

TEMA Nº 2 → CONTRASTE DE HIPÓTESIS EN LOS DISEÑOS DE UNA MUESTRA

1.- TIPOS DE CONTRASTE:

Contrastes paramétricos: Son aquellos que se relacionan con el estudio de un parámetro poblacional (media, varianza, proporción, correlación, etc.) y la variable de estudio proviene de una población con una **función de densidad de probabilidad conocida**. Ejemplo: Determinar si un parámetro poblacional puede tomar un valor concreto / Determinar si las variables medidas en la muestra son independientes en la población. Según Amón (1986):

- ❖ Las hipótesis recaen sobre los parámetros de la distribución de la variable.
- ❖ El nivel de medida es, al menos, de intervalo.
- ❖ Los supuestos sobre la forma de la distribución (normalidad) y/o sobre sus parámetros (varianza conocida, desconocida, etc.) son restrictivos

En general, **el contraste paramétrico es más potente** que su equivalente no paramétrico.

Contrastes no paramétricos: Son aquellos que **no se relacionan con parámetros poblacionales** o se refieren a datos que provienen de una población con una **función de densidad de probabilidad desconocida**. Ejemplo: Determinar la forma de distribución de una variable en la población / Determinar si los datos observados en la muestra son independientes entre sí. Según Amón (1986):

- ❖ Nivel de medida inferior al intervalo.
- ❖ Solo se establecen supuestos generales sobre la distribución (continuidad, simetría, etc.)

En general, la utilización de contrastes no paramétricos **disminuye la potencia de contraste** (probabilidad de aceptar la H_1 cuando esta es cierta)

2.- PASOS DE CONTRASTE DE HIPÓTESIS →

Un **contraste de hipótesis** es un procedimiento para decidir si una hipótesis se acepta como válida o se rechaza. Los principales pasos para tomar la decisión son:

1. Establecer los supuestos.
2. Formular las hipótesis estadísticas
3. Elección del estadístico de contraste apropiado (discrepancia entre los datos del estadístico de la muestra y el parámetro planteado en la hipótesis poblacional),
4. Regla de decisión (la distribución muestral del estadístico nos permite fijar los valores críticos que determinan la zona de rechazo de la H_0). El cálculo del estadístico de contraste nos permite también la determinación del nivel crítico p (probabilidad de que, siendo cierta la H_0 , obtengamos unos datos como los observados en la muestra)
5. Conclusión (rechazamos o no la H_0 a partir del nivel crítico p , o del valor o los valores críticos)
6. Interpretación (en el contexto de la investigación)

3.- CONTRASTE SOBRE LA TENDENCIA CENTRAL DE LOS DATOS →

PARAMÉTRICOS	MEDIA POBLACIONAL CONOCIDA σ^2 DESCONOCIDA σ^2
NO PARAMÉTRICOS	PRUEBA DE LOS SIGNOS PRUEBA DE WILCOXON

CONTRASTES PARAMÉTRICOS SOBRE LA MEDIA POBLACIONAL

Supuestos (Razones que justifican su elección): Disponemos de una muestra aleatoria de “n” observaciones aleatorias e independientes. La variable dependiente sigue una distribución normal en la población y está medida, al menos, a nivel de intervalo. Podemos conocer o no el valor de la varianza poblacional (también aplicable a cualquier otro tipo de distribución, con varianza desconocida, si $n \geq 30$, según el Teorema Central del Límite)

Hipótesis Estadísticas: μ_0 es el valor concreto de la media poblacional acerca del cual queremos comprobar si μ es diferente, menor o mayor // Se aconseja formular primero H_1

Contraste Bilateral:	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$
Contraste Unilateral Izquierdo	$H_0: \mu \geq \mu_0$ $H_1: \mu < \mu_0$
Contraste Unilateral Derecho	$H_0: \mu \leq \mu_0$ $H_1: \mu > \mu_0$

Estadístico de contraste:

Conocida la varianza de la población $\rightarrow Z = (\bar{Y} - \mu_0) / \sigma / \sqrt{n}$

$\bar{Y}$ es la media muestral // $\sigma / \sqrt{n}$ es el error típico de la media

μ_0 es el valor propuesto para la media poblacional por H_0

El Estadístico de Contraste sigue la distribución $N(0,1)$

Desconocida la varianza de la población $\rightarrow T = (\bar{Y} - \mu_0) / S_{n-1} / \sqrt{n}$

$S_{n-1}^2 = (Y - \bar{Y})^2 / n - 1$ es la cuasi varianza muestral (estimador insesgado de varianza poblacional)

El Estadístico de Contraste sigue la distribución T de Student con $n - 1$ gl. (Cuando gl > 30 la distribución T de Student se aproxima a la normal).

