2009

UNED

[TEMA 3]
Análisis de datos para diseños de dos grupos. Muestras independientes.

ÍNDICE

3.1. – INTRODUCCIÓN.

En temas anteriores hemos visto los contrastes de hipótesis más habituales para diseños de una sola muestra cuya utilidad es indiscutible, pero generalmente será conveniente, y en algunos casos necesario, utilizar más de una muestra.

Veamos un ejemplo típico en el que utilizamos un diseño de una muestra. Un profesor opina que aplicando un nuevo método de enseñanza, podría lograr que sus estudiantes comprendieran mejor su asignatura, lo que se traduciría en un incremento de la nota media a final del curso. Sabe que la nota media y la varianza de las calificaciones de años anteriores son: $\mu = 5$, $\sigma^2 = 4$, y considera que estos son los datos de la población. Aplica el nuevo método de enseñanza a una muestra de 36 sujetos obteniendo una nota media: $\bar{Y} = 7$. Se pregunta si los datos de la muestra son compatibles con los datos de cursos pasados. Realiza un contraste de hipótesis y como el estadístico de contraste es significativo ($T = 6$, $p < 0{,}001$) concluye que el nuevo método de enseñanza es más eficaz que el tradicional.

Ahora bien, en los resultados que ha encontrado el profesor pueden influir otros factores. Por ejemplo, si ha sido él quien aplicó el nuevo método de enseñanza, es posible que haya trabajado con más entusiasmo que en cursos anteriores, o bien los alumnos del último curso accedieron a su asignatura mejor preparados, o quizás los estudiantes han tenido menos trabajo en otras asignaturas y han podido dedicar más tiempo a la suya. Si hubiese formado dos grupos las conclusiones serían más claras al estar menos sujetas a explicaciones rivales. Para uno de ellos el método de enseñanza sería el tradicional (grupo de control), mientras que el otro grupo (experimental) estaría formado por los sujetos que aprendieron la asignatura con el nuevo método. Finalmente, nuestro profesor realizaría un contraste de hipótesis para comprobar si existen diferencias entre el grupo experimental y el grupo de control.

También es muy común en psicología el diseño de dos muestras cuando queremos comprobar la eficacia de un tratamiento. En este caso medimos la variable dependiente (por ejemplo, ansiedad) antes y después del tratamiento y comparamos ambas medias para comprobar si la terapia ha sido eficaz.

En otras ocasiones el mismo problema que queremos investigar nos obliga a utilizar dos muestras, porque queremos estudiar diferencias entre dos poblaciones diferentes, como puede ser entre hombres y mujeres, entre ambiente rural y urbano, entre dos clases sociales, etc.

Dividiremos el estudio del análisis de datos en diseños de dos muestras en dos temas. En el Tema 3 trabajaremos con muestras independientes y en Tema 4 con muestras relacionadas. Comenzaremos el presente tema aprendiendo a distinguir entre los dos tipos de muestras (punto 3.3), para ver a continuación como llevar a cabo contrastes de hipótesis paramétricos para dos medias en muestras independientes (punto 3.4).

A continuación trataremos el contrastes de hipótesis para dos varianzas (punto 3.5) y dos proporciones (punto 3.6) en muestras independientes. Finalizaremos el tema con un apartado sobre la magnitud del efecto (punto 3.7), que cada vez está cobrando más importancia en los informes de investigación, para finalizar con un resumen del tema (punto 3.8) y con ejercicios de autocomprobación (punto 3.9).

Los dos temas que veremos a continuación, no deberían plantearnos muchos problemas si hemos asimilado correctamente los anteriores, puesto que matemáticamente los conceptos son los mismos y seguiremos el mismo esquema que en el Tema 2, fijándonos prácticamente los mismos objetivos aplicándolos al estudio de dos muestras, es decir:

3.2 OBJETIVOS DEL TEMA.

- ✓ Distinguir entre muestras independientes y relacionadas.
- ✓ Plantear las hipótesis en función de los objetivos de la investigación.
- ✓ Distinguir entre contraste unilateral y bilateral.
- ✓ Seleccionar el estadístico de contraste más adecuado a las hipótesis planteadas.
- ✓ Conocer la distribución muestral del estadístico seleccionado.
- ✓ Realizar los cálculos oportunos para someter a contrastación empírica las hipótesis planteadas.
- ✓ Relacionar el intervalo de confianza con el estadístico de contraste.
- ✓ Interpretar el nivel p-crítico.
- ✓ Determinar e interpretar el o los valores críticos de la distribución muestral.
- ✓ Tomar una decisión respecto a las hipótesis planteadas.
- ✓ Conocer, comprender e interpretar la magnitud del efecto.

3. 3. MUESTRAS INDEPENDIENTES O RELACIONADAS.

En todas las técnicas estadísticas que vemos en este curso **suponemos que las observaciones dentro de una muestra son independientes**, es decir, que no existe relación entre ellas. Por lo tanto, dentro de un grupo el valor de una determinada puntuación no nos informa en absoluto del valor de otras puntuaciones dentro del mismo grupo. Veamos un par de ejemplos. Un psicobiólogo dispone de 10 ratas para realizar un experimento. Quiere formar dos grupos, que tras ser sometidos a diferentes niveles de estrés correrán el mismo laberinto. Todas las ratas están en la misma jaula. Las 5 primeras que coge forman el primer grupo y las 5 restantes el segundo. Probablemente la primera rata que atrapa el investigador es la más "despistada", la siguiente es un poco menos despistada y así sucesivamente, de forma que el primer grupo está formado por las ratas más torpes, e independientemente del tratamiento la media del primer grupo es superior (tardan más en correr el laberinto). En este caso las puntuaciones dentro de cada grupo están relacionadas y sabiendo en qué orden fue capturada una rata puedo predecir su nivel de "torpeza". Veamos otro ejemplo. Me voy 20 días a Río de Janeiro de vacaciones en la mejor época posible, y el tiempo empeora progresivamente. Calculo la temperatura media de esos días y concluyo que la temperatura en Río es muy fría. Los resultados son significativos. Pero si he tenido la mala suerte de que mi viaje coincida con que la peor borrasca del siglo haya comenzado justo al aterrizar en Río, quizás observe que la temperatura ha descendido día tras día, de forma que conociendo la temperatura de un día cualquiera de mis vacaciones, puedo predecir que la del día siguiente será inferior. La conclusión a la que he llegado es errónea porque los datos que he tomado no son independientes. En el Tema 5, veremos cómo realizar un contraste de hipótesis para comprobar la independencia de las observaciones dentro de una muestra. En cualquier caso, para garantizar la independencia de los datos dentro de un grupo, lo mejor que puede hacerse es seleccionar los elementos de la muestra de forma aleatoria.

Por otro lado, cuando trabajamos **con dos muestras (o más de dos), las muestras pueden ser independientes o relacionadas**. Son independientes cuando no existe relación entre los sujetos de una y otra, lo que podremos garantizar si los sujetos son asignados aleatoriamente a cada una de las muestras.

Tenemos **muestras relacionadas** cuando cada observación en una muestra tiene su pareja en la otra. El caso más evidente es cuando son los **mismos sujetos** los que pasan por diferentes condiciones experimentales. Como comentábamos anteriormente, si queremos probar la eficacia de una terapia contra la ansiedad, podemos seleccionar a un grupo de sujetos en los que medimos su nivel de ansiedad antes del experimento; aplicamos la terapia, y volvemos a medirlo después de la terapia para comparar las medias en ansiedad antes y después. En otras ocasiones no son los mismos sujetos los que se repiten en las muestras, pero hay una **relación sujeto a sujeto** en ambas. Por ejemplo, si disponemos de 10 parejas de hermanos gemelos, podemos formar dos grupos de 5 personas donde cada dos hermanos son asignados, aleatoriamente, a grupos distintos. También podemos contar con padres e hijos, maridos y mujeres, etc. Por último, también podemos utilizar pares de **sujetos que están equiparados en variables que pueden influir en el diseño de la investigación**. Por ejemplo, supongamos que para probar la eficacia de dos métodos de enseñanza, queremos controlar la influencia del cociente intelectual, por lo que tomamos pares de sujetos con un CI semejante formando cada uno de ellos parte de muestras diferentes.

Recordamos que las ventajas y problemas de los diseños de dos grupos son tratados en la asignatura Fundamentos de Investigación (Tema 5).

3.4 CONTRASTES DE HIPÓTESIS SOBRE DOS MEDIAS EN MUESTRAS INDEPENDIENTES.

En este apartado veremos tres contrastes de hipótesis sobre dos medias para muestras independientes en función de los supuestos que hagamos sobre las varianzas poblacionales. Comenzaremos por suponerlas conocidas, pero si no es así, que es lo más habitual, podemos suponer que son iguales o diferentes. Utilizaremos un ejemplo para ilustrar el procedimiento estadístico en cada uno de los casos, y para no repetir innecesariamente las mismas fórmulas, cambiaremos a lo largo de los ejemplos la hipótesis alternativa que, como sabemos, puede ser bilateral, unilateral derecha o unilateral izquierda.

En todo contraste de hipótesis el proceso de inferencia estadística se realiza sobre una distribución teórica que denominamos distribución muestral. Comenzaremos con un ejemplo, con fines obviamente didácticos, cuyo objetivo es, simplemente, que el alumno comprenda cómo se compone la distribución muestral en el caso de dos medias con muestras independientes.

