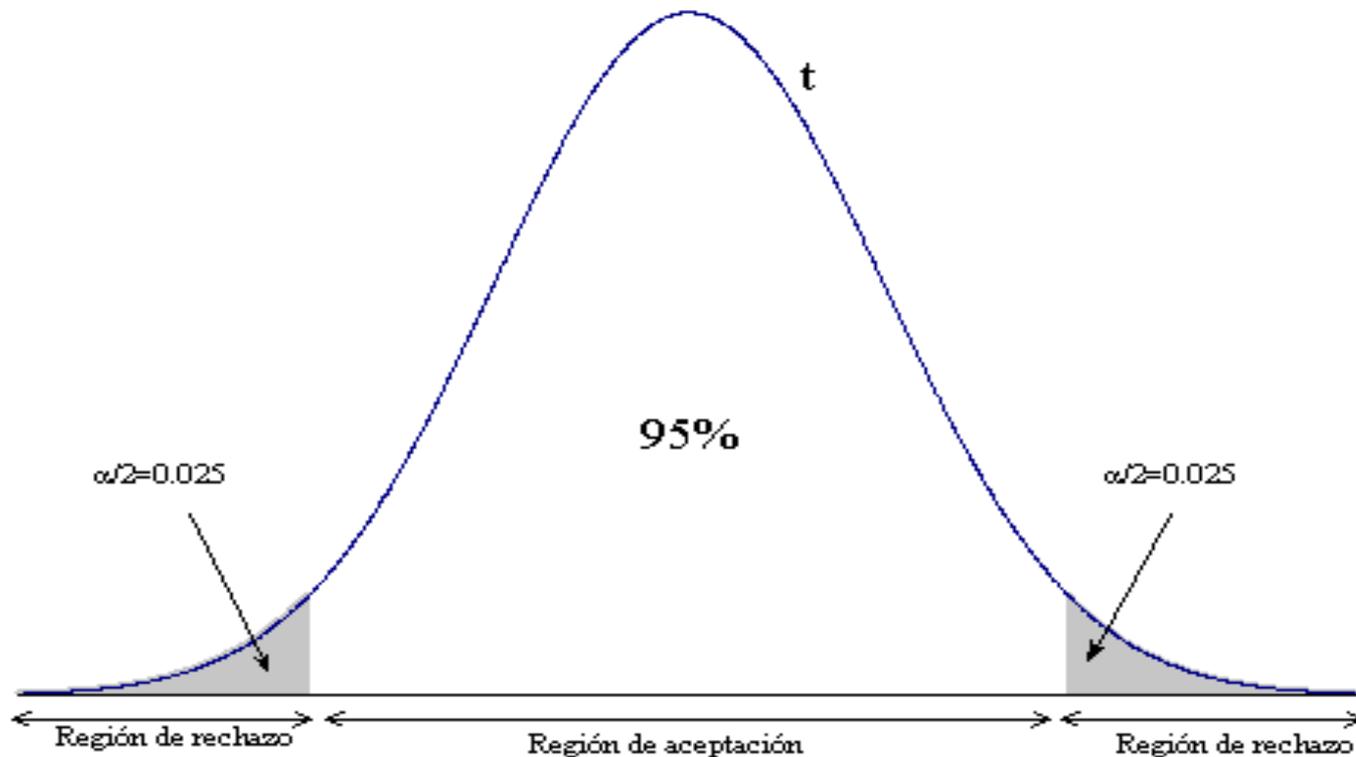
4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

REPRESENTACIÓN GRÁFICA DE LA T-Student



T-STUDENT CON EL SPSS

Ansiedad 2 [Conjunto_de_datos1] - Editor de datos SPSS

Archivo Edición Ver Datos Transformar **Analizar** Gráficos Utilidades Ventana ?

1 : sujeto

	sujeto	ansiedad	ten						
1	1	1							
2	2	1							
3	3	1							
4	4	1							
5	5	1							
6	6	1							
7	7	2							
8	8	2							
9	9	2							
10	10	2							
11	11	2							
12	12	2	2	16	12	8	8		
13									
14									
15									
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									
27									
28									
29									
30									
31									

Informes
 Estadísticos descriptivos
 Tablas
Comparar medias

- Medias...
- Prueba T para una muestra...
- Prueba T para muestras independientes...**
- Prueba T para muestras relacionadas...
- ANOVA de un factor...

