



UNIVERSIDAD DE BUENOS AIRES

ESTADÍSTICA - CÁTEDRA I

CÓDIGO 60

PROF. ADJUNTA A CARGO: DRA. MARÍA SILVIA GALIBERT

TRABAJO PRÁCTICO Nro. 5

AUTORES

Profesores y Docentes de la Cátedra I



“La probabilidad estadística de que las estructuras orgánicas y las reacciones tan precisamente armonizadas que caracterizan a los organismos vivos se generen por accidente es cero. Así es, cero.”

Ilya Prigogine 1917 - 2003 Source: I. Prigogine, N. Gregoir, A. Babbyabtz, *Physics Today* 25, pág. 23-28.

NOTA: A diferencia de las otras prácticas, ésta comienza con el desarrollo de 5 ejercicios modelo dado que, en lo concerniente a Prueba de hipótesis, es fundamental que puedan aprender y seguir al detalle sus pasos y su lógica al momento de resolver. Verán que mientras presentamos las resoluciones, intercalamos algunas precisiones conceptuales que hacen a la comprensión del tema pero que no se espera que aparezcan en las soluciones que Uds. presenten. Para distinguir entre lo que es la solución del ejercicio propiamente dicha y las explicaciones adicionales, hemos optado por remarcar estas últimas en recuadros o utilizando letra cursiva.

PRÁCTICA 5. Prueba de hipótesis

EJERCICIO MODELO I

Un psicopedagogo que trabaja en una escuela secundaria con orientación artística lleva registros históricos de los puntajes de los alumnos en un test de actitudes hacia la matemática. Según sus registros, la media para los estudiantes que completaron 3er año es de 43 puntos con una desviación estándar de 9. Se propone llevar a cabo, en coordinación con los profesores de matemática, un programa de actividades tendientes a mejorar dicha actitud. Pone a prueba su programa con 36 alumnos y al final del período correspondiente les administra el test. Obtiene una media de 46 puntos.

- ¿Puede afirmarse al 5% que el programa respondió a sus objetivos?
- En caso afirmativo evalúe el tamaño del efecto.
- Calcule el nivel crítico de probabilidad o “valor p” e indique qué decisión habría tomado de haber trabajado con un nivel de significación del 1%.
- ¿Cómo mínimo qué valor de la media muestral conduciría a atribuir su diferencia respecto de 43 a que la media poblacional es en realidad mayor que 43 (el programa tuvo efecto) más que al azar; en otras palabras, cuál es el punto crítico a partir del cual la diferencia es estadísticamente significativa al 5%?

RESOLUCIÓN

- a)** Es una prueba de hipótesis sobre una media (μ), siendo conocida la desviación estándar poblacional (σ).

Se sugiere seguir los siguientes pasos:

- 1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimos necesarios (si los hubiere) para llevar a cabo la prueba**

X = Puntaje que obtendrían en el test de actitudes hacia la matemática los estudiantes de la escuela de arte donde trabaja el psicopedagogo, una vez completado el ciclo básico, *si se implementara* el programa. (Nótese que se trata de una población hipotética).

2) Plantear las hipótesis por contrastar

$H_0: \mu = 43$ El efecto del programa es nulo.
 $H_1: \mu > 43$ El programa tiene un efecto positivo.

3) Elegir un nivel de significación

$\alpha = 0,05$

4) Indicar el estadístico de prueba adecuado y explicitar su distribución bajo la hipótesis nula

por el Teorema Central del Límite ya que la muestra es suficientemente grande.

$$\bar{X} \approx N\left(43, \frac{9}{\sqrt{36}}\right) = N(43; 1,5)$$

Bajo H_0

Bajo H_0 el efecto del programa es nulo; por lo que también la desviación estándar de X es la misma que la de los puntajes históricos.

5) Determinar la zona de rechazo de H_0 y establecer la regla de decisión

Zona de rechazo: Siendo un contraste unilateral a derecha, con un nivel de significación de 0,05 y la distribución de la media es normal, la zona de rechazo está dada por todos los valores mayores que el percentil 95 de la distribución normal con media 43 y desviación estándar 1,5; esto es 45,4673, habiendo utilizado la aplicación ... (mencione la aplicación que haya usado; por ejemplo Probability Distribution o EXCEL, como lo hizo en el TP4).

Regla de decisión: Se rechaza H_0 si y sólo si el valor observado de la media es mayor que 45,4673.

6) Calcular el valor observado del estadístico de prueba

En este caso el valor observado de la media es un dato directo del enunciado del problema; en la realidad cuando se trabaja con una base de datos hay que calcular la media, usualmente lo hace el mismo programa.

$\bar{x} = 46$

Obsérvese que en este punto hemos utilizado una letra minúscula para aludir a la media muestral por tratarse de un valor observado a partir de una muestra en particular; mientras que en el punto 4 utilizamos la mayúscula para aludir a la media como una variable aleatoria cuyos valores dependen de la muestra que salga sorteada (es decir, un estadístico). Por eso, trate de diferenciar el concepto de variable (conjunto de valores posibles) del de valor particular que toma esa variable una vez que se tomaron los datos (valor observado). Una notación alternativa de la media observada es \bar{x}_{obs} .

7) Tomar la decisión

Como $46 > 45,4673$ se rechaza H_0 al 5%

8) Dar una conclusión en términos del problema planteado

Se concluye que el programa sería eficaz en el sentido de que los puntajes de actitud hacia la matemática aumentarían en promedio si se implementara el programa.

b) Tamaño del efecto

$$d = \frac{46 - 43}{9} = 0,33, \text{ lo que se interpreta como un efecto más bien pequeño.}$$

Conviene valorar el tamaño del efecto en términos relativos a la información acopiada en torno al problema estudiado. A falta de tal información, Cohen (1988) sugiere el siguiente criterio orientativo: Valores de d entre 0,2 y 0,3 indican un efecto pequeño, alrededor de 0,5 un efecto mediano y mayores que 0,8 un efecto alto.

c) Nivel crítico

$Valor_p = P(\bar{X} > 46 / \mu = 43) = 0,0228 > 0,01$; por lo que no se habría rechazado la hipótesis nula de haberse trabajado con ese nivel de significación.

d) Punto crítico

El punto crítico es $\bar{x}_{crit} = 45,4673$. Por tanto, la diferencia es estadísticamente significativa al 5% para cualquier valor de la media muestral superior a 45,4673.

Nótese que el punto crítico ya había quedado establecido en el paso 5 del punto a, cuando hubo que determinar la zona de rechazo, pues justamente el punto crítico es la frontera que separa "zona de rechazo" de "zona de aceptación" de H_0 .

EJERCICIO MODELO II

El tiempo ejecución de una tarea rutinaria por parte de los empleados de cierta empresa se distribuye normalmente con una media de 50 minutos. Los directivos de la empresa están dispuestos a pagar un programa de entrenamiento para sus empleados si tuvieran evidencia, al 5%, de que éste sería eficaz para mejorar los tiempos de ejecución. Con el fin de tomar la decisión acerca de contratar el servicio de entrenamiento, lo pone a prueba sobre 16 de sus empleados elegidos al azar. Luego del entrenamiento, los tiempos registrados fueron: 44 – 32 – 56 – 20 – 68 – 38 – 50 – 34 – 54 – 44 – 41 – 47 – 30 – 58 – 42 – 46.

- ¿Cuál será la decisión de los directivos?
- Calcule el valor p y formule la decisión en términos del mismo.
- Calcule el tamaño del efecto del entrenamiento dé una valoración del mismo.
- Lea la descripción de la salida computacional para aprender a reconocer los componentes principales de la resolución de la prueba de hipótesis (la información adicional, no esencial se diferencia con letra inclinada).

a) Es una prueba de hipótesis sobre una media (μ), siendo desconocida la desviación estándar poblacional (σ).

1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimo necesarios (si los hubiere) para llevar a cabo la prueba.

X = Tiempo (en minutos) que tardarían en ejecutar la tarea los empleados de la empresa, **si se implementara** el programa de entrenamiento. (Nótese que se trata de una población hipotética).

Supuesto: X se distribuye normalmente.

2) Plantear las hipótesis por contrastar

$H_0: \mu = 50$ El efecto del programa es nulo.

$H_1: \mu < 50$ El programa tiene el efecto esperado disminuyendo el tiempo medio de ejecución.