3.4.1 Distribución muestral de la diferencia de medias para dos muestras independientes.

Supongamos que tenemos dos poblaciones, y que cada una de ellas, para que el ejemplo sea lo más corto posible, contiene 3 observaciones. Denotaremos las puntuaciones mediante la letra latina Y. Presentamos las puntuaciones, media y varianza de dichas poblaciones:

Población 1. $\{Y_{11} = 2; Y_{12} = 5; Y_{13} = 8\}$. $\mu_1 = 5$, $\sigma_1^2 = 6$.
Población 2. $\{Y_{21} = 3{,}5; Y_{22} = 5; Y_{23} = 6{,}5\}$. $\mu_2 = 5$, $\sigma_2^2 = 1{,}5$.

Donde el primer subíndice hace referencia a la población a la que pertenecen y el segundo al orden que cada puntuación ocupa en su población. En la Tabla 3.1 se muestra el cálculo de la media aritmética de todas las sub-muestras de tamaño "n = 2" con reposición para la Población 1, y que formarán la distribución muestral de la media para dicha población en muestras de tamaño "n = 2".

Tabla 3.1

Media aritmética de todas las muestras posibles de tamaño n = 2 para la Población 1.

	$Y_{11} = 2$	$Y_{12} = 5$	$Y_{13} = 8$
$Y_{11} = 2$	$\frac{2+2}{2} = 2$	$\frac{2+5}{2} = 3{,}5$	$\frac{2+8}{2} = 5$
$Y_{12} = 5$	$\frac{5+2}{2} = 3{,}5$	$\frac{5+5}{2} = 5$	$\frac{5+8}{2} = 6{,}5$
$Y_{13} = 8$	$\frac{8+2}{2} = 5$	$\frac{8+5}{2} = 6{,}5$	$\frac{8+8}{2} = 8$

La distribución muestral de la media para la Población 1, está compuesta por los 9 valores de la Tabla 3.1, que ordenados son: {2; 3,5; 3,5; 5; 5; 5; 6,5; 6,5; 8}, con media y varianza:

$$\mu_1 = \frac{\sum Y}{n} = \frac{45}{9} = 5 \qquad \sigma_{\bar{y}_1}^2 = \frac{\sum Y^2}{n} - \bar{Y}^2 = \frac{252}{9} - 5^2 = 3$$

Como vimos en el primer tema, estos valores también podríamos calcularlos en función de la media y varianza de la población.

$$\mu_1 = 5; \ \sigma_{\bar{y}_1}^2 = \frac{\sigma^2}{n} = \frac{6}{2} = 3$$

Repitiendo el proceso seguido para la Población 1, obtenemos la distribución muestral para la Población 2, formada por los valores: {3,5; 4,25; 4,25; 5; 5; 5; 5,75; 5,75; 6,5}, con media y varianza, respectivamente: $\mu_2 = 5 \ \ y \ \ \sigma_{\bar{y}_2}^2 = 0{,}75$.

La distribución muestral de las diferencias $\bar{Y}_1 - \bar{Y}_2$, la compondremos al emparejar todas las muestras de la Población 1 con todas las muestras de la Población 2. En la Tabla 3.2 están reflejadas todas las opciones posibles.

Tabla 3.2

Distribución muestral de las diferencias $\bar{Y}_1 - \bar{Y}_2$. Todas las posibles muestras de la Población 1 pueden emparejarse, aleatoriamente, con cualquier muestra de la Población 2.

		Distribución muestral de la Población 1								
Distribución muestral. Población 2		**2**	**3,5**	**3,5**	**5**	**5**	**5**	**6,5**	**6,5**	**8**
	3,5	-1,5	0	0	1,5	1,5	1,5	3	3	4,5
	4,25	-2,25	-0,75	-0,75	0,75	0,75	0,75	2,25	2,25	3,75
	4,25	-2,25	-0,75	-0,75	0,75	0,75	0,75	2,25	2,25	3,75
	5	-3	-1,5	-1,5	0	0	0	1,5	1,5	3
	5	-3	-1,5	-1,5	0	0	0	1,5	1,5	3
	5	-3	-1,5	-1,5	0	0	0	1,5	1,5	3
	5,75	-3,75	-2,25	-2,25	-0,75	-0,75	-0,75	0,75	0,75	2,25
	5,75	-3,75	-2,25	-2,25	-0,75	-0,75	-0,75	0,75	0,75	2,25
	6,5	-4,5	-3	-3	-1,5	-1,5	-1,5	0	0	1,5

Con los valores de la Tabla 3.2, podemos comprobar que la media y varianza de los datos que contiene son igual a:

$$\mu_1 - \mu_2 = 0, \quad \sigma^2_{\bar{Y}_1 - \bar{Y}_2} = 3{,}75$$

También podríamos haber obtenido la varianza de la distribución muestral mediante la siguiente fórmula:

$$\sigma^2_{\bar{Y}_1 - \bar{Y}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{6}{2} + \frac{1{,}5}{2} = 3{,}75$$

En los contrastes que veremos a continuación vamos a suponer que las poblaciones de las que proceden las muestras que utilizaremos se distribuyen normalmente, o bien que $n_1 \geq 30$ y $n_2 \geq 30$. Esto nos garantiza que las distribuciones muestrales de la media en ambos casos también se distribuyen normalmente, y si esto es así, también se distribuirá normalmente la distribución muestral de las diferencias entre medias.

Observe el lector que la diferencia $\mu_1 - \mu_2$ en el ejemplo anterior es igual a cero, que es lo que generalmente postulará la hipótesis nula, por lo que habitualmente nos limitaremos a este caso.

3.4.2 Varianzas poblacionales conocidas.

Ejemplo 3.1. Un psicólogo escolar utiliza un test de comprensión verbal recientemente traducido del inglés, que proporciona puntuaciones en un nivel de medida de intervalo. Se sabe, por investigaciones anteriores, que las varianzas en la población son para niños y niñas $\sigma_1^2 = 36\ y\ \sigma_2^2 = 49$ respectivamente. Las investigaciones anteriores también indican que la media es la misma en ambos grupos, pero este último aspecto no ha sido comprobado con muestras españolas. El psicólogo considera que la traducción del test no es muy acertada y puede provocar diferencias que en realidad no se deben a la comprensión verbal, por lo que selecciona aleatoriamente una muestra de 100 niños y otra muestra de 200 niñas obteniendo una media igual a 20 para los niños e igual a 17,5 para las niñas. Con un nivel de confianza del 95%. ¿Podemos afirmar que la puntuación media en el test de comprensión verbal es la misma para niños y niñas?

Condiciones y supuestos. Tenemos un diseño de dos muestras independientes (niños y niñas), seleccionadas de dos poblaciones con varianzas conocidas (el psicólogo asume que las varianzas de las poblaciones de niños y niñas son las que reflejan las investigaciones anteriores), donde la variable dependiente (comprensión verbal) proporciona puntuaciones en una escala de intervalo. Aunque no sabemos si las poblaciones se distribuyen normalmente, trabajamos con muestras que son lo suficientemente grandes ($n_1 \geq 30$ y $n_2 \geq 30$). En definitiva se cumplen los siguientes supuestos:

- Variable dependiente con un nivel de medida de intervalo o razón.
- Dos poblaciones que se distribuyen normalmente, o bien $n_1 \geq 30$ y $n_2 \geq 30$.
- Varianzas poblacionales conocidas.

Formular las hipótesis. En este caso el psicólogo piensa que pueden existir diferencias pero no tiene una hipótesis previa sobre la dirección de las mismas, por lo que planteamos un contraste de hipótesis bilateral:

$$\begin{matrix} H_0: \mu_1 - \mu_2 = 0 \\ H_1: \mu_1 - \mu_2 \neq 0 \end{matrix} \quad , \quad \text{o bien:} \quad \begin{matrix} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{matrix}$$

Estadístico de contraste y su distribución muestral. Conocemos las varianzas de las dos poblaciones y trabajamos con muestras grandes, lo que nos permite asumir la normalidad de la distribución muestral de las diferencias entre medias. Siendo el grupo 1 el de niños y el 2 el de niñas, el estadístico de contraste es igual a:

$$Z = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(20 - 17{,}5) - 0}{\sqrt{\frac{36}{100} + \frac{49}{200}}} = 3{,}21$$

Observamos que la fórmula del estadístico de contraste sigue el mismo esquema general visto en el Tema 1, cuantificando la discrepancia entre la diferencia de medias observada entre las dos muestras frente a una diferencia nula planteada en la hipótesis nula medida en unidades de desviación típica. Por tanto, en el numerador tenemos la diferencia entre el valor del estadístico en la muestra $(\bar{Y}_1 - \bar{Y}_2)$ respecto del valor del parámetro que postula la hipótesis nula $(\mu_1 - \mu_2)$.

Habitualmente la hipótesis nula, como en este caso, especificará que no existe diferencia entre las medias poblacionales, por lo que el término $\mu_1 - \mu_2$, es igual a cero. Por este motivo, generalmente calcularemos el estadístico de contraste mediante la siguiente ecuación:

$$Z = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad \text{(Ecuación 3.1)}$$

Podemos calcular el nivel p-crítico en la tabla de curva normal, que como sabemos es la probabilidad de obtener un valor como el observado o más extremo, suponiendo que la hipótesis nula es cierta.

Primero buscamos la probabilidad de encontrar valores superiores a 3,21:

$$z = 3{,}21 \xrightarrow{Tabla\ de\ curva\ normal} p = 0{,}9993 \rightarrow 1 - 0{,}9993 = 0{,}0007$$

Y como el contraste es bilateral multiplicamos por dos el valor obtenido:

$$\textit{Nivel p-crítico} = 2 \times 0{,}0007 = 0{,}0014$$

Establecer la regla de decisión en función del nivel de confianza. El nivel de significación es del 5% y el contraste es bilateral, por lo que los valores críticos que delimitan cuándo mantenemos y cuándo rechazamos la hipótesis nula son las puntuaciones típicas: $z = \pm 1{,}96$. En el Figura 3.1 representamos los datos del problema.