Modelo lineal general
 Modelos mixtos
 Correlaciones
 Regresión
 Loglineal
 Clasificar
 Reducción de datos
 Escalas
 Pruebas no paramétricas
 Series temporales
 Supervivencia
 Respuesta múltiple

Vista de datos / Vista de variables /
 Prueba T para muestras independ. SPSS El procesador está preparado

Inicio curso guarda Microsoft PowerPoint ... Ansiedad 2 [Conjunto... 19:27

T-STUDENT CON EL SPSS

Ansiedad 2 [Conjunto_de_datos1] - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

Prueba T para muestras independientes

1 : sujeto 1

	sujeto	ansiedad	tensión						
1	1	1	1						
2	2	1							
3	3	1							
4	4	1							
5	5	1							
6	6	1							
7	7	2							
8	8	2							
9	9	2							
10	10	2							
11	11	2	2	16	14	10	9		
12	12	2	2	16	12	8	8		
13									
14									
15									
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									
27									
28									
29									
30									
31									

Contrastrar variables:

Variable de agrupación:

Definir grupos...

Opciones...

Aceptar
Pegar
Restablecer
Cancelar
Ayuda

Vista de datos / Vista de variables /

SPSS El procesador está preparado

Inicio curso guarda Microsoft PowerPoint ... Ansiedad 2 [Conjunto... 19:29

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

PRUEBA T-Student PARA MUESTRAS DEPENDIENTES

EJEMPLO

PACIENTE	PESO ANTES (kg)	PESO DESPUÉS (kg)
1	100	95
2	89	84
3	83	78
4	98	93
5	108	103
6	95	90

PRUEBA T-Student PARA MUESTRAS DEPENDIENTES

Extraemos una muestra de 3 individuos y calculamos su peso medio antes de la dieta

89 kg.

Extraemos **otra** muestra de 3 individuos y calculamos su peso después de la dieta

97 kg.

Concluimos por tanto que la dieta no es eficaz...

¡FALSO!

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

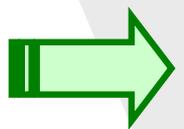
4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

¿DÓNDE ESTÁ EL FALLO...?

La variable de estudio, en este caso **el peso**, es muy distinta de un individuo a otro

¿Cómo controlar esta variabilidad?



Cogiendo una única muestra de pacientes y calculando su peso antes y después de la dieta

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

ESTUDIO DE DATOS APAREADOS

Hacemos un estudio de datos apareados cuando el mismo grupo se mide dos veces:



Los individuos se miden al principio del tratamiento para establecer una línea basal y después de alguna intervención se repite la medición **en los mismos sujetos**

4.1 CONTRASTE DE
HIPÓTESIS4.2 MÉTODOS
PARAMÉTRICOS4.3 TRANSFORMACIONES
DE DATOS4.4 MÉTODOS NO
PARAMÉTRICOS

ESTUDIO DE DATOS APAREADOS

Estos estudios necesitan una manera de controlar los datos entre pacientes. El objetivo es controlar factores extraños que podrían influir en el resultado

La prueba estadística que se utiliza cuando los mismos individuos son objeto de medición de una variable numérica es la **prueba t-Student para datos apareados**

Debemos asumir que la diferencia de las medias sigue una distribución normal

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

¿Y si queremos estudiar la influencia de más de un factor?



ANOVA

ANOVA

- El término “factor” se refiere a la variable por la cual se forman los grupos

POR EJEMPLO

Dividir en grupos en base a su estado de tiroides y terapia

- Al número de grupos definido por un factor se le conoce como **número de niveles**
- En estudios experimentales en medicina a los niveles se les llama **tratamiento**

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

ANOVA

- Muchos de los proyectos en medicina utilizan más de dos grupos
- Hay estudios que analizan la influencia de más de un factor
- Después se comprueban las distintas combinaciones para determinar las diferencias entre los grupos

4.1 CONTRASTE DE HIPÓTESIS

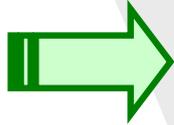
4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

ANOVA

P corregida



Si no se realiza la prueba múltiple, las distintas combinaciones entre los grupos alteran el nivel de significación α

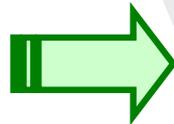
ANOVA

EJEMPLO

Supongamos que queremos estudiar las posibles diferencias entre las medias de 4 grupos dos a dos.