3) Elegir un nivel de significación

$$\alpha = 0,05$$

4) Indicar el estadístico de prueba adecuado y explicitar su distribución bajo la hipótesis nula

$$t = \frac{\bar{X} - 50}{\frac{S}{\sqrt{16}}} \sim t_{15} \text{ bajo } H_0. \text{ Los grados de libertad corresponden al tamaño de muestra menos 1 (16-1=15).}$$

¿Qué nos llevó a elegir este estadístico? Nótese que en el enunciado no hay información sobre σ , por lo que no se puede utilizar la distribución directa de $\bar{X} \sim N\left(50, \frac{\sigma}{\sqrt{16}}\right)$

Así, es necesario estandarizar a \bar{X} y estimar σ a partir de S , lo que da lugar al estadístico presentando supra cuya distribución es t de Student con 15 grados de libertad (recordemos también que los grados de libertad se calculan como $n-1$)

5) Determinar la zona de rechazo de H_0 y establecer la regla de decisión

Zona de rechazo: Siendo un contraste unilateral a izquierda, con un nivel de significación de 0,05, la zona de rechazo está dada por todos los valores menores que el percentil 5 de la distribución t de Student con 15 grados de libertad; esto es -1,753.

Regla de decisión: Se rechaza H_0 si y sólo si el valor observado del estadístico, t_{obs} , es menor que -1,753.

6) Calcular el valor observado del estadístico de prueba

De los datos se sigue que $\bar{x} = 44$ y la desviación estándar muestral es $s = 11,916$. Sustituyendo se obtiene que:

$$t_{obs} = \frac{44 - 50}{\frac{11,916}{4}} = -2,014$$

7) Tomar la decisión

Como $-2,014 < -1,753$ se rechaza H_0 al 5%.

8) Dar una conclusión en términos del problema planteado

Se concluye que el programa sería eficaz en el sentido de que los tiempos de ejecución de la tarea disminuirían en promedio si se implementara el programa. Por tanto, los directivos de la empresa, a partir de los datos de esta muestra, decidirían implementarlo.

b) Valor $p = P(t_{15} < -2,014 \text{ bajo } H_0) = 0,03115$.

Como $0,03115 < 0,05$ se rechaza H_0 .

c) $d = \left| \frac{44-50}{11,916} \right| = 0,5035$, lo que corresponde a un efecto moderado.

d) Veamos ahora una salida computacional que se obtendría si los datos del problema –en concreto, las observaciones muestrales- se cargaran en una base y se procesaran con un software estadístico. Tomaremos como ejemplo la salida de los programas Statistix y Geogebra.

Statistix 8.0		EJERCICIO MODELO					
One-Sample T Test							
Null Hypothesis: $\mu = 50$							
Alternative Hyp: $\mu < 50$							
Variable	Mean	SE	95% Conf Interval		T	DF	P
			Lower	Upper			
TIEMPO	44.000	2.9791	37.650	50.350	-2.01	15	0.0312

One-Sample T test significa literalmente Prueba T para una sola muestra. Hay una sola muestra porque las hipótesis que se contrastan son sobre la media (μ) de una sola población.

Las siguientes dos líneas aluden claramente a las hipótesis nula y alternativa; en lugar de la notación con letra griega " μ " se utiliza la palabra mu.

Debajo de la palabra Mean (media) se halla el valor observado de la media para la muestra en cuestión.

SE es la abreviatura de Standard Error (error estándar) y es el error estándar de la media muestral. Error estándar de un estimador es otra manera de llamar a su desviación estándar. Entonces SE es la desviación estándar de la media muestral, lo que va en el denominador del estadístico de prueba; es decir, el cociente entre la desviación estándar estimada de la variable ($s=11,916$) y la raíz del tamaño de muestra (4). Note que el valor 2,9791 es el resultado de dividir 11,916 por 4.

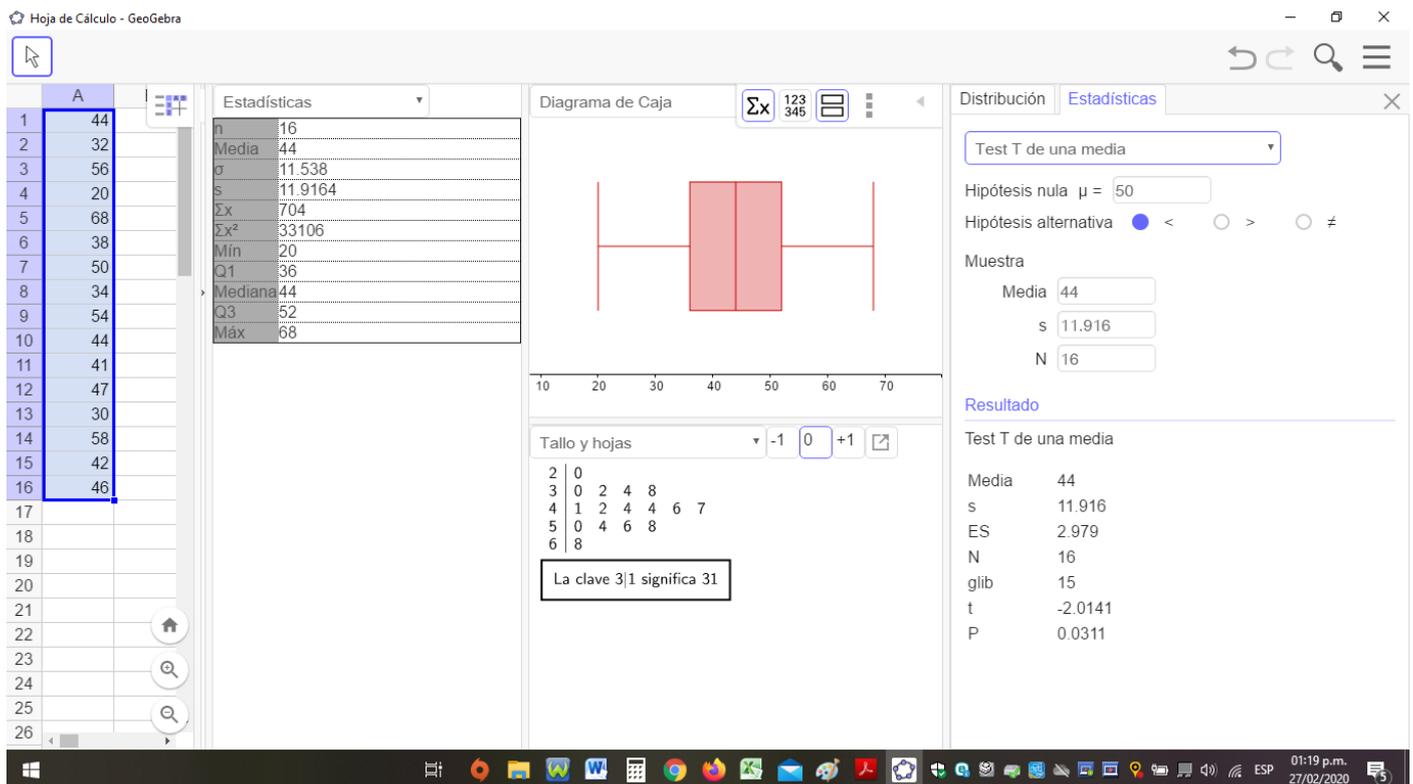
Los valores debajo de Lower y Uper de 95% Conf Interval corresponden a la estimación de la media mediante un intervalo del 95% de confianza. La estimación por intervalos de confianza es otro método de inferencia estadística, diferente al de la prueba de hipótesis, que no forma parte del programa de esta materia.

T corresponde al valor observado del estadístico de prueba. Coincide con lo que en nuestra notación es t_{obs} .

DF es la abreviatura de Degree Freedom = grados de libertad.

P es el valor p (p-value) que se compara con el nivel de significación que el investigador elija para tomar la decisión.

Ahora veamos la salida de GeoGebra:



Como se ve, la salida se organiza en cuatro columnas o pantallas.

Primera: Base de datos.

Segunda: Resúmenes estadísticos (Descriptivo)

Tercera: Diagrama de Caja y Bigotes y diagrama de Tallo-hoja. Es útil que la salida los muestre cuando se realiza el test de hipótesis porque sirven para chequear el supuesto de normalidad de los datos que se necesita para validar el test T.

Cuarta: Es la salida específica del test de hipótesis T sobre una media.