Relación entre la varianza y la cuasi varianza muestral $\rightarrow (n-1) S_{n-1}^2 = n S^2$ // Luego $S_{n-1}^2 = \frac{n S^2}{n - 1}$

Intervalo de confianza:

$$\bar{Y} \pm Z_{\alpha/2} \cdot \sigma / \sqrt{n} = L_S \text{ y } L_I \text{ (conocida } \sigma^2 \text{ poblacional)}$$

$$\bar{Y} \pm T_{n-1; \alpha/2} \cdot S_{n-1} / \sqrt{n} = L_S \text{ y } L_I \text{ (desconocida } \sigma^2 \text{ poblacional)}$$

CONTRASTES NO PARAMÉTRICOS SOBRE LA MEDIANA POBLACIONAL

PRUEBA DE LOS SIGNOS: Considera que el valor de la mediana bajo la H_0 es correcto; por tanto, en la población, existirán similares valores por encima y por debajo (mismo número de signos positivos y negativos al calcular la diferencia entre cada valor y la mediana propuesta en la H_0)

Supuestos (Razones que justifican su elección): muestra aleatoria de “n” observaciones aleatorias e independientes de una población con mediana desconocida. Contraste sobre el Promedio Poblacional. Desconocemos el cumplimiento de los supuestos paramétricos. Asumimos que la distribución de la variable es **continua** por lo que no debemos encontrar diferencias nulas, si aparecen, se descartan). Datos directos, sin agrupar, medidos, al menos, en una escala ordinal.

Hipótesis Estadísticas:

Contraste Bilateral:	$H_0: \eta = \eta_0$ $H_1: \eta \neq \eta_0$
Contraste Unilateral Izquierdo	$H_0: \eta \geq \eta_0$ $H_1: \eta < \eta_0$
Contraste Unilateral Derecho	$H_0: \eta \leq \eta_0$ $H_1: \eta > \eta_0$

η_0 es el valor concreto de la mediana poblacional // Se aconseja formular primero la H_1

Estadístico de contraste: (S_+ ó S_-) se obtienen al restar cada una de las observaciones del valor de la mediana propuesto por la H_0 . Si alguna diferencia es nula, se elimina, reduciéndose el tamaño de la muestra.

Cuando la muestra es grande se aproxima a la normal mediante $[S_+ \text{ (ó } S_-) - n/2] / \sqrt{n/4}$

Distribución muestral del estadístico: modelo binomial con parámetros $n = S_+ + S_-$ y $\pi = 0.50$.

PRUEBA DE WILCOXON: Conocida como “prueba de rangos y signos”, gana en precisión con respecto a la prueba de los Signos, ya que aprovecha mejor la información disponible; tiene en cuenta, a la vez, el signo de las diferencias y la magnitud de las mismas. Wilcoxon es una prueba más precisa.

Supuestos (Razones que justifican su elección): Disponemos de una muestra aleatoria de “n” observaciones aleatorias e independientes. Se trata de un contraste sobre el Promedio Poblacional. Desconocemos el cumplimiento de los supuestos paramétricos. La distribución de la variable en la población ha de ser **simétrica y continua**. Datos directos, sin agrupar. Nivel de medida de los datos, al menos, de intervalo.

Hipótesis Estadísticas: similares a la prueba de signos

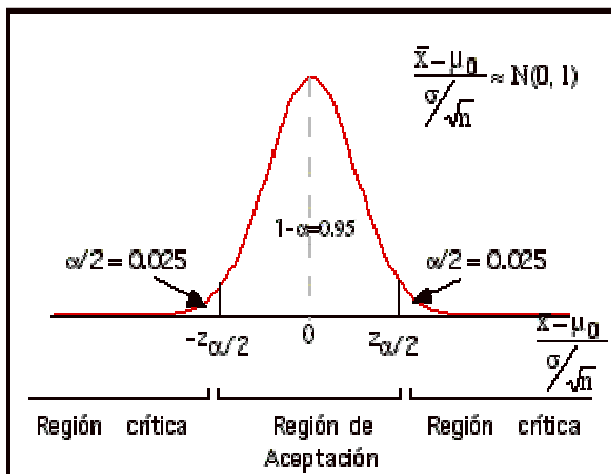
Cálculo del Estadístico de Contraste: ($\sum R_+$ ó $\sum R_-$) suma de rangos positivos y negativos. Conocemos los datos obtenidos en la muestra // Calculamos las diferencias entre cada observación y la mediana propuesta por la H_0 . Se eliminan las diferencias nulas, reduciendo el tamaño muestral “n” // Estas diferencias, en valor absoluto, se ordenan de menor a mayor, y se sustituye cada una de ellas por el rango que le corresponde // Si hay empates, se adopta el “criterio del rango medio”; es decir, se le atribuye a cada una el rango promedio de los rangos que les hubieran correspondido de ser diferentes // A cada rango se le añade el signo + o – de la puntuación de diferencia correspondiente y se suman, obteniendo R_+ y R_- .