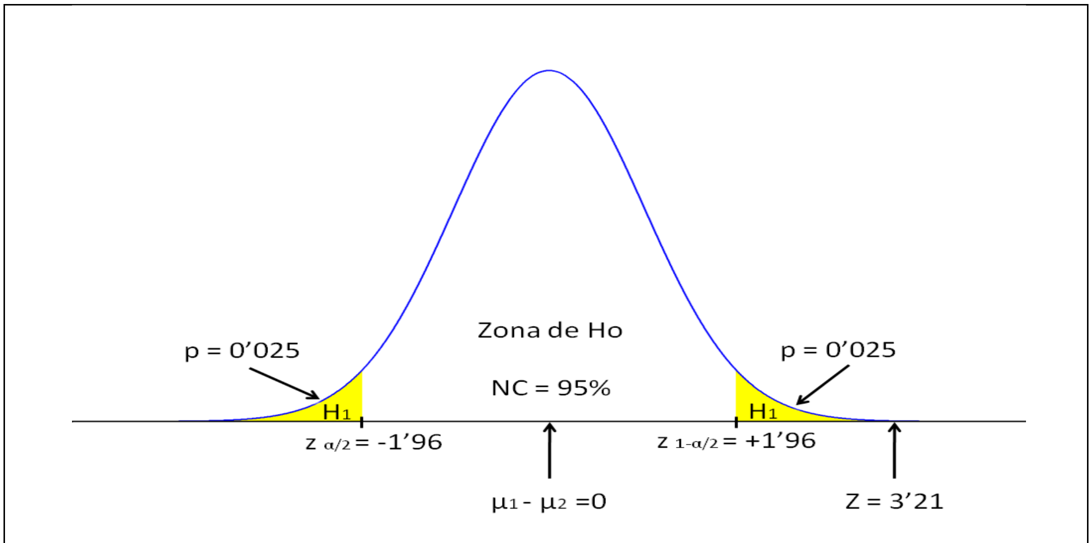


Figura 3.1. *Distribución muestral de las diferencias entre medias según especifica la hipótesis nula. Varianzas poblacionales conocidas.*

Conclusión. Vemos claramente en la Figura 3.1, que el estadístico de contraste (Z = 3,21) no pertenece a la zona de valores compatibles con H_0 que definen las puntuaciones $\pm 1{,}96\ (3{,}21 > 1{,}96)$, por lo que rechazamos la hipótesis nula. En otras palabras, el estadístico de contraste (la discrepancia observada) supera la diferencia que cabría esperar por simple azar. En general, en un contraste bilateral, mantendremos la hipótesis nula cuando el estadístico de contraste no alcance el valor crítico: $z_{\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}$, y la rechazaremos cuando: $Z < z_{\frac{\alpha}{2}}\ o\ bien\ Z > z_{1-\frac{\alpha}{2}}$

Si utilizamos el nivel p-crítico para concluir qué decisión tomar con respecto a H_0, llegamos a la misma conclusión, puesto que 0,0014 < 0,05 (en general, $p < \alpha$). Como se ha expuesto en los temas anteriores, el comparar el nivel crítico con el nivel de significación nos proporciona más información que la comparación del estadístico de contraste con el valor crítico, puesto que vemos claramente que es muy improbable que siendo la hipótesis nula verdadera obtengamos dos muestras cuyas medias tengan una diferencia como la observada. El resultado sería significativo incluso a un nivel de confianza superior al 99%.

Interpretar el resultado en función del contexto de la investigación. Las sospechas del psicólogo parecen fundadas. Las diferencias entre niños y niñas en fluidez verbal son significativas, y pueden deberse a la deficiente traducción del test.

Intervalo de confianza. Si estuviéramos interesados en calcular el intervalo de confianza, lo haríamos mediante la expresión:

$$(\bar{Y}_1 - \bar{Y}_2) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad \text{(Ecuación 3.2)}$$

Que en nuestro caso queda:

$$(20 - 17{,}5) \pm 1{,}96 \sqrt{\frac{36}{100} + \frac{49}{200}} \longrightarrow 2{,}5 \pm 1{,}52 \longrightarrow (0{,}98;\ 4{,}02)$$

Es decir, con un nivel de confianza del 95% la diferencia entre la media de los niños y la media de las niñas en el test de fluidez verbal oscila entre 0,98 y 4,02 puntos a favor de los primeros. Al no contener el valor cero, no podemos admitir la hipótesis nula: $\mu_1 - \mu_2 = 0$.

3.4.3 Varianzas poblacionales desconocidas pero supuestas iguales.

Ejemplo 3.2. En un estudio sobre depresión en personas mayores llevado a cabo en un centro geriátrico, se quiere comprobar si las personas ingresadas que no reciben visitas de sus familiares tienen una puntuación media en depresión superior a aquellas personas cuyos familiares les visitan con frecuencia. Para comprobar esta hipótesis, se seleccionaron aleatoriamente 41 personas que no reciben visitas obteniéndose una puntuación media de 20 puntos con una cuasivarianza igual a 100, mientras que en una muestra aleatoria de 31 personas que si reciben visitas con frecuencia la media fue igual a 15 con una cuasivarianza igual a 90. Suponiendo que las varianzas en la población son iguales para ambos grupos, y con un nivel de confianza del 99% ¿podemos decir que los datos obtenidos avalan la hipótesis de partida?

Condiciones y supuestos. Los requisitos en este caso son iguales que en el caso anterior. La única diferencia es que no conocemos las varianzas poblacionales, si bien las suponemos iguales. Comprobamos pues que se cumplen los siguientes puntos:

- Variable dependiente con un nivel de medida de intervalo o razón. Suponemos que el test de depresión proporciona medidas en una escala de intervalo.
- No sabemos si la distribución en la población es normal, pero salvamos este obstáculo utilizando dos muestras con 30 o más observaciones cada una.
- Varianzas poblacionales desconocidas y supuestas iguales. Veremos posteriormente cómo contrastar diferencias entre dos varianzas. En cualquier caso, la diferencia entre las varianzas de las muestras es pequeña.

Formular las hipótesis. Partimos de la idea de que la depresión media es superior en las personas que no reciben visitas de sus familiares (Grupo 1) respecto de las personas que reciben con frecuencia visitas de sus familiares (Grupo 2), por lo que realizamos un contraste unilateral derecho. Las hipótesis en este caso han de ser:

$$\begin{aligned} H_0&: \mu_1 - \mu_2 \leq 0 \\ H_1&: \mu_1 - \mu_2 > 0 \end{aligned} \qquad \text{o bien,} \qquad \begin{aligned} H_0&: \mu_1 \leq \mu_2 \\ H_1&: \mu_1 > \mu_2 \end{aligned}$$

Estadístico de contraste y su distribución muestral. El estadístico de contraste en este caso se distribuye según t de Student con $n_1 + n_2 - 2$ grados de libertad, y adopta la siguiente expresión:

$$T = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Como comentábamos anteriormente, el término $\mu_1 - \mu_2$, habitualmente es igual a cero, por lo que calcularemos el estadístico de contraste, mediante la siguiente ecuación.

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \text{(Ecuación 3.3)}$$

Con los datos del ejemplo tendremos: $n_1 + n_2 - 2 = 41 + 31 - 2 = 70$ grados de libertad, siendo el estadístico de contraste igual a:

$$T = \frac{20 - 15}{\sqrt{\frac{40 \times 100 + 30 \times 90}{41 + 31 - 2}\left(\frac{1}{41} + \frac{1}{31}\right)}} = \frac{5}{2{,}33} = 2{,}15$$

Establecer la regla de decisión en función del nivel de confianza. Buscamos en las tablas de t de Student el valor crítico, que en este caso es igual a la puntuación que supera al 99% de la distribución para 70 grados de libertad: $t_{70;0{,}99} = 2{,}381$ (véase la Figura 3.2)

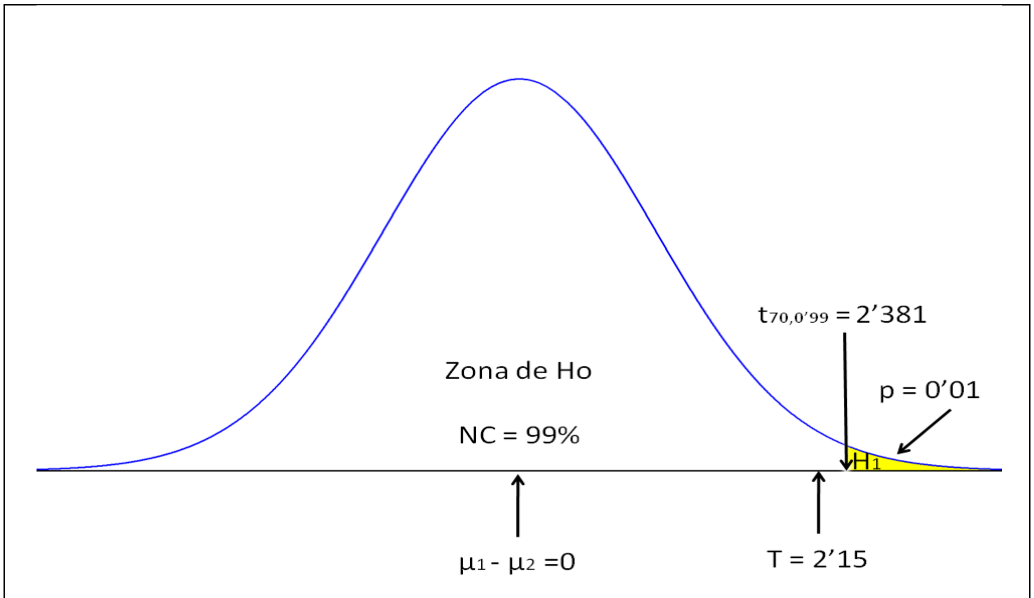


Figura 3.2. *Distribución muestral de las diferencias entre medias según especifica la hipótesis nula. Varianzas poblacionales desconocidas pero supuestas iguales ($\sigma_1^2 = \sigma_2^2$).*

El nivel p-crítico es igual a p = 0,0175. No podemos calcularlo exactamente en las tablas del apéndice, pero podemos utilizarlas para hallar un valor aproximado. Observamos en la tabla t de Student, que para 70 grados de libertad nuestro estadístico de contraste se encuentra entre las puntuaciones 1,994 y 2,381 (1,994 < 2,15 < 2,381) que dejan por encima de si respectivamente las proporciones: 0,025 y 0,01, luego el nivel p-crítico se encontrará entre estos dos valores (0,01 < p < 0,025).