EJERCICIO MODELO III

En una investigación sobre violencia contra la mujer en estudiantes universitarias, Vázquez et al (2010)¹ observaron que 159 jóvenes, entre las 1043 que fueron elegidas al azar para el estudio, experimentaron algún tipo de maltrato en algún momento de sus vidas por parte de su pareja o de alguna persona del sexo opuesto diferente de su pareja. El muestreo fue estratificado por ciclo (1ero., 2do. y 3ero.) y por área de conocimiento (ciencias humanas y sociales, biosanitarias y experimentales) para las mujeres matriculadas en las universidades gallegas de Santiago de Compostela, Coruña y Vigo.

¹El artículo completo sobre el que se basa este ejercicio puede leerse en <http://www.psicothema.com/pdf/3715.pdf>.

¿Puede afirmarse con un nivel de significación del 1% que el porcentaje de jóvenes universitarias de la correspondiente población que ha recibido este tipo de maltrato no llega al

a) 16%?

b) 20%?

Para responder a los interrogantes planteados en a) y b) calcule y utilice el nivel crítico (valor p).

c) Examine la salida computacional correspondiente al ítem a) y reconozca los componentes esenciales de la prueba de hipótesis.

RESOLUCIÓN

a) Se trata de una prueba sobre una proporción poblacional p .

1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimos necesarios (si los hubiere) para llevar a cabo la prueba

X = Condición de haber recibido o no maltrato por parte de su pareja u otra persona del sexo opuesto, en la población de estudiantes universitarias gallegas (Santiago de Compostela, Coruña y Vigo) de los tres primeros ciclos de las carreras de ciencias humanas y sociales, biosanitarias y experimentales.

Supuestos: Las observaciones muestrales son independientes una de otra y con la misma probabilidad de haber recibido maltrato. Es decir que; para cada elección al azar, X es una variable Bernoulli con el mismo parámetro p sobre el cual se desea hacer inferencias.

2) Plantear las hipótesis por contrastar Hipótesis estadísticas

$H_0: p \geq 0,16$ Al menos el 16% de las universitarias en cuestión ha recibido maltrato.

$H_1: p < 0,16$ Menos del 16% de las universitarias en cuestión ha recibido maltrato.

3) Elegir un nivel de significación Nivel de significación

$\alpha = 0,01$

4) Indicar el estadístico de prueba adecuado y explicitar su distribución bajo la hipótesis

Siendo las hipótesis sobre una proporción poblacional, el estadístico es la proporción muestral o frecuencia relativa, que se denota por \hat{p} = Cantidad de éxitos en la muestra / n .

$$\hat{p} \approx N(0,16; \sqrt{\frac{0,16 \cdot 0,84}{1043}}) = N(0,16; 0,0114)$$

por el Teorema Central del Límite ya que la muestra es grande y $n \cdot p_0 = 1043 \cdot 0,16 = 167 > 5$ y $n(1-p_0) = 1043 \cdot 0,84 = 876 > 5$.

Bajo H_0

5) Calcular el valor observado del estadístico de prueba

$$\hat{p} = \frac{159}{1043} \approx 0,1524$$

6) Calcular el nivel crítico o valor p

La prueba es unilateral a izquierda; por lo que el valor p es $P(\hat{p} < 0,1524 \text{ bajo } H_0) = 0,2525$ (habiendo calculado la probabilidad normal con un utilitario).

7) Tomar la decisión

Como $0,2525 > 0,01$ no se rechaza H_0 al 1%.

8) Dar una conclusión en términos del problema planteado

No puede afirmarse al 1% (y ni siquiera al 10%) que el porcentaje de mujeres de la población considerada que ha recibido maltrato por parte de sus parejas u otros varones sea inferior al 16%.

b) Ídem a) con $p_0 = 0,20$

1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimos necesarios (si los hubiere) para llevar a cabo la prueba

Ídem a)

2) Hipótesis estadísticas

$H_0: p \geq 0,20$ Al menos el 20% de las universitarias en cuestión ha recibido maltrato.

$H_1: p < 0,20$ Menos del 20% de las universitarias en cuestión ha recibido maltrato.

3) Elegir un nivel de significación Nivel de significación

Ídem a) $\alpha = 0,01$

4) Indicar el estadístico de prueba adecuado y explicitar su distribución bajo la hipótesis nula

por el Teorema Central del Límite ya que la muestra es grande y $n \cdot p_0 = 1043 \cdot 0,20 = 208,6 > 5$ y $n(1-p_0) = 1043 \cdot 0,80 = 834,4 > 5$.

$$\hat{p} \approx N(0,20; \sqrt{\frac{0,20 \cdot 0,80}{1043}}) = N(0,20; 0,0124)$$

Bajo H_0

5) Calcular el valor observado del estadístico de prueba

$$\hat{p} = \frac{159}{1043} \approx 0,1524$$

6) Calcular el nivel crítico o valor p

La prueba es unilateral a izquierda; por lo que el valor p es $P(\hat{p} < 0,1524 \text{ bajo } H_0) = 0,00006$.

7) Tomar la decisión

Como $0,00006 < 0,01$ se rechaza H_0 al 1%.

8) Dar una conclusión en términos del problema planteado

Hay evidencia al 1% (e inclusive al 1‰) que menos de la quinta parte de las mujeres de la población considerada ha recibido maltrato por parte de sus parejas u otros varones.

c)

Statistix 8.0 05/02/2020,
15:54:47

One-Sample Proportion Test

Sample Size	1043		
Successes	159		
Proportion	0.15244		
Null Hypothesis: P = 0.16			
Alternative Hyp: P < 0.16			
Difference	-0.00756		
Standard Error	0.01113		
Z (uncorrected)	-0.67	P	0.2528
Z (corrected)	-0.62	P	0.2665
95% ConfidenceInterval			
Uncorrected	(0.13063, 0.17426)		
Corrected	(0.13015, 0.17474)		

One-Sample Proportion Test Test de proporción para una muestra

SampleSize	1043	Tamaño muestral: 1043
Successes	159	Cantidad de éxitos (cantidad de unidades en las que se registró la característica de interés; en este caso, cantidad de mujeres que sufrieron violencia por parte del otro sexo) = 159

Proportion	0.15244	Proporción: $159 / 1043 = 0,15244$
------------	---------	------------------------------------

NullHypothesis: P = 0.16 Hipótesis nula y alternativa como se plantearon en el paso 2.
AlternativeHyp: P < 0.16

Difference -0.00756 Diferencia entre el valor observado de la proporción (0,15244) y el valor bajo la hipótesis nula (0,16): $0,15244 - 0,16 = 0,00756$

Standard Error 0.01113 Error estándar (desviación estándar o error típico) de la proporción muestral

Z (uncorrected) -0.67 Puntaje Z asociado a la proporción muestral o proporción muestral estandarizada,

La distribución de la variable proporción muestral estandarizada es aproximadamente normal estándar. No lo usamos en el desarrollo anterior porque utilizamos una aplicación para el cálculo de las probabilidades normales. Los programas usan internamente el valor tipificado para el cálculo de cualquier probabilidad normal; siempre ese cálculo se hace a través de la normal estándar.

P 0.2528 valor p utilizando la distribución normal *sin hacer la corrección por continuidad. La discrepancia con el valor del paso 6 es por la cantidad de decimales considerada en cada caso.*

Z (corrected) -0.62 Idem lo anterior pero con la corrección por continuidad

P 0.2665 Valor p utilizando la distribución normal con corrección por continuidad.

95% ConfidenceInterval

Uncorrected (0.13063, 0.17426)

Corrected (0.13015, 0.17474)

Intervalos del 95% de confianza con y sin la corrección continuidad.

Aunque no forma parte del programa de esta materia es de fácil interpretación. Tomando la primera línea como ejemplo, se puede decir que estimamos la proporción de mujeres que han sufrido este tipo de violencia entre un 13,06% y un 17,43%, estimación acerca de la que hay un 95% de confianza.

La corrección por continuidad tiene lugar cuando la distribución de una variable aleatoria discreta se aproxima por una continua. Esa corrección se hace para mejorar la aproximación. El estadístico "proporción muestral" es una variable discreta porque proviene de un conteo: Cantidad de éxitos, dividido el total de observaciones. Todo el desarrollo precedente supone que esta cantidad de éxitos es binomial; por tanto, al aplicar la distribución normal, lo que se está haciendo es una aproximación de la binomial por la normal. El programa da los resultados con y sin la corrección.