Tendríamos $4*2=8$ posibilidades de cometer un error de tipo I con un nivel $\alpha=0.05$

La posibilidad de que cada comparación significativa fuera falsa sería de 5%



El total de posibilidad de declarar una de las comparaciones como significativa, de forma incorrecta, sería de un 40%

ANOVA

Una manera de compensar las comparaciones múltiples es disminuir el nivel α , dividiendo α entre el número de comparaciones hechas

EJEMPLO

En el caso anterior, si se hacen 4 comparaciones, α se divide entre 4 para obtener una comparación de $0.05/4=0.0125$

Con este método cada comparación debe ser significativa al nivel de 0.0125 para declararla como tal

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

ANOVA

La forma de analizar los datos con observaciones múltiples se llama **ANOVA**

Éste método protege al investigador contra el “**error**” **inflación**, preguntando primero si hay diferencias entre las medias de los grupos

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

ANOVA - CARACTERÍSTICAS

Se asumen grupos de variable con distribución normal

Debe haber homogeneidad en las varianzas

Las variables son independientes, es decir, no se relaciona en forma alguna con el valor de otra

Si el **ANOVA** da significativo, estudiamos la diferencia entre las medias

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

Si las observaciones están muy sesgadas no debe emplearse la t-Student

¿QUÉ HACER?

En este caso las observaciones deben ser **transformadas o readaptadas**

También se pueden usar métodos no paramétricos

¿QUÉ ES UNA TRANSFORMACIÓN DE DATOS?

Transformar observaciones significa expresar los valores en otra escala

POR EJEMPLO

Si el peso se expresa en libras, se multiplica por 2.2 para obtener el valor en kg.

- De esta forma se pueden usar pruebas estadísticas que de otra manera no serían apropiadas.
- También resultan adecuadas cuando la dispersión es muy alta.

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

Las **transformaciones** más comunes son las **logarítmicas**, tanto en base 10 como logaritmo neperiano



¡OJO!
Hay que tener cuidado con los valores iguales a cero

Las transformaciones logarítmicas se emplean con frecuencia cuando se trata de valores de laboratorio que tienen distribución sesgada o cuando hay mucha dispersión

Otra transformación, menos utilizada, es la **raíz cuadrada**

4.1 CONTRASTE DE
HIPÓTESIS4.2 MÉTODOS
PARAMÉTRICOS4.3 TRANSFORMACIONES
DE DATOS4.4 MÉTODOS NO
PARAMÉTRICOS

¿QUÉ SON LOS MÉTODOS NO PARAMÉTRICOS?

Son pruebas estadísticas que no generan premisas sobre la distribución de las observaciones

Usar la prueba t requiere que se dé por supuesto que las diferencias siguen una distribución normal, lo cual es especialmente importante cuando los tamaños de muestra son pequeños ($n < 30$)

La mayoría de los test no paramétricos implican el análisis del rango de los datos, por lo que no se utilizan los valores de la muestra

4.1 CONTRASTE DE HIPÓTESIS

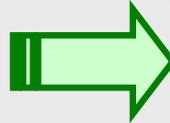
4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

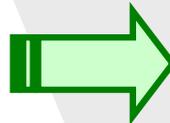
TIPOS DE VARIABLES EN LOS MÉTODOS NO PARAMÉTRICOS

Variable independiente



cuantitativa

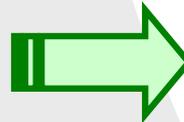
Variable dependiente



cualitativa dicotómica

MÉTODOS NO PARAMÉTRICOS PARA UN SOLO GRUPO

Variables independientes



Test de la prueba de signo

Variables dependientes



- Test de la prueba de signo aplicado a la diferencia (de medias, proporción,...)

- Prueba de Wilcoxon (U-Mann Whitney)

Se recomienda la prueba de Wilcoxon por encima de la prueba de signo para comparaciones pareadas, por ser más potente

4.1 CONTRASTE DE HIPÓTESIS

4.2 MÉTODOS PARAMÉTRICOS

4.3 TRANSFORMACIONES DE DATOS

4.4 MÉTODOS NO PARAMÉTRICOS

MÉTODOS NO PARAMÉTRICOS PARA VARIOS GRUPOS

La alternativa no paramétrica al ANOVA es el **contraste de Kruskal-Wallis**

Sirve para contrastar la hipótesis de que “k” muestras alternativas provienen de la misma población

En el caso de existir diferencias podemos hacer comparaciones a posteriori