Distribución muestral del estadístico: distribución de valores de Wilcoxon. Para utilizar la tabla W de Wilcoxon se selecciona el sumatorio más pequeño y se toma la decisión mediante los valores críticos.

PROBLEMAS EJEMPLO

CONTRASTE SOBRE LA MEDIA (VARIANZA POBLACIONAL CONOCIDA)

Supongamos que pertenecemos al consejo regulador de la denominación de origen de los vinos de Ribera de Duero. Sabemos que los vinos jóvenes de años anteriores tienen un grado alcohólico medio de 12.5 grados y una varianza de 0.5^2 y que los datos en la población se ajustan a una distribución normal $N(12.5, 0.5)$. Para el año actual, el consejo regulador, de acuerdo con todos sus miembros, ha decidido cambiar algunos de los pasos del proceso de fabricación. El primer problema que se plantea es: ¿Se ha modificado el grado alcohólico al modificar el proceso de fabricación? La definición del problema a estudiar nos permite determinar la población que queremos estudiar, los vinos jóvenes de ribera de Duero en el año actual; la variable que queremos medir, el grado alcohólico de los mismos, y la hipótesis de trabajo inicial ¿Se ha modificado el grado alcohólico? Para realizar el contraste de hipótesis, extraemos una muestra aleatoria de 14 observaciones que aportan una media de 12.529 grados y una cuasi varianza de 0.338^2 establecemos un nivel de confianza del 95%.



Supuestos: Se trata de una muestra aleatoria con datos medidos, al menos, en una escala de intervalo. Asumimos que la variable se distribuye normalmente en la población con varianza conocida.

Hipótesis: Contraste bilateral $H_0: \mu = 12.5$
 $H_1: \mu \neq 12.5$

Se trata de determinar si los datos de la muestra arrojan evidencia a favor o en contra de la H_0 .

Estadístico de contraste: Discrepancia entre el estimador (muestra) y el parámetro de la H_0 .

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{12.529 - 12.5}{0.5 / \sqrt{14}} = 0.217$$

Valores críticos: con $\alpha = 0.05$ y un contraste bilateral $\rightarrow \pm 1.96$ (región de aceptación)

Decisión estadística: El valor del estadístico de contraste (0.217) se sitúa en la región de aceptación (entre los valores críticos); por tanto aceptamos H_0 . El nivel crítico (p) asociado al E. Contraste (0.217) es $2 \cdot 0.4129 = 0.8258$ ($1 - 0.5871 = 0.4129$) que resulta mayor que el nivel de significación fijado ($\alpha = 0.05$); por tanto, aceptamos H_0 . A partir del intervalo de confianza, aceptamos H_0 (12.5 está incluido)

$$12.529 \pm (1.96 \cdot 0.5 / \sqrt{14}) \rightarrow (12.27 \text{ y } 12.79)$$

Conclusión: El cambio en el proceso de fabricación no ha modificado significativamente el grado alcohólico.

Interpretación: Con la información de la que disponemos no hemos encontrado evidencia suficiente de que la media sea distinta de 12.5

Vamos a plantear **otro supuesto** que modificaría la definición de las hipótesis $\rightarrow$ Sospechamos, antes de realizar el experimento, que la modificación en el proceso de fabricación aumenta el grado alcohólico.

Supuestos: Similares; **Hipótesis** (Contraste unilateral derecho): $H_0: \mu = 12.5$ y $H_1: \mu > 12.5$

Estadístico de contraste: Similar

Valores críticos: con $\alpha = 0.05$ y un contraste unilateral derecho $\rightarrow 1.65$

Decisión estadística: El valor del estadístico de contraste (0.217) se sitúa en la región de aceptación (por debajo del valor crítico 1.65); por tanto aceptamos H_0 .