EJERCICIO MODELO IV

Los Trastornos de la Conducta Alimentaria (TCA) han despertado un creciente interés durante los últimos años entre los profesionales de la salud debido al incremento de su incidencia y prevalencia². Los TCA tienen una etiología policausal, interactuando de una forma compleja factores predisponentes, factores desencadenantes y mantenedores (Garner y Garfinkel, 1980; Toro y Vilardell, 1987). Dentro de estos factores se ha postulado la insatisfacción con la propia figura (IC) como uno de los factores claves y se encontraron importantes relaciones entre IC y la instauración de TCA.

²Términos utilizados en Epidemiología: *Prevalencia* se denomina a la proporción de individuos de un grupo o una población que presentan una característica o evento determinado en un momento o en un período determinado. Refiere a todos los individuos afectados. *Incidencia* es una medida del número de casos nuevos de una enfermedad en un período de tiempo determinado.

Un grupo de investigadores³ realiza un estudio seleccionando dos muestras aleatorias e independientes de 29 chicas y 24 chicos de alrededor de 15 años que cursan el 3º año en una escuela privada de Tudela (Navarra). Les administran el Eating Disorders Inventory (EDI), test original de Garner, Olmsted y Polivy (1983), que es un autoinforme para evaluar las características conductuales y cognitivas de la anorexia y la bulimia nerviosa. Consta de 64 ítems agrupados en 8 subescalas. Hay una 2ª versión (EDI-2) con 3 subescalas más quedando con 91 ítems. Una de las subescalas mide IC. En esta escala IC los investigadores obtuvieron los siguientes puntajes:

Chicas	13	11	13	10	16	13	13	5	17	17	21	9	8	19	16	2	11	15	10	25	13	22	13	15
	9	15	6	15	10																			
Chicos	8	10	7	3	0	8	16	7	3	12	6	11	4	5	2	4	10	4	10	8	7	8	13	15

a) Procese los datos del problema mediante el uso de un software e interprete la salida resultante para responder: ¿Es la diferencia observada en la IC entre chicas y chicos estadísticamente significativa al 5%? De ser así, ¿cuál es el tamaño del efecto del sexo sobre la IC?

En el desarrollo de la resolución, explicita los pasos de la prueba de hipótesis.

b) Revise la conclusión a la que llegó en a). ¿Puede decirse, entonces, que el sexo es un factor explicativo en el escenario que estamos considerando? Justifique.

RESOLUCIÓN

Se trata de una prueba de hipótesis sobre una diferencia de medias de muestras independientes provenientes de poblaciones de varianza desconocida.

1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimos necesarios (si los hubiere) para llevar a cabo la prueba

X_1 = Puntaje en la escala IC de adolescentes mujeres escolarizadas de Tudela.

X_2 = Puntaje en la escala IC de adolescentes varones escolarizados de Tudela.

Supuestos:

- $X_1 \sim N(\mu_1, \sigma_1)$ y $X_2 \sim N(\mu_2, \sigma_2)$ Este supuesto de distribución normal de las variables no sería necesario si cada una de las muestras fuera suficientemente grande pero en este caso son muestras pequeñas.

- X_1 y X_2 son variables aleatorias independientes entre sí.

2) Plantear las hipótesis por contrastar

$H_0: \mu_1 = \mu_2$ Mujeres y varones no difieren en promedio en su IC.

$H_1: \mu_1 \neq \mu_2$ Mujeres y varones difieren en promedio en su IC.

³Nota. Este ejercicio, con datos ficticios, está basado sobre la investigación de Baile et al (2003). Artículo completo en http://www.um.es/analesps/v19/v19_2/02-19_2.pdf

Las hipótesis precedentes pueden expresarse de manera totalmente equivalente así:

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

O así:

$$H_0: \mu_2 - \mu_1 = 0$$

$$H_1: \mu_2 - \mu_1 \neq 0$$

3) Elegir un nivel de significación

$$\alpha = 0,05$$

4) Indicar el estadístico de prueba adecuado y explicitar su distribución bajo la hipótesis

El software proporcionará un valor observado del estadístico de prueba y el valor p asociado, el cual calculará teniendo en cuenta la distribución del mismo bajo la hipótesis nula. A continuación se explica el fundamento de los resultados que aparecerán en la salida computacional correspondiente.

Para construir el estadístico de prueba se considera, en principio, la diferencia de medias muestrales en la misma dirección que fueron planteadas las hipótesis. Si H_0 afirma $\mu_1 - \mu_2 = 0$, el estadístico de prueba contiene a $\bar{X}_1 - \bar{X}_2$; si H_0 afirma $\mu_2 - \mu_1 = 0$ el estadístico de prueba contiene la expresión $\bar{X}_2 - \bar{X}_1$.

En cuanto a la distribución de la diferencia de medias muestrales, ésta es normal si las variables lo son o aproximadamente normal si las muestras son suficientemente grandes. Pero se necesita conocer los parámetros bajo la hipótesis nula.

La esperanza de $\bar{X}_1 - \bar{X}_2$ es $\mu_1 - \mu_2$ que, bajo H_0 , es 0. Por tanto la esperanza o media de $\bar{X}_1 - \bar{X}_2$ es 0 bajo H_0 .

Se necesita también el error estándar (o desviación estándar) de $\bar{X}_1 - \bar{X}_2$. De la teoría de probabilidades se deduce que éste es $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$. Pero σ_1 y σ_2 son desconocidos y deben ser estimados con sus correspondientes desviaciones estándar muestrales s_1 y s_2 ; por tanto, lo que tendremos será un error típico de la diferencia de medias estimado: que lo denotaremos $s_{\bar{X}_1 - \bar{X}_2}$. Al no conocer el error estándar verdadero de la diferencia de medias muestrales es necesario estandarizarlas; por tanto el estadístico de prueba tendrá la forma:

*$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}} \sim t_\nu$$
 donde ν denota los grados de libertad que dependen de los supuestos de homogeneidad de varianzas.*

El estadístico de prueba de este test y su distribución varía según se sostenga o no el supuesto de homogeneidad de varianzas; esto es: $\sigma_1^2 = \sigma_2^2$, afirmación que constituye en sí misma una hipótesis estadística que se puede confrontar con la alternativa $\sigma_1^2 \neq \sigma_2^2$. Por tanto, lo correcto es primero realizar esta prueba y, si se rechaza la igualdad de varianzas se procede de una determinada manera y, si no, de otra. No es el objetivo de este curso ahondar en este aspecto sino saber reconocerlo en la lectura de las salidas computacionales correspondientes.

Utilizamos la letra t porque al estimar el error estándar, la distribución, ya no será normal sino t de Student aunque, si las muestras son grandes, puede considerarse aproximadamente normal. La expresión para el error estándar estimado dependerá de los supuestos; será una expresión si se sostienen el supuesto de homogeneidad de varianzas u otra en caso contrario. No nos detendremos en estas fórmulas porque lo hará el programa por nosotros pero tendremos en cuenta estas disquisiciones a la hora de interpretar las salidas que nos da el software.

5) Calcular el valor observado del estadístico de prueba

Dependerá de cómo se haya estimado el error estándar de la diferencia de medias según el supuesto de homogeneidad de varianzas. De todos modos, como ya dijimos, no vamos a aprender a calcularlo a mano, sino que lo leeremos en la salida del software.

6) Calcular el nivel crítico o valor p

La prueba es bilateral; por lo que el valor p es $2P(t > |t_{obs}|)$

Análisis de la salida computacional

```

Statistix 8.0                                difmedias, 06/02/2020,
14:11:09
Diferencia Medias Muestras Independientes

Two-Sample T Tests for PUNTAJEIC by SEXO

SEXO           Mean       N         SD         SE
MUJERES          13.172       29          5.0573       0.9391
VARONES          7.5417      24          4.0858       0.8340
Difference        5.6307

Null Hypothesis: difference = 0
Alternative Hyp: difference <> 0

Assumption      T         DF         P          95% CI for Difference
Lower         Upper
Equal Variances  4.39         51          0.0001       3.0578       8.2037
Unequal Variances 4.48        51.0        0.0000       3.1092       8.1523

Test for Equality      F         DF         P
ofVariances      1.53        28,23       0.1498

```

En la salida precedente hay dos pruebas de hipótesis: sobre la homogeneidad de varianzas y sobre la diferencia de medias.