Conclusión. Como podemos apreciar en el Figura 3.2, el valor del estadístico de contraste no supera al valor crítico ($2{,}15 < 2{,}381$) por lo que la diferencia encontrada no es significativa con un nivel de confianza del 99%. En general, y como en situaciones anteriores, en un contraste unilateral derecho mantendremos la hipótesis nula cuando el estadístico de contraste no supere el valor crítico, es decir, si $T < t_{n_1+n_2-2,1-\alpha}$, y la rechazaremos en caso contrario, es decir, cuando $T > t_{n_1+n_2-2,1-\alpha}$. Si comparamos el nivel p-crítico con el nivel de significación, llegamos a la misma conclusión (0,0175 > 0,01).

Interpretar el resultado en función del contexto de la investigación. Al nivel de confianza del 99% los resultados no indican que la puntuación media en depresión es mayor en el grupo de sujetos que no reciben visitas respecto de los que sí las reciben. Pero los resultados sí son significativos al nivel de confianza del 95%, como apreciamos al comparar el nivel de significación con el nivel crítico. Quizás fuera conveniente profundizar en la relación entre ser visitado o no por los familiares y puntuar más alto en depresión en las personas que permanecen ingresadas en centros geriátricos.

Intervalo de confianza. Utilizamos para su cálculo la expresión que puede verse en la Ecuación 3.4:

$$(\bar{Y}_1 - \bar{Y}_2) \pm t_{n_1+n_2-2,\frac{\alpha}{2}} \sqrt{\frac{(n_1-1)\hat{S}_1^2 + (n_2-1)\hat{S}_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \qquad \text{(Ecuación 3.4)}$$

Que en nuestro caso queda:

$$(20-15) \pm 2{,}648\sqrt{\frac{40\times 100 + 30 \times 90}{41+31-2}\left(\frac{1}{41}+\frac{1}{31}\right)} \longrightarrow 5 \pm 6{,}16 \longrightarrow (-1{,}16;\ 11{,}16)$$

Observamos que el intervalo de confianza contiene el valor cero, luego al nivel de confianza del 99% asumimos que las diferencias entre las medias en la población pueden tomar este valor, y por lo tanto no podemos rechazar la hipótesis nula.

3.4.4 Varianzas poblacionales desconocidas y supuestas distintas.

Ejemplo 3.3. Un laboratorio desarrolla un fármaco con el que se pretende reducir la ansiedad. Para comprobarlo, se extrajeron dos muestras aleatorias de cinco observaciones cada una que suponemos procedentes de poblaciones que se distribuyen normalmente con distinta varianza. A los sujetos de la primera muestra se les administró el fármaco y los de la segunda una sustancia placebo. Posteriormente se les midió la ansiedad a todos los sujetos mediante un test en el que cuanto más elevada es la puntuación mayor es la ansiedad. Las resultados de ambas muestras fueron:

Grupo 1 (con fármaco): 10; 20; 30; 20; 5
Grupo 2 (sin fármaco): 30; 50; 30; 60; 20

Con un nivel de confianza del 95%, ¿podemos afirmar que el fármaco efectivamente reduce la ansiedad?

Condiciones y supuestos. Al igual que en los ejemplos anteriores, necesitamos que la variable dependiente esté medida a nivel de intervalo. En cuanto a las poblaciones de las que proceden las varianzas, necesitamos suponerlas normalmente distribuidas porque el tamaño de las muestras es pequeño (con $n_1 \geq 30$ y $n_2 \geq 30$, no es necesario suponer distribuciones normales en ambas poblaciones). En este caso tampoco conocemos las varianzas poblacionales, aunque ahora las suponemos distintas.

Formular las hipótesis. De acuerdo con la hipótesis del laboratorio esperamos que la puntuación media sea inferior en el Grupo 1, por lo que hemos de plantear un contraste de hipótesis unilateral izquierdo.

$$H_0: \mu_1 - \mu_2 \geq 0 \quad \text{o bien,} \quad H_0: \mu_1 \geq \mu_2$$
$$H_1: \mu_1 - \mu_2 < 0 \quad\quad\quad\quad H_1: \mu_1 < \mu_2$$

Estadístico de contraste y su distribución muestral. El estadístico de contraste lo calculamos mediante la Ecuación 3.5:

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}}} \qquad \text{(Ecuación 3.5)}$$

Observe el lector que hemos omitido el término $\mu_1 - \mu_2$, porque en este caso, como es habitual, es igual a cero. Omitiremos a partir de este punto dicho término en contrastes posteriores.

El estadístico de contraste sigue una distribución muestral cuyos grados de libertad calculamos mediante la Ecuación 3.6, cuyo resultado se redondea al entero más próximo:

$$gl = \frac{\left(\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}\right)^2}{\frac{\left(\hat{S}_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(\hat{S}_2^2/n_2\right)^2}{n_2 - 1}} \qquad \text{(Ecuación 3.6)}$$

Con los datos del ejemplo 3.3 tenemos que calcular las medias y varianzas insesgadas de ambas muestras ya que no nos son proporcionadas directamente en el enunciado del ejercicio.

Primero calculamos las varianzas de ambos grupos:

$$\bar{Y}_1 = \frac{\sum Y_1}{n_1} = \frac{85}{5} = 17, \quad S_1^2 = \frac{\sum Y_1^2}{n} - (\bar{Y}_1)^2 = \frac{1825}{5} - 17^2 = 76$$

$$\bar{Y}_2 = \frac{\sum Y_2}{n_2} = \frac{190}{5} = 38, \quad S_2^2 = \frac{\sum Y_2^2}{n} - (\bar{Y}_2)^2 = \frac{8300}{5} - 38^2 = 216$$

Las cuasivarianzas o varianzas insesgadas, serán:

$$\hat{S}_1^2 = S_1^2 \times \frac{n}{n-1} = 76 \times \frac{5}{4} = 95$$

$$\hat{S}_2^2 = S_2^2 \times \frac{n}{n-1} = 216 \times \frac{5}{4} = 270$$

Con lo que calculamos el estadístico de contraste y los grados de libertad.

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}}} = \frac{17 - 38}{\sqrt{\frac{95}{5} + \frac{270}{5}}} = -2{,}46$$

$$gl = \frac{\left(\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}\right)^2}{\frac{\left(\hat{S}_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(\hat{S}_2^2/n_2\right)^2}{n_2 - 1}} = \frac{\left(\frac{95}{5} + \frac{270}{5}\right)^2}{\frac{(95/5)^2}{5-1} + \frac{(270/5)^2}{5-1}} = 6{,}50 \approx 6$$

Establecer la regla de decisión en función del nivel de confianza. Buscamos en las tablas t de Student el valor que supera a una proporción igual a 0,05 para 6 grados de libertad, obteniendo un valor igual a: $t_{6;0{,}05} = -1{,}943$. En el Figura 3.3 representamos los datos del ejemplo.

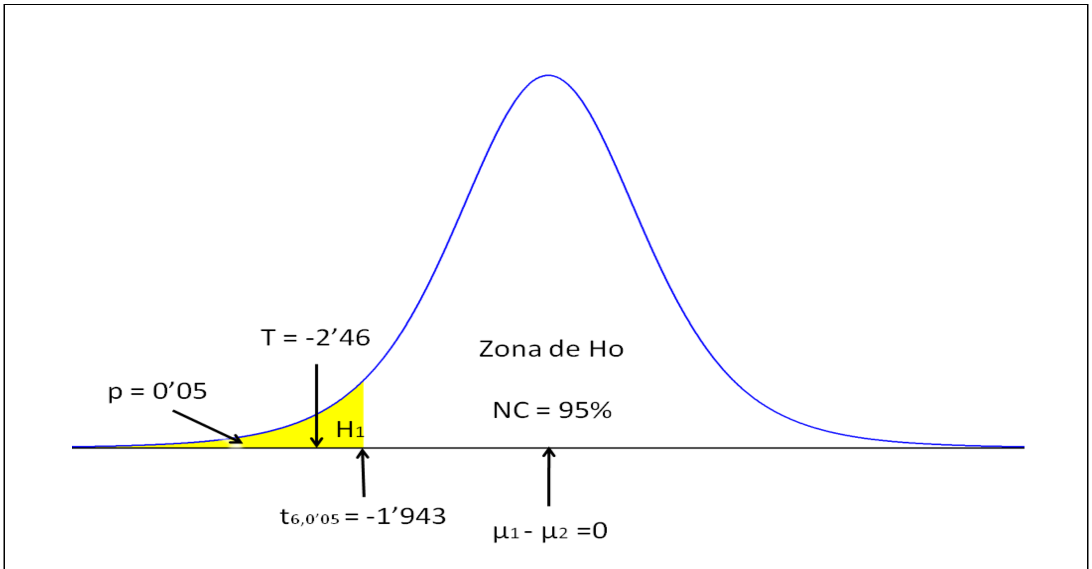


Figura 3.3. *Distribución muestral de las diferencias entre medias según especifica la hipótesis nula. Varianzas poblacionales desconocidas pero supuestas distintas ($\sigma_1^2 \neq \sigma_2^2$).*

Conclusión. El valor del estadístico de contraste es una puntuación más extrema que el valor crítico que hemos buscado en la tabla t de Student ($-2{,}46 < -1{,}943$), por lo que rechazamos la hipótesis nula. Con la misma lógica que en todos los contrastes, en general en un contraste unilateral izquierdo mantendremos la hipótesis nula cuando se cumpla que, $T > t_{g.l.,\alpha}$ y la rechazaremos si $T < t_{g.l.,\alpha}$

En cuanto al nivel p-crítico, en la tabla t de Student, para 6 grados de libertad, tenemos que: $(-3{,}143 < -2{,}46 < -2{,}447)$, por lo que deducimos que el nivel p-crítico estará comprendido entre las probabilidades de encontrar valores iguales o inferiores a estas dos puntuaciones, es decir: $(0{,}01 < p < 0{,}025)$.