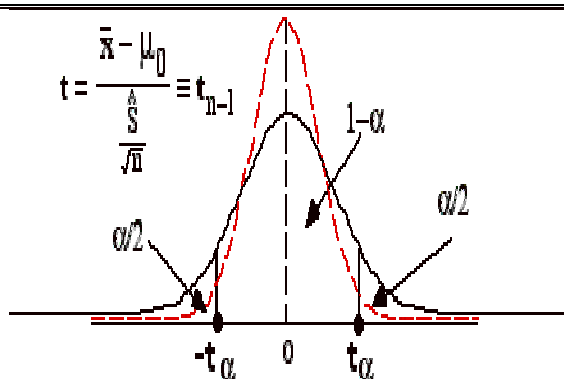
El nivel crítico (p): asociado al estadístico de contraste (0.217) es **0.4129** ($1 - 0.5871$) que resulta mayor que el nivel de significación fijado ($\alpha = 0.05$); por tanto, aceptamos H_0

Conclusión: El cambio en la fabricación no ha modificado significativamente el grado alcohólico.

Interpretación: Con la información de la que disponemos no hemos encontrado evidencia suficiente de que la media no sea igual a 12.5

CONTRASTE SOBRE LA MEDIA (VARIANZA POBLACIONAL DESCONOCIDA)

Siguiendo con el **ejemplo del grado alcohólico de los vinos** de la denominación de origen de Ribera de Duero, los resultados del contraste bilateral desconocida la varianza de la población (suponiendo que no nos hubieran facilitado este parámetro), serían:



Supuestos: Se trata de una muestra aleatoria con datos medidos, al menos, en una escala de intervalo. Asumimos que la variable se distribuye normalmente en la población con varianza desconocida.

Hipótesis: Contraste bilateral $H_0: \mu = 12.5$
 $H_1: \mu \neq 12.5$

Se trata de determinar si los datos de la muestra arrojan evidencia a favor o en contra de la H_0 .

Estadístico de contraste: Discrepancia entre el estimador (muestra) y el parámetro de la H_0 .

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{12.529 - 12.5}{0.338 / \sqrt{14}} = 0.316$$

Valores críticos: con $\alpha = 0.05$ y un contraste bilateral $\rightarrow t_{18, 0.025} = \pm 2.101$ (región aceptación)

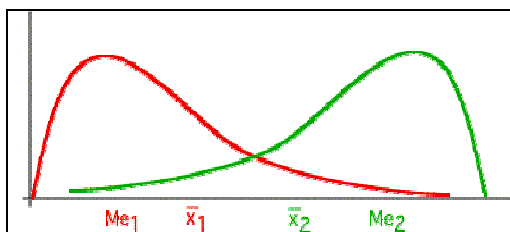
Decisión estadística: El valor del estadístico de contraste (0.316) se sitúa en la región de aceptación (entre los valores críticos); por tanto aceptamos H_0 . El nivel crítico (p) asociado al E. Contraste (0.316) es aproximadamente $(1 - 0.65 = 0.35) \rightarrow 0.35 \cdot 2 = 0.7 > 0.05$ (Aceptamos H_0). A partir del intervalo de confianza, aceptamos H_0 (12.5 está incluido)

$$12.529 \pm (1.96 \cdot 0.338 / \sqrt{14}) \rightarrow (12.35 \text{ y } 12.705)$$

Conclusión: El cambio en el proceso de fabricación no ha modificado significativamente el grado alcohólico.

Interpretación: Con la información de la que disponemos no hemos encontrado evidencia suficiente de que la media sea distinta de 12.5

CONTRASTES NO PARAMÉTRICOS (PRUEBA DE LOS SIGNOS Y PRUEBA DE WILCOXON)



En la figura se muestra como para distribuciones asimétricas es mucho más intuitiva la comparación de las medianas que la comparación de las medias, ya que estas están afectadas por los valores muy extremos de la distribución.

Problema ejemplo: Una determinada empresa utiliza, desde hace muchos años, un test de aptitudes para seleccionar a sus empleados. Para dicho test tiene establecido que la mediana está en 180. El psicólogo de la empresa quiere probar que efectivamente la mediana sigue siendo correcta (probar la H_0 de que la mediana en la población es igual a 180). Para ello obtiene una muestra de 15 trabajadores y les aplica el test obteniendo los siguientes resultados:

189 233 195 160 212 176 231 202 193 174 166 248 199 213 185

Suponemos que la variable en la población es simétrica y continua (medida en una escala de intervalo) y que no sabemos nada acerca de la forma de la población. Con estos datos y únicamente con fines pedagógicos aplicaremos las pruebas de los Signos y la de Wilcoxon (sería directamente aplicable Wilcoxon).