Primero hay que decidir si hay o no homogeneidad de varianzas. Para ello miramos la parte final de la tabla bajo el título: Test for Equality of Variances. La F corresponde al valor observado del estadístico de prueba que tiene la distribución de Fisher (el tema no forma parte de este curso) con los grados de libertad (DF) correspondientes. Lo que debemos mirar es el valor p. En este caso es $0.1498 > 0.10$; por lo que no se rechazaría la hipótesis nula de igualdad de varianzas utilizando un nivel de significación del 10% (se recomienda este nivel cuando se espera validar la hipótesis nula para aumentar la potencia). Luego, podemos suponer que las varianzas son iguales, entonces, para concluir acerca de la diferencia de medias hay que mirar debajo de donde dice Assumption (supuesto). En virtud de la prueba anterior estamos en la situación de Equal Variances, luego hay que mirar esa línea. El valor observado del estadístico de prueba fue $t_{obs} = 4.39$, la distribución es t de Student con 51 grados de libertad ($29 + 24 - 2$) y el valor p es 0,0001.

7) Tomar la decisión

El valor p es $0,0001 < 0,05$; por tanto, al 5%, se rechaza la hipótesis nula de igualdad de medias.

8) Dar una conclusión en términos del problema planteado

En promedio, difieren las mujeres de los varones en su puntaje en la escala de Insatisfacción Corporal; es decir, que esa diferencia muestral de 5,6307 puntos a favor de las mujeres, es estadísticamente significativa. Pero puede decirse más que esto, ya que la media de las mujeres es mayor que la de los varones, es decir, que el sexo se manifiesta como un factor explicativo de la diferencia en IC.

Para estimar el tamaño del efecto de tal factor, se debería estimar la varianza poblacional (que suponemos igual entre mujeres y varones habida cuenta del resultado del correspondiente test de hipótesis). Entonces la varianza se estima como un promedio ponderado de las dos varianzas muestrales:

$$s_{comb}^2 = \frac{28 \times 5,0573^2 + 23 \times 4,0858^2}{28 + 23} = 21,57$$

Luego, $s = 4,6444$ estima a σ de modo que $d = \frac{\bar{X}_1 - \bar{X}_2}{s_{comb}} = \frac{5,6307}{4,6444} = 1,21$, lo cual es un efecto grande. Como conclusión global podría afirmarse que hay una marcada tendencia a que las adolescentes mujeres experimenten mayor insatisfacción corporal que sus pares varones en la población considerada.

EJERCICIO MODELO V

Considere los datos del ejercicio 4 de la práctica 3. Haga la prueba de hipótesis correspondiente para decidir al 10% si el Estilo de Personalidad es independiente de la funcionalidad familiar o no. Reconozca los elementos del test en la correspondiente salida computacional.

RESOLUCIÓN

Es una prueba de hipótesis de independencia entre dos variables.

1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimos necesarios (si los hubiere) para llevar a cabo la prueba

X= Estilo de personalidad.

Y= Funcionalidad familiar.

2) Plantear las hipótesis por contrastar

H₀: Las variables X e Y son independientes.

H₁: Las variables X e Y no son independientes.

La manera estadística de expresar en símbolos estas hipótesis es:

$$H_0: p_{xy} = p_x \cdot p_y \text{ para todo par de valores } (x,y)$$

siendo p_{xy} la probabilidad conjunta y p_x, p_y las correspondientes marginales.

$$H_1: p_{xy} \neq p_x \cdot p_y \text{ para algún par de valores } (x,y)$$

3) Elegir un nivel de significación

$$\alpha = 0,01$$

4) Indicar el estadístico de prueba adecuado y explicitar su distribución bajo la hipótesis nula

$$\chi^2 = \sum_{i=1}^3 \sum_{j=1}^4 \frac{(f_{ij}^o - f_{ij}^e)^2}{f_{ij}^e} \sim \chi_{2 \times 3}^2 \sim \chi_6^2$$

Para que el estadístico siga esa distribución, se necesita que no más del 20% de las celdas tengan frecuencia esperada mayor o igual a 5. En el problema que nos ocupa, hay 6 celdas con frecuencia esperada menor a 5. Para resolver este problema se agrupan las clases de las variables de una manera conveniente (las más afines). Agruparemos en una clase los estilos de personalidad Cooperativo.

5) Calcular el valor observado del estadístico de prueba

$\chi^2_{obs} = 9,27$ fue calculado en la práctica 3.

6) Calcular el nivel crítico o valor p

Como es un test unilateral a derecha, valor $p = P(\chi^2_6 > 9,27) = 0,15895$

7) Tomar la decisión

Como $0,15895 > 0,10$ se no se rechaza H_0 al 10%.

8) Dar una conclusión en términos del problema planteado

El estilo de personalidad (clasificada en las categorías reagrupadas) es independiente de la funcionalidad familiar.

Lectura de la salida computacional

FAMILIA		PERSONALI				
		INTR-INHIB	COO-SOC-CO	VIOLENTO	RESP-SENS	
FUNCIONAL	Obs	26	8	2	10	46
	Expected	22.56	7.08	7.52	8.85	
	Cell Chi-Sq	0.53	0.12	4.05	0.15	
MOD-FUNC	Obs	54	16	20	16	106
	Expected	51.98	16.31	17.33	20.38	
	Cell Chi-Sq	0.08	0.01	0.41	0.94	
DISFUNCION	Obs	22	8	12	14	56
	Expected	27.46	8.62	9.15	10.77	
	Cell Chi-Sq	1.09	0.04	0.88	0.97	
		102	32	34	40	208
Overall Chi-Square		9.27				
P-Value		0.1589				
Degrees of Freedom		6				
Cases Included		12				Missing Cases 0

La lectura es directa. Se halla la tabla de contingencia con las frecuencias conjuntas observadas, esperadas y el valor ji-cuadrado de la celda, tal como se trabajó en la práctica 3. También están las frecuencias marginales. El valor ji-cuadrado general: 9,27 y su correspondiente valor $p = 0,1589$ que se compara con el nivel de significación elegido para tomar la decisión.

EJERCICIOS PROPUESTOS

1.- En una muestra aleatoria de 25 adultos de la ciudad de Buenos Aires se obtuvo este año una media de 42 en una prueba de memoria espacial. Estudios anteriores afirman que esas puntuaciones, a nivel poblacional, estaban normalmente distribuidas con media 40 y una desviación estándar de 2. ¿Es la media de la muestra de este año significativamente diferente de 40 al nivel del 1%? Suponga que el desvío poblacional actual se mantuvo igual al histórico.-

2.- Un psicometrista está realizando estudios de validación de un test que diseñó para detectar trastorno de atención en niños de entre 8 y 12 años. El test está estandarizado en escala CI. Los puntajes se suponen distribuidos normalmente en la población de interés. Se consideran con atención "normal" a los niños que tienen un puntaje entre 77 y 124; los puntajes inferiores a 77 se toman como indicadores de baja atención y los mayores a 124, como de alta capacidad de atención. El psicometrista administró el test a una muestra aleatoria de 6 niños que asisten a un centro de salud para tratarse por haber recibido el diagnóstico de Trastorno por déficit atencional (TDA).

a) A partir de los resultados que se muestran abajo, realice la prueba de hipótesis correspondiente para decidir al 5% si, en promedio, los niños con TDA que asisten a ese centro de salud puntúan por debajo del punto de corte para pertenecer al grupo de sujetos con baja atención. Puede resolver a mano o bien utilizando algún software estadístico; independiente de la opción que elija, presente los pasos para la resolución de una prueba de hipótesis.

Puntajes: 66 – 80 – 75 – 72 – 79 – 68

b) Conceptualice: ¿Qué puede decir con respecto a la potencia de este test y cómo se relaciona la misma con el tamaño de muestra?

3.- La empresa A entrenó a 80 empleados elegidos al azar en la aplicación de una técnica para hacer uso del sentido del humor como reforzador motivacional. Pasados seis meses, decide evaluar los resultados obtenidos. Se propone mantener el uso de la técnica si más del 75% de los empleados considera que ha sido beneficiosa para motivarlos en su tarea. Así, les administró una encuesta a los 80 participantes que, entre otras cosas, preguntaba: ¿considera que esta técnica lo motiva para realizar su labor diaria? Sesenta y ocho empleados respondieron afirmativamente. A la luz de este resultado ¿aconsejaría usted a la empresa que continúe con la aplicación de la técnica? Utilice una prueba con un nivel de significación 0,10 para responder al interrogante.