Interpretar el resultado en función del contexto de la investigación. A un nivel de confianza del 95% concluimos que la media en ansiedad es inferior para el grupo que tomó el fármaco, por lo que concluimos que reduce la ansiedad.

3.4.5 Consideraciones sobre los contrastes de hipótesis en dos muestras independientes.

En el primer contraste de hipótesis incluíamos en los supuestos que las varianzas poblacionales son conocidas, lo que difícilmente podremos asumir en un caso práctico. Si no conocemos las medias de las poblaciones con las que trabajamos, difícilmente podremos considerar que sí conocemos sus varianzas.

Lo más habitual, por lo tanto, será asumir que las varianzas poblacionales son desconocidas, y en este caso el contraste más utilizado es la prueba T descrita en el punto 3.4.3, en el que suponemos varianzas poblacionales iguales. Este supuesto, al que denominaremos **homocedasticidad**, es muy común en otras técnicas estadísticas, como veremos en los temas en que compararemos las medias de más de dos grupos (Análisis de la Varianza) o en el Análisis de Regresión. De hecho, algunos manuales de estadística tan sólo describen este procedimiento para contrastar diferencias entre dos medias de muestras independientes. La cuestión estriba en que podamos asumir la normalidad de la distribución muestral de las diferencias, lo que podremos garantizar si las muestras que utilizamos son grandes. Si la distribución muestral es normal y los tamaños de ambas muestras son iguales, podemos despreocuparnos de las varianzas poblacionales y suponer sin más que son iguales, sin que por ello peligre la validez del contraste de hipótesis que estamos realizando.

Ahora bien, habrá casos en los que la opción más acertada será suponer varianzas poblacionales distintas, y por lo tanto tendremos que utilizar el estadístico de contraste visto en el apartado 3.4.4. En la literatura científica sobre este tema se proponen diferentes procedimientos para ajustar los grados de libertad de la distribución muestral. No es nuestro objetivo profundizar en esta cuestión, que el lector interesado puede consultar en la bibliografía recomendada en este texto, por lo que nos hemos limitado a describir la solución propuesta por Welch (1938) que posiblemente sea la más utilizada. En definitiva, el procedimiento de Welch nos ofrece un valor inferior para los grados de libertad en relación a si tomamos "$n_1 + n_2 - 2$", el contraste por lo tanto es más conservador, siendo más difícil rechazar la hipótesis nula.

Muchos investigadores sugieren que ha de realizarse previamente un contraste de hipótesis sobre la igualdad las varianzas, de manera que si aceptamos la hipótesis nula ($H_0: \sigma_1^2 = \sigma_2^2$) las supondremos iguales y en caso contrario diferentes. Veremos a continuación cómo llevar a cabo dicho contraste, que tampoco está exento de problemas que también podemos consultar en la bibliografía recomendada.

3.5 CONTRASTE DE HIPÓTESIS SOBRE DOS VARIANZAS EN MUESTRAS INDEPENDIENTES.

Ejemplo 3.4. Según Eysenck (1981), hombres y mujeres tienen la misma puntuación media en cociente intelectual (CI), pero distinta varianza, siendo esta superior para los hombres. Para comprobar la hipótesis de Eysenck, seleccionamos aleatoriamente una muestra de 41 hombres y otra de 31 mujeres. Tras aplicar un test de inteligencia en ambas muestras, observamos que la cuasivarianza en el grupo de hombres es igual a 289, mientras que en el de mujeres vale 225. Con un nivel de confianza del 99% ¿avalan estos datos la hipótesis de Eysenck?

Condiciones y supuestos. Asumimos que las puntuaciones que nos proporcionan los tests de inteligencia miden este constructo en una escala de intervalo, y que la variable medida se distribuye normalmente tanto en la población de hombres como en la de mujeres. En general, los supuestos necesarios son:

- Variable dependiente con un nivel de medida de intervalo o razón.
- Dos poblaciones con variables normalmente distribuidas, o bien $n_1 \geq 30$ y $n_2 \geq 30$.

Formular las hipótesis. Plantearemos un contraste unilateral derecho, en el que la hipótesis alternativa corresponderá a la sugerida por Eysenck, e indicará que la variabilidad en inteligencia es superior en el grupo de hombres. Estadísticamente planteamos las hipótesis de la siguiente forma:

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} \leq 1 \qquad H_1: \frac{\sigma_1^2}{\sigma_2^2} > 1$$

o bien:

$$H_0: \sigma_1^2 \leq \sigma_2^2 \qquad H_1: \sigma_1^2 > \sigma_2^2$$

Elección del estadístico de contraste y su distribución muestral. El estadístico de contraste, sigue una distribución muestral "F" de Fisher, y es calculado según la Ecuación 3.7 :

$$F = \frac{\hat{S}_1^2}{\hat{S}_2^2} = \frac{289}{225} = 1{,}28 \qquad \text{(Ecuación 3.7)}$$

Los grados de libertad del numerador y denominador son, respectivamente: $n_1 - 1 = 41 - 1 = 40$, y $n_2 - 1 = 31 - 1 = 30$.

El cálculo del nivel p-crítico mediante las tablas de las que disponemos generalmente no podrá ser muy aproximado. Observamos en dichas tablas que el primer valor que nos ofrecen para 40 y 30 grados de libertad es igual a 1,573, al que supera una proporción igual a 0,10, luego con las tablas tan sólo podemos saber que el nivel p-crítico es mayor que 0,10 ($p > 0{,}10$). Con un programa informático adecuado concluiríamos que el valor exacto de p es 0,2432.

Establecer la regla de decisión en función del nivel de confianza. A un nivel de confianza del 99% para 40 y 30 grados de libertad, el valor crítico es igual a 2,299. La Figura 3.4 refleja los datos del problema.

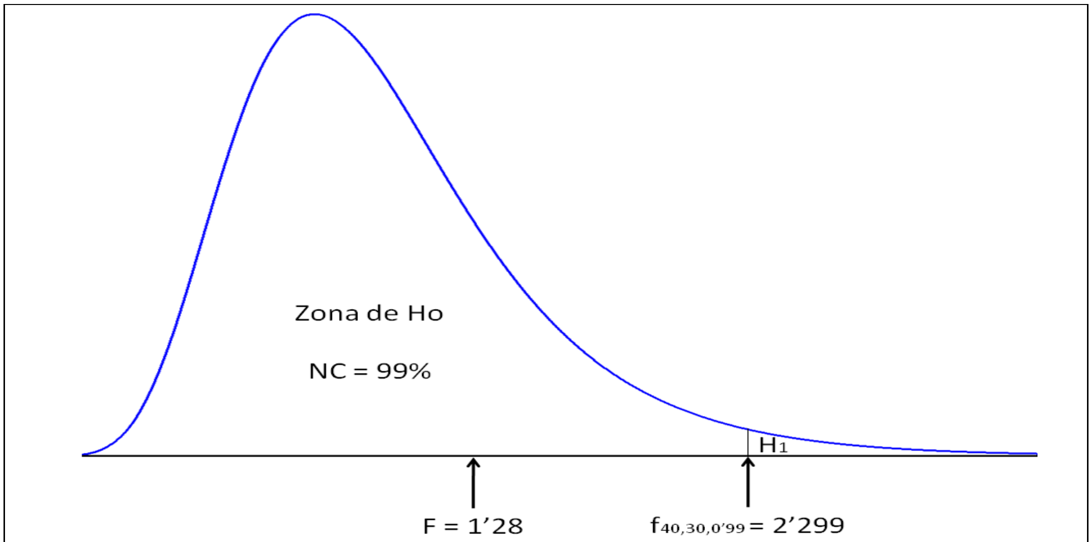


Figura 3.4. *Distribución muestral del cociente de dos varianzas en muestras independientes.*

Conclusión. A la vista de los resultados mantenemos la hipótesis nula a un nivel de confianza del 99%, puesto que el valor del estadístico de contraste es inferior al valor crítico, que en una distribución F de Fisher con 40 y 30 grados de libertad deja por encima a una proporción igual a 0,01. Deducimos del valor del nivel p-crítico, que los resultados obtenidos están lejos de ser significativos para cualquier nivel de confianza razonable. Concluimos por lo tanto que la varianza de hombres y mujeres en inteligencia es la misma.