Supuestos → Muestra aleatoria de $n = 15$ observaciones. Distribución de la variable en la población (simétrica y continua). Nivel de medida de los datos, al menos, escala de intervalo.

Hipótesis (C. Bilateral) → $H_0: \eta = 180$ y $H_1: \eta \neq 180$ // $\alpha = 0,05$

Estadísticos de Contraste: →

Signos: Previamente ordenamos las puntuaciones (la mediana es un índice de posición). Restando 180 a cada una de las 15 puntuaciones obtenidas por los sujetos, obtenemos 11 (S_+) y 4 (S_-)

160	166	174	176	185	189	193	195	199	202	212	213	231	233	248	ORDEN
-20	-14	-6	-4	5	9	13	15	19	22	32	33	51	53	68	RESTA
-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	SIGNOS

Probabilidad binomial con $n = 15$ (participantes) y $p = 0,5$ (probabilidad de éxito si la H_0 es cierta)

Wilcoxon: Más precisa. Después de ordenar las puntuaciones de menor a mayor y ordenar las diferencias con la mediana en valor absoluto, obtenemos:

160	166	174	176	185	189	193	195	199	202	212	213	231	233	248	ORDEN
-20	-14	-6	-4	5	9	13	15	19	22	32	33	51	53	68	RESTA
-4	5	-6	9	13	-14	15	19	-20	22	32	33	51	53	68	ORDEN
-	+	-	+	+	-	+	+	-	+	+	+	+	+	+	SIGNOS
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	RANGOS

Sumando los rangos cuyas diferencias tienen valor positivo y negativo:

$R_+ (2+4+5+7+8+10+ 11+12+13+14+15 = 101)$ y $R_- (1+3+6+9 = 19)$

Utilizaremos como estadístico $W \rightarrow R_- (1+3+6+9 = 19)$ dado que es el más pequeño $\rightarrow W = 19$

Decisión estadística:

Signos: Función de distribución Binomial (Tabla II); para $n = 15$ y probabilidad = 0,50, encontramos el valor 0,0592 $\rightarrow P(S_- \leq 4) = 0,0592$. Para averiguar el nivel crítico p , multiplicamos por 2 la probabilidad asociada al estadístico de contraste ($0,0592 \cdot 2 = 0,1184$) dado que es bilateral.

Wilcoxon: Tablas $\rightarrow$ para un contraste bilateral con $\alpha/2 = 0,025$ y $n = 15$; el **valor crítico es $\rightarrow 26$**

Conclusión:

Signos: Como $p (0,1184)$ es mayor que $\alpha (0,05)$ **Mantenemos H_0** . También tabla (Test de signos dos direcciones) para $n = 15$ y $\alpha = 0,05$, y un contraste bilateral, el valor crítico es igual a 3 $\rightarrow$ como $S_- = 4 > 3$ (Mantenemos H_0)

Wilcoxon (más precisa): Como 19 (se toma la suma menor) $< 26 \rightarrow$ **Rechazamos la H_0**

Interpretación:

Signos: Teniendo en cuenta los resultados, **es posible** seguir manteniendo que la Mediana de la población de referencia es 180, con un $\alpha = 0,05$.

Wilcoxon: Teniendo en cuenta los resultados, **no es posible** seguir manteniendo que la Mediana de la población de referencia es 180, con un $\alpha = 0,05$.

4.- CONTRASTE SOBRE LA VARIANZA POBLACIONAL →

Supuestos (Razones que justifican su elección): Disponemos de una muestra aleatoria de “n” observaciones aleatorias e independientes. La variable dependiente sigue una distribución normal y está medida, al menos, a nivel de intervalo.

Hipótesis Estadísticas:

Contraste Bilateral:	$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 \neq \sigma_0^2$
Contraste Unilateral Izquierdo	$H_0: \sigma^2 \geq \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$
Contraste Unilateral Derecho	$H_0: \sigma^2 \leq \sigma_0^2$ $H_1: \sigma^2 > \sigma_0^2$

σ_0^2 es el valor concreto de la varianza poblacional acerca del cual queremos comprobar si la varianza poblacional es diferente, menor o mayor // Se aconseja formular primero H_1 .