4.- Un equipo de estudiosos sostiene que el entrenamiento basado en la resolución creativa de problemas favorece el rendimiento. Se asignaron al azar nueve adolescentes a cada uno de dos grupos. Un grupo fue entrenado en la resolución creativa de problemas y el otro no. Luego se les dio una serie de problemas para resolver. El número de problemas para los cuales cada adolescente presentó una solución posible fue:

Grupo entrenado: 12 16 19 8 10 13 9 15 14

Grupo no entrenado: 15 5 11 8 9 5 6 11 10

Suponga que la "cantidad de problemas resueltos por un adolescente" se distribuye normalmente. Formule las hipótesis convenientes. Contrástelas utilizando un programa computacional y resuelva detallando los distintos pasos. Al tomar la decisión, indique el mínimo nivel de significación entre 0,10, 0,05 o 0,01 al que podría probarse la hipótesis de los estudiosos con los resultados observados.

5.- En tesis de maestría de Guattini, M. (2015) (<http://lildbi.fcm.unc.edu.ar/lildbi/tesis/Guattini-marina-m.pdf>) se reportan la siguiente tabla de contingencia (tabla 17, pág. 52) donde se estudia la relación entre la información sobre la Diabetes Mellitus 2 y el estado nutricional de enfermos de diabetes adscriptos al Programa Nacional de Diabetes y que retiran medicamentos de la farmacia del Hospital Dr. Nanzer de Sto. Tomé, Sta. Fe.

INFORMACIÓN SOBRE DM2	DOMINIO NUTRICIÓN	
	Saludable	No saludable
Conducta no deseada	123	26
Conducta deseada	24	2

Realice una prueba de hipótesis, utilizando un software, para decidir si hay asociación entre la información y el estado nutricional de los pacientes. Indique a qué nivel de significación tomó la decisión. De hallar asociación, dé una medida adecuada de la misma.

6.- Conceptualice

- i) ¿Qué se entiende por hipótesis estadística?
- ii) Dé un ejemplo de una población en la que tenga por objetivo estudiar:
 - a) una variable cuantitativa. Formule acerca de la misma una H_0 y H_1 . Indique qué significaría en ese contexto cometer error de tipo I y cometer error de tipo II.
 - b) dos variables cualitativas cuya asociación desea investigar. Formule las hipótesis nula y alternativa correspondientes. Interprete los dos tipos de errores posibles, siempre en el contexto de su ejemplo.
- iii) Sintetice la lógica del test de hipótesis.
- iv) Defina nivel de significación y cómo incide su valor en la potencia de la prueba.
- v) ¿Por qué el valor p es más informativo que el valor crítico para tomar la decisión?
- vi) ¿Por qué es necesario calcular un tamaño de efecto?
- vii) ¿A qué debe su nombre de “nula” la hipótesis H_0 ?

7.- Ubique en el material de la cátedra el artículo de Novo, M. et al. (2009).

Lea los apartados: Resumen, *Participants* y la primera parte de Results donde alude a Table 1.

- a) En *Participants* se informa $\chi^2 (1, N = 116) = 0.31, ns$. ¿Qué significa esto en la descripción de la muestra?
- b) i) ¿De cuántas pruebas de hipótesis se informan resultados en *Table 1*?
 - ii) ¿Qué tipo de pruebas de hipótesis son? ¿Dónde se informan los grados de libertad dentro de la tabla?
 - iii) Elija tres de ellas, explicita las hipótesis nula y alternativa, y verifique los valores p .
 - iv) ¿Qué frase del resumen alude a los resultados de estas pruebas de hipótesis?

En los siguientes ítems elija la única opción correcta.

8.- Mariano es publicista y quiere demostrarle a su cliente que, a partir de la campaña que ideó y le vendió, logró aumentar la proporción de consumidores de la marca de su cliente. Para ello va a realizar un muestreo y una prueba de hipótesis. ¿Cuál de los siguientes niveles de significación elegiría Mariano para favorecer arribar a la conclusión que desea?

- a. 0,01
- b. 0,05
- c. 0,10

9.- En un estudio sobre el efecto de un entrenamiento en meditación (mindfulness) sobre las habilidades sociales se administró la *Escala de Habilidades Sociales (EHS)*, de Gismero (2000) a un grupo de 25 estudiantes antes (Pre) y después (Post) de dicho entrenamiento.

Parte I

¿Qué test de hipótesis corresponde realizar?

- Un test t de diferencia de medias de muestras independientes.
- Un test t para la media de la diferencia de datos pareados.
- Un test χ^2 de independencia entre el entrenamiento y las habilidades sociales.

Parte II

Los datos se procesaron con Statistix y se obtuvo la salida computacional siguiente:

Paired T Test for Post - Pre

Null Hypothesis: difference = 0

Alternative Hyp: difference > 0

Mean	11.640
Std Error	0.8080
Mean - H0	11.640
Lower 95% CI	9.9723
Upper 95% CI	13.308
T	14.41
DF	24
P	0.0000

Cases Included 25 Missing Cases 0

Del análisis de la salida se concluye que:

- el entrenamiento tuvo efecto estadísticamente significativo mejorando las habilidades sociales.
- hubo un aumento de 11,64 puntos en la media que no resultó estadísticamente significativo.
- el tamaño del efecto del entrenamiento en Mindfulness fue de 14,41 puntos.

10.- García Jiménez, M. et al (2020) estudian la relación entre abandonar un proceso judicial abierto por violencia de género y el contacto de las mujeres que abrieron dichos procesos con sus parejas. Consideraron si las mujeres renunciaron a continuar con el proceso al inicio del mismo o cuando estuviera avanzado y, por otra parte, si el contacto con sus parejas estaba siendo, en ese momento de renuncia, nulo, ocasional o frecuente. Es decir que tomaron en cuenta 2 momentos (inicial o avanzado) y 3 tipos de contacto (nulo, ocasional o frecuente) y generaron así una tabla de contingencia de 2 x 3. El valor observado del estadístico de χ^2 fue 7,19. El valor p fue

- 0,027 y se concluyó al 5% que las variables estaban asociadas.
- 0,007 y se concluyó al 1% que las variables estaban asociadas.
- 0,973 y se concluyó que las variables eran independientes.

11.- El *valor p* para testear si un tratamiento tiene efecto beneficioso fue 0,0025.

¿Cuál de las siguientes interpretaciones es correcta?

- El tratamiento tiene un efecto importante.
- El tratamiento tiene un efecto cuyo tamaño debe ser calculado.
- La diferencia entre recibir o no recibir el tratamiento no es significativa.

12.- En cierto experimento de sensibilidad auditiva los investigadores quieren validar la afirmación de que el 80% de las personas es capaz de reconocer la diferencia entre dos frecuencias de sonido. Para ello realizan la prueba con 64 personas y obtienen que 44 de ellas pueden reconocer la diferencia. Con un nivel de significación del 1% resulta que

- se valida la afirmación, ya que el valor p es mayor que 0,01.
- se concluye que el porcentaje es menor, ya que el valor p = 0,0052.
- se concluye que hay una diferencia significativa al 1%, ya que el valor p = 0,01039.

13.- Las puntuaciones en el Inventario de Burnout de Maslach-Encuesta General (Schaufeli, Leiter, Maslach, & Jackson, 1996) de los transportistas de una ciudad española se distribuyen normalmente con media $\mu = 2,3$ y desviación estándar $\sigma = 1,5$. Se quiso analizar si la puntuación promedio de los transportistas de la ciudad de La Plata superaría a la media de la de los transportistas de la ciudad española en dicha escala. Para ello fueron seleccionados 36 transportistas de la ciudad de La Plata y se registró su puntaje en el mencionado inventario. Se obtuvo una media de 2,8. Suponiendo que la desviación estándar en la población bonaerense es similar a la de la población española, se realizó la prueba de hipótesis adecuada al 5 %. Se concluyó que

- la diferencia observada no es estadísticamente significativa; es atribuible al azar.
- la diferencia de medias poblacionales entre los transportistas de La Plata y los españoles es de 0,5.
- los transportistas de La Plata están más estresados, en promedio, que los españoles.