Interpretar el resultado en función del contexto de la investigación. Eysenck (1981) afirmaba que el hecho de que los hombres mostrasen mayor variabilidad en inteligencia, implica que hay más hombres que mujeres con CI muy altos y con CI muy bajos. Literalmente afirmaba “Esto está de acuerdo con la observación común de que la mayoría de los genios en ciencias, en artes o en otras ocupaciones, así como con defectos mentales, son hombres”. Los datos que manejamos no avalan la hipótesis de Eysenck al no mostrar diferencias significativas en cuanto a la variabilidad en la inteligencia de ambos grupos.

Propiedad recíproca de la distribución F. Aunque en este contraste no ha sido necesario utilizarla, recordamos la propiedad recíproca de la distribución F que vimos en la asignatura Introducción al Análisis de Datos, y que nos sirve para calcular probabilidades que no aparecen en la tabla:

$$f_{n_1-1,n_2-1,\alpha/2} = \frac{1}{f_{n_2-1,n_1-1,1-\alpha/2}}$$

3.6 CONTRASTES DE HIPÓTESIS SOBRE DOS PROPORCIONES EN MUESTRAS INDEPENDIENTES.

Tenemos ahora una variable dependiente dicotómica o dicotomizada, en la que podemos distinguir entre dos sucesos que denominamos “éxito” y “fracaso”, y nos preguntamos si la proporción de éxitos difiere o no en dos poblaciones distintas. Este es un caso muy común en Psicología, donde con frecuencia trabajamos directamente con datos dicotómicos. Por ejemplo: hombre y mujer, acierto o fracaso en una determinada tarea, posición a favor o en contra, recuperación o no de una enfermedad tras aplicar una determinada terapia, etc. Otras veces dicotomizamos una variable que originalmente es continua, como aprobados y suspensos, admitidos y no admitidos en función de las puntuaciones en un test, etc.

Una de las ventajas de trabajar con datos dicotómicos reside en que los supuestos en los que no basamos no son tan fuertes como en el caso de variables continuas. Tan sólo necesitamos que los tamaños de las muestras sean razonables $(n_1 > 30\ y\ n_2 > 30)$ para poder asumir que la distribución muestral de las diferencias de proporciones es normal.

El estadístico de contraste que aplicaremos en este apartado dependerá de cómo planteemos las hipótesis estadísticas. En primer lugar veremos cómo proceder cuando queramos comprobar si la diferencia entre dos proporciones es igual, mayor o menor que cero. Posteriormente trataremos el caso en el que estemos interesados en probar si la diferencia de proporciones es igual, mayor o menor que un determinado valor distinto de cero.

Ejemplo 3.5. En unas determinadas oposiciones se presentaron 200 graduados por la UNED, de los que aprobaron 70, mientras que de 300 candidatos de otras universidades aprobaron 60. Con un nivel de confianza del 95%, ¿podemos afirmar que la proporción de aprobados es la misma entre los graduados de la UNED y los de otras universidades?

Condiciones y supuestos. Tenemos una variable dicotómica y dos muestras grandes que proceden de poblaciones independientes. En general, los supuestos necesarios son:

- Observaciones aleatorias e independientes.
- Variable dependiente dicotómica o dicotomizada.
- Muestras grandes. $n_1 \geq 30\ y\ n_2 \geq 30$.

Formular las hipótesis. Queremos comprobar si la proporción de aprobados es igual o diferente en dos poblaciones distintas, es decir, como no tenemos ninguna hipótesis de partida, planteamos un contraste bilateral.

$$\begin{matrix} H_0: \pi_1 - \pi_2 = 0 \\ H_1: \pi_1 - \pi_2 \neq 0 \end{matrix} \quad o\ bien, \quad \begin{matrix} H_0: \pi_1 = \pi_2 \\ H_1: \pi_1 \neq \pi_2 \end{matrix}$$

Estadístico de contraste y su distribución muestral. Al plantear la hipótesis nula la igualdad de proporciones, el estadístico de contraste que debemos utilizar es el que muestra la Ecuación 3.8:

$$Z = \frac{p_1 - p_2}{\sqrt{P(1-P)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \text{(Ecuación 3.8)}$$

Bajo el supuesto de que la hipótesis nula es verdadera y las dos proporciones poblacionales son iguales, la proporción P es la proporción ponderada de éxitos obtenidos en las dos muestras (véase la Ecuación 3.9), como mejor estimador de la proporción poblacional, π:

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} \qquad \text{(Ecuación 3.9)}$$

Calculamos en primer lugar las proporciones muestrales y el término P:

$$p_1 = \frac{70}{200} = 0{,}35, \qquad p_2 = \frac{60}{300} = 0{,}20, \qquad P = \frac{200 \times 0{,}35 + 300 \times 0{,}20}{200 + 300} = 0{,}26$$

La proporción P, no es más que la proporción total de éxitos, que también podemos calcular como:

$$P = \frac{70 + 60}{200 + 300} = 0{,}26$$

El estadístico de contraste es igual a:

$$Z = \frac{p_1 - p_2}{\sqrt{P(1-P)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0{,}35 - 0{,}20}{\sqrt{0{,}26(1-0{,}26)\left(\frac{1}{200} + \frac{1}{300}\right)}} = 3{,}75$$

Establecer la regla de decisión en función del nivel de confianza. Para un nivel de confianza del 95% y un contraste bilateral, los valores críticos son $z = \pm 1{,}96$. En la Figura 3.5 representamos los datos del problema.

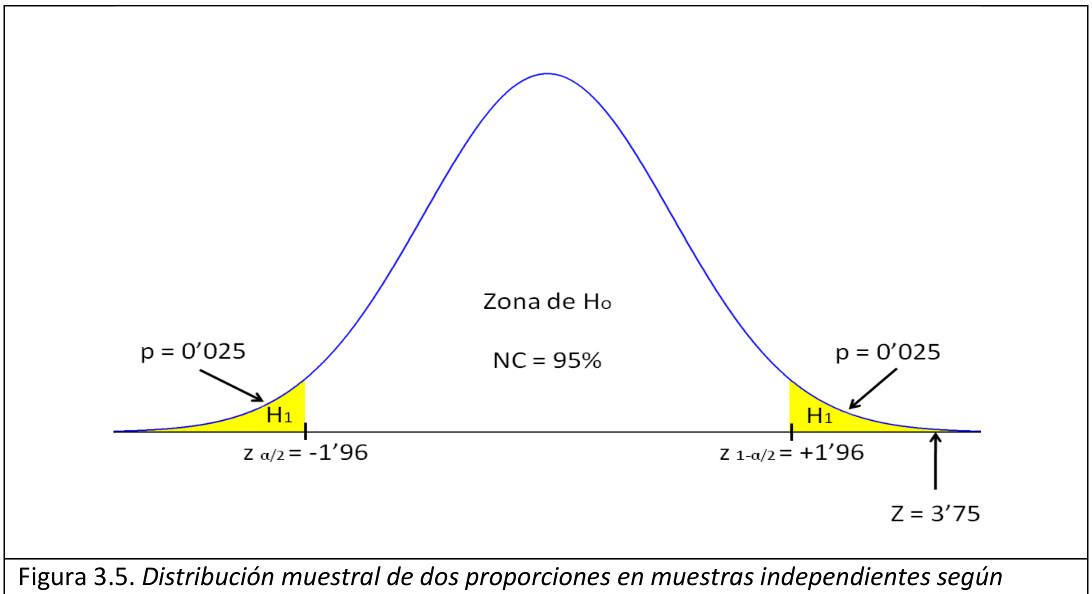


Figura 3.5. *Distribución muestral de dos proporciones en muestras independientes según especifica la hipótesis nula.*

Conclusión. A un nivel de confianza del 95% rechazamos la hipótesis nula, puesto que el estadístico de contraste cae fuera del intervalo que definen los valores críticos $\pm 1,96$, como podemos apreciar en el Figura 3.5. En cuanto al nivel p-crítico, deducimos que es inferior a 0,0002 (que corresponde a la probabilidad de encontrar valores superiores a 3,59, máxima puntuación que aparece en las tablas), por lo que el resultado obtenido supera ampliamente el nivel de confianza fijado por el investigador.

Interpretar el resultado en función del contexto de la investigación. Los resultados muestran que la proporción de opositores que aprueban es diferente entre los graduados de la UNED y otras universidades. Concretamente, la proporción de aprobados es diferente en las dos poblaciones.

Ejemplo 3.6. Según los datos que maneja el director de una academia especializada en preparar a sus alumnos para unas determinadas oposiciones, la proporción de aprobados entre los titulados de la UNED es 0,15 puntos superior a la proporción de aprobados de los titulados procedentes de otras universidades. El director sospecha que el presente curso dicha proporción será superior a 0,15. Para comprobar esta hipótesis extrae una muestra aleatoria de 60 alumnos procedentes de la UNED y otra de 100 alumnos procedentes de otras universidades. Somete a ambas muestras a un examen con el temario de las oposiciones, que es superado por 33 alumnos de la UNED y 30 de otras universidades. Con un nivel de confianza del 95%, podemos afirmar que los datos que maneja el director de la academia son correctos?

Condiciones y supuestos. Tenemos dos muestras independientes con una variable dependiente dicotómica, y estamos interesados en comprobar si la diferencia entre las proporciones poblacionales es superior a 0,15.