Estadístico de contraste → $X^2 = (n-1) S_{n-1}^2 / \sigma_0^2$

σ_0^2 es el valor propuesto para la varianza poblacional por H_0 y S_{n-1}^2 = la cuasi varianza muestral. El Estadístico de Contraste sigue una distribución Chi-Cuadrado con (n-1) gl.

$$\text{Intervalo de confianza} \rightarrow \frac{(n-1) S_e^2}{X^2_{(n-1) (1-\alpha/2)}} < \sigma^2 < \frac{(n-1) S_e^2}{X^2_{(n-1) \alpha/2}}$$

Si el contraste se hace a partir de la varianza poblacional, se sustituye (n-1) por n.

5.- CONTRASTE SOBRE LA PROPORCIÓN POBLACIONAL →

Supuestos (Razones que justifican su elección): Disponemos de una muestra aleatoria de “n” observaciones aleatorias e independientes. La variable dependiente es dicotómica. La proporción es el cociente entre el número de veces que aparece la observación dividido entre el número total de observaciones)

Hipótesis Estadísticas:

Contraste Bilateral:	$H_0: \pi = \pi_0$ $H_1: \pi \neq \pi_0$
Contraste Unilateral Izquierdo	$H_0: \pi \geq \pi_0$ $H_1: \pi < \pi_0$
Contraste Unilateral Derecho	$H_0: \pi \leq \pi_0$ $H_1: \pi > \pi_0$

π_0 es el valor concreto de la proporción poblacional acerca del cual queremos comprobar si la proporción poblacional es diferente, menor o mayor // Se aconseja formular primero H_1

Estadístico de contraste: $Z = (P - \pi_0) / \sqrt{(\pi_0 (1 - \pi_0) / n)}$

P es la proporción muestral de la característica que interesa.

π_0 es el valor propuesto para la proporción poblacional por la H_0 .

σ_0 es el error típico de la distribución muestral de la proporción.

El Estadístico de Contraste sigue la distribución N (0,1) cuando “n” es suficientemente grande, lo que significa que “n · p” > 5 ó n = 25. Cuando no se cumple esta condición (muestras pequeñas), se utiliza la distribución binomial.

Intervalo de Confianza → $P \pm |Z_{\alpha/2}| \sqrt{(\pi_0 (1 - \pi_0) / n)}$

PROBLEMAS EJEMPLO

Contraste de hipótesis sobre la varianza: Los resultados de un reciente estudio indican que 81 estudiantes de primero de Psicología de la UNED, elegidos mediante muestreo aleatorio simple, estudian una media de 6 horas a la semana, con una cuasi desviación típica de 3. La población universitaria europea estudia, también, una media de 6 horas a la semana, pero con una desviación típica de 2. Queremos saber si la variabilidad encontrada en nuestra muestra es estadísticamente diferente de la de la población de estudiantes europeos. Trabajaremos a un nivel de confianza del 95 %.

Supuestos: Muestra aleatoria de 81 estudiantes de la UNED ($\bar{X} = 6$ y $S_x = 3$). Datos de la población de estudiantes europea ($\mu = 6$ y $\sigma = 2$) para $\alpha = 0,05$. La variable dependiente está medida, al menos, a nivel de intervalo. Se plantea un CH sobre la varianza.

Hipótesis: Contraste bilateral (resultado estadísticamente diferente). Las hipótesis siempre se plantean sobre la población (la población universitaria europea tiene una $\sigma = 2$; luego la varianza de la población a contrastar será $\rightarrow \sigma^2 = 4$. Luego $\rightarrow H_0: \sigma^2 = 4; H_1: \sigma^2 \neq 4$

Estadístico de contraste: $\chi^2 = (n - 1) S^2_{n-1} / \sigma^2_0 \rightarrow \chi^2 = 80 \cdot 9 / 4 = 180$

Decisión estadística:

- A partir de los valores críticos (80 gl; $\alpha/2 = 0,025$ y $1-\alpha/2 = 0,975$) $\rightarrow 57'15$ y $106'63$. Rechazamos la H_0 porque el valor muestral del estadístico de contraste (180) > el valor crítico superior ($\chi^2 = 106'63$)
- Si tenemos en cuenta el nivel crítico p ($\chi^2 = 180$ con 80 gl $\rightarrow$ encontramos que p es < 0'005) y dado que $p < \alpha$ (0'05) $\rightarrow$ Rechazamos H_0 .
- Considerando el intervalo de confianza: Rechazamos H_0 ($\sigma^2 = 4$ no está en el intervalo)

Intervalo de confianza $\rightarrow \frac{80 \cdot 9}{106'63} < \sigma^2 < \frac{80 \cdot 9}{57'11}$ (6'75 y 12'61)

Conclusión e interpretación: Dado que hemos rechazado la H_0 con $\alpha = 0,05 \rightarrow$ La variabilidad en el tiempo de estudio que emplean los estudiantes de Psicología de 1º de la UNED no es igual que la variabilidad en el tiempo de estudio semanal de la población universitaria europea.