14.- Se ha realizado un estudio acerca de los factores psicosociales que intervienen en el comportamiento tabáquico en adolescentes españoles. Entre otras variables de personalidad se ha registrado el puntaje de ansiedad rasgo obtenido mediante el cuestionario de Autoevaluación de Ansiedad Estado/Rasgo en niños (STAIC). Se clasificó a 20 jóvenes seleccionados al azar en dos grupos: 10 adolescentes fumadores (Grupo 1) y 10 adolescentes no fumadores (Grupo 2). Se realizó la prueba de hipótesis con el programa Statistix para estudiar si el nivel de ansiedad rasgo difiere significativamente al 10% entre estos dos grupos de adolescentes. Se supone que estos puntajes se distribuyen normalmente. Se obtuvo la siguiente salida:

Variable	Mean	N	SD	SE
Grupo 1	35.200	10	3.0111	0.9522
Grupo 2	34.100	10	2.3781	0.7520
Difference	1.1000			

Null Hypothesis: difference = 0
Alternative Hyp: difference <> 0

Assumption	T	DF	P	95% CI for Difference	
				Lower	Upper
Equal Variances	0.91	18	0.3766	-1.4492	3.6492
Unequal Variances	0.91	17.1	0.3772	-1.4590	3.6590

Test for Equality of Variances	F	DF	P
	1.60	9,9	0.2465

Según la salida precedente se concluye que, **en la población considerada**

- los fumadores son en promedio más ansiosos que los no fumadores.
- fumadores y no fumadores no difieren en promedio en cuanto a su ansiedad.
- el tabaco tiene un efecto moderado elevando los niveles de ansiedad.

REFERENCIAS

García Jiménez, M., Cala, M., Trigo, M. y Barberá, E. (2020). Indicadores de liberación de la violencia de género en la pareja relacionados con el momento de renunciar al procedimiento judicial en España. *Psicothema*, 32, 1, 40-46.

Garner, D. y Garfinkel, P. (1980). Cultural Expectations of Thinness in Woman. *Psychological Reports* 47, 483-491.

Gismero, E. (2000). *EHS Escala de Habilidades Sociales*. Madrid: TEA: Publicaciones de Psicología Aplicada.

Novo, M., Fariña, F., Seijo, D., Vázquez, M. y Arce, R. (2019). Assessing the effects of an education program on mental health problems in separated parents. *Psicothema* 31,3, 284-291.

Toro, J. y Vilardell, E. (1987). *Anorexia Nerviosa*. Barcelona: Martínez Roca.

Vázquez, F., Torres, A., Otero, P., Blanco, V y López, M. (2010). Prevalencia y factores de riesgo de la violencia contra la mujer en estudiantes universitarias españolas. *Psicothema* 22, 2, 196-201.

RESOLUCIÓN

1.- En una muestra aleatoria de 25 adultos de la ciudad de Buenos Aires se obtuvo este año una media de 42 en una prueba de memoria espacial. Estudios anteriores afirman que esas puntuaciones, a nivel poblacional, estaban normalmente distribuidas con media 40 y una desviación estándar de 2. ¿Es la media de la muestra de este año significativamente diferente de 40 al nivel del 1%? Suponga que el desvío poblacional actual se mantuvo igual al histórico.-

Se trata de una Prueba de hipótesis para la media de una población con varianza poblacional conocida.

1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimos necesarios (si los hubiere) para llevar a cabo la prueba

X = Puntaje obtenido en el test que mide memoria espacial por parte de un adulto de la CABA este año

Supuesto: X se distribuye normalmente.

2) Plantear las hipótesis por contrastar

$H_0: \mu = 40$ La media de este año es la misma que la media histórica, es decir, 40.

$H_1: \mu \neq 40$ La media de este año difiere de 40.

3) Elegir un nivel de significación

$\alpha = 0,01$

4) Indicar el estadístico de contraste adecuado y explicitar su distribución bajo la hipótesis nula

$\bar{X} \approx N\left(40, \frac{2}{\sqrt{25}}\right) = N(40; 0,4)$

Como $n < 30$ un supuesto necesario es que la variable se distribuye normalmente

Bajo H_0

5) Determinar la zona de rechazo de H_0 y establecer la regla de decisión

Zona de rechazo: Siendo un contraste bilateral, con un nivel de significación de 0,01 y la distribución de la media es normal, la zona de rechazo está dada por todos los valores pertenecientes a la probabilidad 0,005 en extremo superior y en el inferior de la distribución normal con media 40 y desviación estándar 0,4; esto es **41,03** y **38,97**.

Regla de decisión: Se rechaza H_0 si y sólo si el valor observado de la media es mayor que 41,03 o menor que 38,97.

6) Calcular el valor observado del estadístico de contraste

$\bar{x} = 42$

7) Tomar la decisión

Como $42 > 41,03$ se rechaza H_0 al 1%.

8) Dar una conclusión en términos del problema planteado

La media muestral es significativamente diferente de 40; es decir, que la media poblacional difiere de 40. Dado que se rechazó H_0 por los valores mayores a la media postulada en H_0 , puede concluirse, con más precisión, que la media poblacional de este año es superior a 40.

2.- Un psicometrista está realizando estudios de validación de un test que diseñó para detectar trastorno de atención en niños de entre 8 y 12 años. El test está estandarizado en escala CI. Los puntajes se suponen distribuidos normalmente en la población de interés. Se consideran con atención "normal" a los niños que tienen un puntaje entre 77 y 124; los puntajes inferiores a 77 se toman como indicadores de baja atención y los mayores a 124, como de alta capacidad de atención. El psicometrista administró el test a una muestra aleatoria de 6 niños que asisten a un centro de salud para tratarse por haber recibido el diagnóstico de Trastorno por déficit atencional (TDA).

a) A partir de los resultados que se muestran abajo, realice la prueba de hipótesis correspondiente para decidir al 5% si, en promedio, los niños con TDA que asisten a ese centro de salud puntúan por debajo del punto de corte para pertenecer al grupo de sujetos con baja atención. Puede resolver a mano o bien utilizando algún software estadístico; independiente de la opción que elija, presente los pasos para la resolución de una prueba de hipótesis.

Puntajes: 66 – 80 – 75 – 72 – 79 – 68

Se trata de una Prueba de Hipótesis para la media de una población con varianza desconocida (prueba T para una sola muestra).

1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimos necesarios (si los hubiere) para llevar a cabo la prueba

X = Puntaje en un test para detectar trastorno de atención de un niño que asiste a un centro de salud para recibir tratamiento por TDA

Supuestos: La variable se distribuye normalmente.

2) Plantear las hipótesis por contrastar

$H_0: \mu = 77$ El puntaje medio en el test para detectar TDA de niños que asisten a un centro de salud para recibir tratamiento por TAD es 77. (El test para detectar TDA diseñado por el psicometrista no da evidencias de validez).

$H_1: \mu < 77$ El puntaje medio en el test para detectar TDA de niños que asisten a un centro de salud para recibir tratamiento por TAD es menor a 77. (Probar esta hipótesis iría a favor de la validez del test diseñado por el psicometrista; ya que el mismo funcionaría bien para detectar a los niños con problemas).

3) Elegir un nivel de significación

$$\alpha = 0,05$$

4) Indicar el estadístico de prueba adecuado y explicitar su distribución bajo la hipótesis nula

Dado que no hay información sobre σ es necesario estimar σ a partir de S , resultando:

$$t = \frac{\bar{X} - 77}{\frac{S}{\sqrt{6}}} \sim t_5 \text{ bajo } H_0.$$

→ *Trabajando a mano, la prueba sigue del siguiente modo:*

5) Determinar zona de rechazo y establecer la regla de decisión

Zona de rechazo: Prueba unilateral a la izquierda, con nivel de significación de 0,05, la zona de rechazo está dada por todos los valores menores que el percentil 5 de la distribución t de Student con 5 grados de libertad; esto es: -2,015 [Es valor puede obtenerse usando la aplicación Probability Distributions, opción "t"].

Regla de decisión: Se rechaza H_0 si y sólo si el valor observado del estadístico, t_{obs} es menor que -2,015.

6) Calcular el valor observado del estadístico de prueba

$$t_{obs} = \frac{73,33 - 77}{\frac{5,72}{2,45}} = -1,575$$

7) Tomar la decisión

Como $-1,575 > -2,015$ no se rechaza H_0 al 5%.

8) Dar una conclusión en términos del problema planteado

El puntaje promedio que obtuvieron los niños de la muestra no es significativamente menor a 77.

→ *Si decide trabajar con un software estadístico (aquí mostramos al final de los pasos la salida que se obtiene con el programa Statistix), la prueba sigue del siguiente modo:*

5) Calcular el valor observado del estadístico de prueba

El valor observado del estadístico de prueba que informa la salida es -1,57.

6) Calcular el nivel crítico o valor p

Tal como se lee en la salida, el valor p asociado a -1,57 es 0,0884.