Hipótesis. Planteamos un contraste unilateral derecho.

$$H_0: \pi_1 - \pi_2 \leq 0{,}15$$
$$H_1: \pi_1 - \pi_2 > 0{,}15$$

Estadístico de contraste y su distribución muestral. Según la Ecuación 3.10:

$$Z = \frac{(p_1 - p_2) - D}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}} \qquad \text{(Ecuación 3.10)}$$

Con los datos del ejemplo (siendo D el valor propuesto en H_0):

$$Z = \frac{(0{,}55 - 0{,}30) - 0{,}15}{\sqrt{\dfrac{0{,}55(1-0{,}55)}{60} + \dfrac{0{,}30(1-0{,}30)}{100}}} = 1{,}27$$

Establecer la regla de decisión en función del nivel de confianza. Al nivel de confianza del 95%, para un contraste unilateral derecho, el valor crítico es igual a 1,64.

Conclusión. Dado que el estadístico de contraste es menor que el valor crítico, mantenemos la hipótesis nula. Buscando en las tablas de curva normal la probabilidad de encontrar puntuaciones típicas iguales o superiores a 1,27, deducimos que el nivel p-crítico es igual a: $p = 0{,}102$.

Interpretación del resultado en función del contexto de la investigación. A pesar de que la proporción de aprobados entre los alumnos de la UNED supera a la de otras universidades en una proporción igual a 0,25, no podemos afirmar, con los datos que tenemos, que en la población esta proporción será superior a 0,15.

3.7 TAMAÑO DEL EFECTO.

Como ya se indicó previamente, la magnitud o tamaño del efecto es el nombre que se da a una familia de índices que miden el efecto que tiene un tratamiento. Es decir, es un índice que se aplica cuando hay implicados al menos dos grupos, uno de tratamiento y otro de control. Difiere de los contrastes clásicos en que es independiente del tamaño muestral. Este tipo de índices son de uso frecuente en el ámbito del meta-análisis aplicado a la psicología, educación, etc.

Aunque hay una amplia variedad de fórmulas, vamos a ilustrar el proceso básico de obtención del efecto con el denominado índice *"d"*. Básicamente este índice no es más que la estandarización de una diferencia en dos medias, una que sería la del grupo al que se ha aplicado un determinado tratamiento, y otra, la del llamado grupo control. La fórmula del índice viene dada en la Ecuación 3.11.

$$d = \frac{|\bar{Y}_{Tratamiento} - \bar{Y}_{Control}|}{\hat{\sigma}} \qquad \text{(Ecuación 3.11)}$$

Siendo $\hat{\sigma}$ la desviación típica conjunta de ambos grupos, pero cuando sus varianzas son homogéneas. Para el caso en que las varianzas no sean homogéneas, hay otras fórmulas más genéricas en donde se sustituye $\hat{\sigma}$ por el promedio ponderado de las desviaciones típicas insesgadas de ambos grupos (véase la Ecuación 3.12), es decir:

$$\sqrt{\frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{n_1 + n_2 - 2}} \qquad \text{(Ecuación 3.12)}$$

Para interpretar el resultado de este índice, y teniendo en cuenta que es un medida estandarizada, Cohen (1988) propuso una gradación de la magnitud del efecto en "pequeño: d = 0,2", "mediano: d= 0,5" y "grande: d = 0,8 o superior". Veamos un sencillo ejemplo.

Ejemplo 3.7. Se realiza un experimento por el que se trata de estudiar si la verbalización del proceso facilita la realización de tareas manuales complejas. Se seleccionan aleatoriamente 60 sujetos y se asignan 30 a cada uno de dos grupos: el experimental, en el cual los sujetos verbalizan la tarea, y el de control, en el que los sujetos realizan la tarea en silencio. Como variable dependiente se registra el tiempo en segundos que se requiere para completar la tarea, y los valores promedio por grupo son los siguientes:

	$\bar{Y}$	S_{n-1}
Grupo Experimental	**205**	**35**
Grupo Control	**237**	**38**

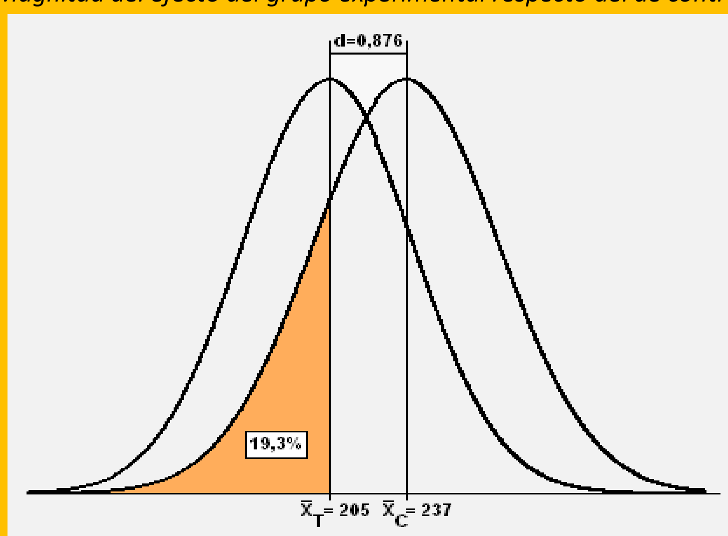
Cuantificar la mejora en rapidez que se produce al verbalizar la tarea.

$$d = \frac{|\bar{Y}_{Tratamiento} - \bar{Y}_{Control}|}{\sqrt{\dfrac{(n_1-1)\hat{S}_1^2 + (n_2-1)\hat{S}_2^2}{n_1+n_2-2}}} = \frac{|205-237|}{\sqrt{\dfrac{29\times 35^2 + 29\times 38^2}{30+30-2}}} = 0{,}876$$

0,876 es la distancia estandarizada entre las medias de los dos grupos, y su probabilidad asociada es 0,807, lo que indica que el 80,7% de los sujetos del grupo experimental tardan menos tiempo que el promedio de los sujetos que no verbalizan. Sólo un 19,3% de los niños que no verbalizan tardan menos tiempo que el promedio de los que sí lo hacen. En la Figura 3.6 se observa la situación del grupo de control respecto del experimental

Figura 3.6

Magnitud del efecto del grupo experimental respecto del de control



3.8 RESUMEN.

En el resumen del Tema 2 veíamos cuáles son los pasos que hemos de seguir en cualquier contraste de hipótesis. Dicho resumen sigue vigente puesto que hemos seguido los mismos pasos para aplicar los contrastes que hemos desarrollado. Además son pocos los conceptos estadísticos nuevos que hemos introducido. Por ejemplo, en el caso de dos medias procedentes de muestras independientes, cambia la manera de obtener la distribución muestral, pero al igual que en el Tema 2, llegamos a una distribución muestral que se distribuye según la curva normal, cuando conocemos las varianzas poblacionales o T de Student, cuando desconocemos las varianzas poblacionales. También cambian las hipótesis postuladas que, lógicamente, ahora plantean diferencias entre dos medias poblacionales.

Gran parte de los conceptos tratados en este tema ya los hemos visto en la asignatura Fundamentos de Investigación, por lo que recomendamos repasar el Tema 5 mediante el texto de dicha asignatura.

Entre los conceptos nuevos destacamos la distinción entre muestras independientes y relacionadas, que es imprescindible para aplicar el análisis de datos adecuado cuando trabajamos con dos muestras. También es nuevo el apartado dedicado a la magnitud del efecto, índice que es independiente del tamaño muestral y que complementa al valor del estadístico de contraste ayudándonos a interpretar la magnitud de las diferencias observadas.

Cada uno de los contrastes tratados ha sido ilustrado a través de un ejemplo, por lo que no hemos visto cómo proceder cuando cambian las hipótesis planteadas. Por ello, a continuación presentamos en las Tablas 3.4 a la 3.7 el resumen de cada una de las pruebas de este tema.

Tabla 3.4. *Contrastes de hipótesis para dos medias en muestras independientes.*			
SUPUESTOS	- Observaciones independientes. - Nivel de medida de intervalo o razón. - Distribuciones normales en la población ó $(n_1 \geq 30 \ y \ n_2 \geq 30)$.		
Varianzas poblacionales	**Conocidas**	**Desconocidas y supuestas iguales**	**Desconocidas y supuestas distintas**
HIPÓTESIS	*Igual en los tres casos. En función de la hipótesis científica plantearemos un contrate:* Bilateral $H_0: \mu_1 - \mu_2 = 0$ $H_1: \mu_1 - \mu_2 \neq 0$	Unilateral izquierdo $H_0: \mu_1 - \mu_2 \geq 0$ $H_1: \mu_1 - \mu_2 < 0$	Unilateral derecho $H_0: \mu_1 - \mu_2 \leq 0$ $H_1: \mu_1 - \mu_2 > 0$
ESTADÍSTICO DE CONTRASTE	$Z = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}}}$
DISTRIBUCIÓN MUESTRAL	Normal tipificada $N(0,1)$	"t" de Student $g.l. = n_1 + n_2 - 2$	"t" de Student $gl = \frac{\left(\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}\right)^2}{\frac{\left(\hat{S}_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(\hat{S}_2^2/n_2\right)^2}{n_2 - 1}}$

REGLA DE DECISIÓN

Distribución normal (varianzas conocidas).			
	bilateral	Unilateral Derecho	Unilateral Izquierdo
Mantener H_0	$z_{\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}$	$Z < z_{1-\alpha}$	$Z > z_{\alpha}$
Rechazar H_0	$Z < z_{\frac{\alpha}{2}}$ ó $Z > z_{1-\frac{\alpha}{2}}$	$Z > z_{1-\alpha}$	$Z < z_{\alpha}$

Distribución "t" de Student (varianzas desconocidas).			
	bilateral	Unilateral Derecho	Unilateral Izquierdo
Mantener H_0	$t_{g.l.,\frac{\alpha}{2}} < T < t_{g.l.,1-\frac{\alpha}{2}}$	$T < t_{g.l.,1-\alpha}$	$T > t_{g.l.,\alpha}$
Rechazar H_0	$T < t_{g.l.,\frac{\alpha}{2}}$ ó $T > t_{g.l.,1-\frac{\alpha}{2}}$	$T > t_{g.l.,\alpha}$	$T < t_{g.l.,\alpha}$