Contraste de hipótesis sobre la proporción (Febrero 2008): Se quiere conocer la postura de la población española frente a la inmigración ilegal. Se fija $\alpha = 0'04$, se extraen aleatoria e independientemente 750 españoles y se les pide que manifiesten si están a favor o en contra. De los 750 españoles, 435 se manifiestan en contra y 315 a favor. Se desea estudiar si son compatibles estos resultados con la hipótesis que establece que las 3/5 partes de la población española están en contra de la inmigración ilegal.

Supuestos: Variable dependiente $\rightarrow$ postura de la población española frente a la inmigración ilegal (a favor o en contra). Se trata de una variable cualitativa dicotómica.

$P = (435/750 = 0'58)$; se cumple ($n \cdot p > 5$) $\rightarrow (750 \cdot 0'58 = 435)$ Muestra grande (aproximación a la normal $\rightarrow$ CH Proporción

Hipótesis: Contraste bilateral (no se especifica ninguna dirección a favor o en contra). Sabemos que $\pi = 3/5 = 0,6$ y la hipótesis nula se centra si nuestros datos son compatibles con este resultado (teóricamente poblacional). Luego $\rightarrow H_0: \pi = 0'6; H_1: \pi \neq 0'6$

Estadístico de contraste: Nuestra proporción es $\rightarrow 435/750 = 0'58$. Con estos datos aplicamos la fórmula del estadístico de contraste $\rightarrow Z = (0'58 - 0'6) / (\sqrt{0'6 \cdot 0'4 / 750}) = -0'02 / 0'0179 = -1'12$.

Valores críticos: Al tratarse de un contraste bilateral deben ser dos valores. Buscando en las tablas de la curva normal con $\alpha = 0'04$ (bilateral $\alpha/2 = 0'02$ y $1-\alpha/2 = 0'98$) encontramos las puntuaciones típicas (- 2'05 y 2'05)

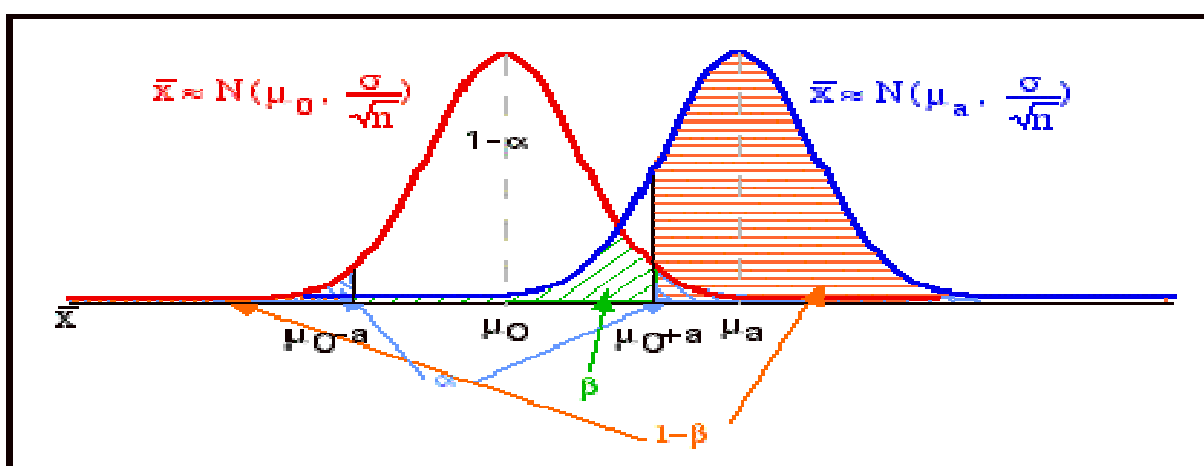
Decisión estadística:

- Teniendo en cuenta los anteriores valores críticos (- 2'05 y 2'05), Aceptamos la H_0 porque el valor muestral del estadístico de contraste (- 1'12) > el valor crítico inferior (- 2'05)
- Si tenemos en cuenta el nivel crítico p ($Z_{(-1'12)} = 0'1314$; luego $2 \cdot 0'1314 = 0'2628$). Como p (0'2628) > que α (0'04) $\rightarrow$ Aceptamos H_0 .
- Considerando el intervalo de confianza: Aceptamos H_0 (0'6 está en el intervalo)

$$P \pm |Z_{\alpha/2}| \sqrt{(\pi_0 (1 - \pi_0) / n)} \rightarrow 0'58 \pm 2'05 \sqrt{(0'58 (1 - 0'58) / 750)} = (0'543 \text{ y } 0'617)$$

Conclusión e interpretación: Dado que hemos aceptado la hipótesis nula ($\pi = 0'6$) podemos afirmar que el 60% de la población española están en contra de la inmigración ilegal; por tanto, el 40% estará a favor.

6.- CÁLCULO DE LA POTENCIA DE CONTRASTE $\rightarrow$



Potencia de contraste ($1-\beta$): Representa la probabilidad de rechazar la H_0 siendo falsa = la probabilidad de aceptar una H_1 verdadera (probabilidad de obtener un resultado estadísticamente significativo). $1-\beta$ = Probabilidad de no cometer Error tipo II. La disminución de (α = Error de tipo I provoca un aumento automático de β = Error tipo II)

Problema ejemplo: La aplicación de un test entre docentes tiene, por experiencias anteriores, una media de 55 y una varianza de 121, **N (55, 121)**. Un psicólogo educativo considera que en la actualidad el promedio se ha incrementado, pasando a ser de 60 puntos. Para contrastar esta hipótesis somete al test a 50 individuos extraídos aleatoriamente de la población y obtiene una media de 58. Fijando un nivel de significación $\alpha = 0,01$. **¿Puede aceptarse la hipótesis del psicólogo?**

Hipótesis estadísticas (contraste unilateral derecho): $H_0: \mu_0 = 55$ y $H_1: \mu_1 > 55$ ($\mu_1 = 60$)

Descripción de la variable y de los datos muestrales: Variable medida a nivel de intervalo, distribución normal, con varianza conocida **N (55, 121)** y observaciones aleatorias e independientes.

Estadístico de contraste y cálculo: Bajo la $H_0: \mu_0 = 55$, tipificamos la media muestral = 58.

$$Z = (\bar{Y} - \mu_0) / \sigma_Y \rightarrow Z = (58 - 55) / (11 / \sqrt{50}) = 1,93$$

Decisión estadística:

- A partir del **valor crítico** (para $\alpha = 0,01$. Se trata de un contraste unilateral, acumula α en una sola cola $\rightarrow Z = 2,33$). Como 1'93 (valor del estadístico de contraste) < 2'33 (valor crítico) $\rightarrow$ Se acepta la H_0 .
- A partir del **nivel crítico p** (al valor del estadístico de contraste 1'93 le corresponde en las tablas una probabilidad de p ($Z < 1,93$) = 0,9732 $\rightarrow 1-0'9732 = 0'0268$). Al tratarse de un contraste unilateral $p = 0'0268 > \alpha = 0'01$; Por tanto, Se acepta la H_0

Interpretación de los resultados: Para un nivel de significación $\alpha = 0,01$, no podemos afirmar que la media poblacional se ha incrementado en la actualidad.

¿Cuál es la probabilidad de cometer error tipo II? (Rechazar la H_1 siendo cierta)

Bajo la Hipótesis nula ($H_0: \mu_0 = 55$) se determina el límite entre la región de aceptación y rechazo (Contraste unilateral derecho)

$$Z = (\bar{Y}_c - \mu_0) / \sigma_Y \rightarrow Z = 2,33 \qquad \text{-----/-----}$$
$$Z = (Y_c - 55) / (11 / \sqrt{50}); \quad Y_c = 58,62$$

Se busca la (Z) que corresponde al valor 58,62 en la distribución de la $H_1 \rightarrow (H_1: \mu_1 = 60)$

$$Z = (\bar{Y}_c - \mu_1) / \sigma_Y \qquad \text{-----/-----}$$
$$Z = (58,62 - 60) / (11 / \sqrt{50}); \quad Z = (-0,89)$$

Probabilidad de obtener una $Z \leq (-0,89) \rightarrow Z(-0,89) = 0,1867 \rightarrow \beta = 0,1867$ (error tipo II)

Potencia de Contraste $\rightarrow (1 - \beta) = 0,8133$