Tabla 3.5. *Contrastes de hipótesis para dos proporciones en muestras independientes.*		
SUPUESTOS	- Observaciones independientes. - Variable dependiente dicotómica o dicotomizada. - $n \geq 30$	
HIPÓTESIS	*En función de la hipótesis científica plantearemos un contrate:* Bilateral: $H_0: \pi_1 - \pi_2 = D$; $H_1: \pi_1 - \pi_2 \neq D$ Unilateral izquierdo: $H_0: \pi_1 - \pi_2 \geq D$; $H_1: \pi_1 - \pi_2 < D$ Unilateral derecho: $H_0: \pi_1 - \pi_2 \leq D$; $H_1: \pi_1 - \pi_2 > D$	
ESTADÍSTICO DE CONTRASTE	$\boldsymbol{D = 0}$ $Z = \frac{p_1 - p_2}{\sqrt{P(1-P)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ Donde: $P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$	$\boldsymbol{D \neq 0}$ $Z = \frac{(p_1 - p_2) - D}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$
DISTRIBUCIÓN MUESTRAL	Normal tipificada. $N(0,1)$	
REGLA DE DECISIÓN	(ver tabla siguiente)	

	bilateral	Unilateral Derecho	Unilateral Izquierdo
Mantener H_0	$z_{\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}$	$Z < z_{1-\alpha}$	$Z > z_{\alpha}$
Rechazar H_0	$Z < z_{\frac{\alpha}{2}}$ ó $Z > z_{1-\frac{\alpha}{2}}$	$Z > z_{1-\alpha}$	$Z < z_{\alpha}$

<table>
<tr><td colspan="2">Tabla 3.6. Contrastes de hipótesis para dos varianzas en muestras independientes.</td></tr>
<tr><td>SUPUESTOS</td><td>- Variable dependiente con un nivel de medida de intervalo o razón.
- Dos poblaciones con variables normalmente distribuidas, o bien $n_1 \geq 30$ y $n_2 \geq 30$.</td></tr>
<tr><td>HIPÓTESIS</td><td>
<table>
<tr><td>Bilateral</td><td>Unilateral izquierdo</td><td>Unilateral derecho</td></tr>
<tr><td>$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$
$H_0: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$</td><td>$H_0: \frac{\sigma_1^2}{\sigma_2^2} \geq 1$
$H_0: \frac{\sigma_1^2}{\sigma_2^2} < 1$</td><td>$H_0: \frac{\sigma_1^2}{\sigma_2^2} \leq 1$
$H_0: \frac{\sigma_1^2}{\sigma_2^2} > 1$</td></tr>
</table>
</td></tr>
<tr><td>ESTADÍSTICO DE CONTRASTE</td><td>$F = \frac{\hat{S}_1^2}{\hat{S}_2^2}$</td></tr>
<tr><td>DISTRIBUCIÓN MUESTRAL</td><td>“f” de Fisher
$\text{g. l. numerador} = \text{n}_1 - 1$
$\text{g. l. numerador} = \text{n}_2 - 1$</td></tr>
<tr><td>REGLA DE DECISIÓN</td><td>
<table>
<tr><td></td><td>Mantener H_0</td><td>Rechazar H_0</td></tr>
<tr><td>Bilateral</td><td>$f_{n_1-1,n_2-1,\alpha/2} < F < f_{n_1-1,n_2-1,1-\alpha/2}$</td><td>$F < f_{n_1-1,n_2-1,\alpha/2}$
O bien:
$F > f_{n_1-1,n_2-1,1-\alpha/2}$</td></tr>
<tr><td>Unilateral Derecho</td><td>$F < f_{n_1-1,n_2-1,1-\alpha}$</td><td>$F > f_{n_1-1,n_2-1,1-\alpha}$</td></tr>
<tr><td>Unilateral Izquierdo</td><td>$F > f_{n_1-1,n_2-1,\alpha}$</td><td>$F < f_{n_1-1,n_2-1,\alpha}$</td></tr>
</table>
</td></tr>
</table>

3.9. EJERCICIOS DE AUTOCOMPROBACIÓN.

1. En un caso práctico en el que estamos interesados en contrastar la diferencia entre las medias de dos poblaciones distintas, lo más habitual es que las varianzas poblacionales sean: a) conocidas; b) desconocidas; c) ninguna de las anteriores es correcta.

2. Si la magnitud del efecto es grande: a) la diferencia entre dos medias es significativa; b) el tamaño de las muestras es grande; c) la diferencia de medias puede ser significativa o no dependiendo del tamaño de las muestras.

3. Si en un contraste unilateral trabajamos con un nivel de confianza del 95%, llamamos a la probabilidad "$1 - 0.95 = 0.05$": a) nivel de significación; b) nivel p-crítico; c) las dos opciones anteriores son verdaderas.

4. En un contraste bilateral sobre dos varianzas al nivel de confianza del 95%, los tamaños muestrales son: $n_1 = 21\ y\ n_2 = 31$, por lo que los valores críticos son: a) 0,426 y 2,195; b) 0,456 y 2,195; c) 0,426 y 2,347.

5. En un contraste de hipótesis unilateral sobre dos proporciones hemos obtenido un estadístico de contrate igual a $Z = 1'80$. ¿Con qué nivel de confianza es significativo? a) 0,95; b) 0,99; c) ninguna de las anteriores.

En un estudio sobre sensibilidad gustativa, un investigador desea comprobar si la sensibilidad gustativa para la fructosa es menor en los fumadores que en los no fumadores, para lo que dispone de dos muestras aleatorias. Tras calcular el umbral absoluto para esta sustancia en ambos grupos, los resultados son:

Fumadores: $\bar{Y}_1 = 2'3,\quad \hat{S}_1^2 = 5'2,\quad n_1 = 34$

No fumadores: $\bar{Y}_2 = 1'25,\quad \hat{S}_2^2 = 4'3,\quad n_2 = 38$

Recordando que a mayor umbral absoluto menor sensibilidad, asumiendo que las varianzas poblacionales son iguales y con un nivel de confianza del 95%, conteste a las siguientes cuestiones:

6. La hipótesis nula es: a) $\mu_1 - \mu_2 \neq 0$; b) $\mu_1 - \mu_2 \leq 0$; c) $\mu_1 - \mu_2 \geq 0$

7. El valor del estadístico de contraste es igual a: a) 1,667; b) 1,994; c) 2,046.

8. La magnitud del efecto es igual a: a) 0,05; b) 0,483; c) 1,05.

9. Utilizando las tablas del texto, el valor p-crítico está comprendido entre los valores: a) 0,025 y 0,05; b) 0,01 y 0,025; c) 0,01 y 0,005.

10. La conclusión de este estudio es: a) mantener la hipótesis nula porque $p < \alpha$; b) rechazar la hipótesis nula porque $p < \alpha$; c) rechazar la hipótesis nula porque el estadístico de contraste es menor que el valor crítico.

3.11. SOLUCIONES A LOS EJERCICIOS DE AUTOCOMPROBACIÓN.

1. La solución es "b". En un caso real difícilmente conoceremos las varianzas poblacionales.

2. La respuesta correcta es "c". Aunque el tamaño del efecto sea grande, si la muestra es pequeña podemos obtener resultados no significativos.

3. la respuesta correcta es "a", el nivel p-crítico es una probabilidad asociada al estadístico de contraste.

4. Tenemos que buscar en la tabla f de Fisher los valores que superan las probabilidades 0,025 y 0,975 con 20 y 30 grados de libertad. En la tabla encontramos:

$$f_{20;30;975} = 2{,}195$$

Y aplicando la propiedad recíproca de la distribución f:

$$f_{20;30;0{,}025} = \frac{1}{f_{30;20;0{,}975}} = \frac{1}{2{,}349} = 0{,}426$$

Luego la respuesta correcta es "a".

5. La solución es "a". El estadístico de contraste supera al valor crítico al nivel de confianza del 95% (1,80 > 1,64), pero no supera el valor crítico correspondiente al nivel de confianza del 99% (1,80 < 2,33).

6. El investigador sospecha que el umbral absoluto es mayor en los fumadores (menor sensibilidad), que será lo que postula la hipótesis alternativa, luego la hipótesis nula es la que especifica la opción "b".

7. Tras realizar los cálculos oportunos comprobamos que la opción correcta es "c"

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{(n_1-1)\hat{S}_1^2 + (n_2-1)\hat{S}_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{2{,}3 - 1{,}25}{\sqrt{\frac{33 \times 5{,}2 + 37 \times 4{,}3}{70}\left(\frac{1}{34}+\frac{1}{38}\right)}} = 2{,}046$$

8. La opción correcta es "b".

$$d = \frac{|\bar{Y}_{Tratamiento} - \bar{Y}_{Control}|}{\sqrt{\frac{(n_1-1)\hat{S}_1^2 + (n_2-1)\hat{S}_2^2}{n_1+n_2-2}}} = \frac{|2{,}3 - 1{,}25|}{\sqrt{\frac{33 \times 5{,}2 + 37 \times 4{,}3}{70}}} = 0{,}483$$

9. Buscando en las tablas t de Student para 70 grados de libertad, observamos que el estadístico de contraste (T = 2,046) está comprendido entre los valores 1,994 y 2,381 que dejan por encima de sí las probabilidades 0,025 y 0,01, respectivamente. Luego la respuesta correcta es "b".

10. La respuesta correcta es "b".