7) Tomar la decisión

Como $p = 0,0884 > 0,05$, no se rechaza H_0 .

8) Dar una conclusión en términos del problema planteado

El puntaje promedio que obtuvieron los niños de la muestra no es significativamente menor a 77.

Salida computacional (Statistix)

Statistix 8.0							
One-Sample T Test							
Null Hypothesis: $\mu = 77$							
Alternative Hyp: $\mu < 77$							
Variable	Mean	SE	95% Conf Interval		T	DF	P
			Lower	Upper			
Puntaje	73.333	2.3333	67.335	79.331	-1.57	5	0.0884
Cases Included 6		Missing Cases 0					

b) Conceptualice: ¿Qué puede decir con respecto a la potencia de este test y cómo se relaciona la misma con el tamaño de muestra?

En el enunciado se dice que psicometrista quiere validar su test de TDA. Por eso lo administra a una muestra de una población de niños con este problema y, si su test midiera bien, debería dar puntajes por debajo del umbral de normalidad (77). Sin embargo no se encuentra con este resultado, por lo que podría concluir que no tiene evidencias de que su test de TDA sea válido.

Pero, al concluir no rechazando H_0 , existe la posibilidad de cometer error de tipo II: no detectar que H_0 es falsa. Lo contrario, la probabilidad de rechazarla cuando es falsa, es la potencia de la prueba de hipótesis. Cuando el tamaño de muestra es pequeño, suele ser baja la potencia de una prueba de hipótesis, su sensibilidad para detectar la falsedad de H_0 . El psicometrista podría concluir que su test de TDA no es válido cuando en realidad tomó la decisión con una prueba de hipótesis de baja potencia. Convendría haber trabajado con una muestra más grande para controlar mejor la probabilidad de cometer error de tipo II y aumentar la potencia de la prueba de hipótesis.

3.-La empresa A entrenó a 80 empleados elegidos al azar en la aplicación de una técnica para hacer uso del sentido del humor como reforzador motivacional. Pasados seis meses, decide evaluar los resultados obtenidos. Se propone mantener el uso de la técnica si más del 75% de los empleados considera que ha sido beneficiosa para motivarlos en su tarea. Así, les administró una encuesta a los 80 participantes que, entre otras cosas, preguntaba: ¿considera que esta técnica lo motiva para realizar su labor diaria? Sesenta y ocho empleados respondieron afirmativamente. A la luz de este resultado ¿aconsejaría usted a la empresa que continúe con la aplicación de la técnica? Utilice una prueba con un nivel de significación 0,10 para responder al interrogante.

Se trata de una Prueba de hipótesis sobre una proporción poblacional

1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimos necesarios (si los hubiere) para llevar a cabo la prueba

X = Condición de haber respondido afirmativamente o no a la pregunta “¿Considera que esta técnica lo motiva para realizar su labor diaria?” por parte de un empleado de la empresa A que participó del entrenamiento para aplicar una técnica que hace uso del sentido del humor como reforzador motivacional

Supuestos: las observaciones muestrales son independientes una de las otras y con la misma probabilidad de responder afirmativamente a la pregunta. Es decir, que para cada elección al azar, X es una variable Bernoulli con el mismo parámetro p sobre el que se desea hacer inferencias.

2) Plantear las hipótesis por contrastar

$H_0: p \leq 0,75$ Como mucho el 75% de los empleados reportaron que la técnica los motiva para realizar su labor diaria.

$H_1: p > 0,75$ Más del 75% de los empleados reportaron que la técnica lo motiva para realizar su labor diaria.

3) Elegir un nivel de significación

$$\alpha = 0,10$$

4) Indicar el estadístico de prueba adecuado y explicitar su distribución bajo la hipótesis nula

por el Teorema Central del Límite ya que la muestra es grande y $n \cdot p_0 = 80 \cdot 0,75 = 60 > 5$ y $n(1 - p_0) = 80 \cdot 0,25 = 20 > 5$.

$$\hat{p} \approx N(0,75; \sqrt{\frac{0,75 \cdot 0,25}{80}}) = N(0,75; 0,0484)$$

Bajo H_0

5) Calcular el valor observado del estadístico de contraste

$$\hat{p} = \frac{68}{80} = 0,85$$

6) Calcular el nivel crítico o valor p

La prueba es unilateral a derecha; por lo que el valor p es $P(\hat{p} > 0,85 \text{ bajo } H_0) = 0,0194$.

7) Tomar la decisión

Como $0,0194 < 0,10$ se decide rechazar H_0 al 10%.

8) Dar una conclusión en términos del problema planteado

Puede afirmarse que más del 75% de los empleados reportaron que la técnica los motiva para realizar su labor diaria. En conclusión, aconsejaría continuar con el uso de la técnica.

4.- Un equipo de estudiosos sostiene que el entrenamiento basado en la resolución creativa de problemas favorece el rendimiento. Se asignaron al azar nueve adolescentes a cada uno de dos grupos. Un grupo fue entrenado en la resolución creativa de problemas y el otro no. Luego se les dio una serie de problemas para resolver. El número de problemas para los cuales cada adolescente presentó una solución posible fue:

Grupo entrenado: 12 16 19 8 10 13 9 15 14

Grupo no entrenado: 15 5 11 8 9 5 6 11 10

Suponga que la "cantidad de problemas resueltos por un adolescente" se distribuye normalmente. Formule las hipótesis convenientes. Contrástelas utilizando un programa computacional y resuelva detallando los distintos pasos. Al tomar la decisión, indique el mínimo nivel de significación entre 0,10, 0,05 o 0,01 al que podría probarse la hipótesis de los estudiosos con los resultados observados.

Se trata de una Prueba de Hipótesis para la diferencia de medias de muestras independientes.

Se nos indica trabajar con un programa computacional. De procesarse los datos del enunciado con el programa Statistix, se obtendría:

Two-Sample T Tests for No Entrenad vs Entrenad						
Variable	Mean	N	SD	SE		
No Entrenad	8.8889	9	3.2956	1.0985		
Entrenad	12.889	9	3.5512	1.1837		
Difference	-4.0000					
Null Hypothesis: difference = 0						
Alternative Hyp: difference < 0						
Assumption	T	DF	P	95% CI for Difference		
				Lower	Upper	
Equal Variances	-2.48	16	0.0124	-7.4235	-0.5765	
Unequal Variances	-2.48	15.9	0.0124	-7.4251	-0.5749	
Test for Equality of Variances	F	DF	P			
	1.16	8.8	0.4189			
Cases Included 18 Missing Cases 0						

1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimos necesarios (si los hubiere) para llevar a cabo la prueba

X1 = Cantidad de problemas para los cuales propondrían una solución los adolescentes si no fueran entrenados en la resolución creativa de problemas.

X2 = Cantidad de problemas para los cuales propondrían una solución los adolescentes si fueran entrenados en la resolución creativa de problemas.

Supuestos:

- $X_1 \sim N(\mu_1, \sigma_1)$ y $X_2 \sim N(\mu_2, \sigma_2)$
- X_1 y X_2 son variables aleatorias e independientes entre sí.

2) Plantear las hipótesis por contrastar

$H_0: \mu_1 = \mu_2$ El entrenamiento no incide en la cantidad de problemas para los cuales se presenta una solución.

$H_1: \mu_1 < \mu_2$ El entrenamiento aumenta la cantidad de problemas para los cuales se presenta una solución.

3) Indicar el estadístico de prueba adecuado y explicitar su distribución bajo la hipótesis nula

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}} \sim t_{16}$$

4) Calcular el valor observado del estadístico de prueba

Dado que, en función del test de homogeneidad de varianzas, no se rechaza la igualdad de varianzas ($p = 0,4189 > 0,01$), se trabaja con la línea Equal Variances.

Así, el valor observado del estadístico de prueba que informa la salida es -2,48

5) Calcular el nivel crítico o valor p

Tal como se lee en la salida, el valor p asociado al estadístico de prueba es 0,0124

6) Tomar la decisión

Como valor $p = 0,0124$ se rechaza H_0 al 5% y aún a niveles menores como, por ejemplo al 2% pero no al 1%.

7) Dar una conclusión en términos del problema planteado

El entrenamiento en la resolución creativa de problema incide favorablemente en la cantidad de problemas a los que un adolescente logra presentar una solución.

5.- En la tesis de maestría de Guattini, M. (2015) (<http://lildbi.fcm.unc.edu.ar/lildbi/tesis/Guattini-marina-m.pdf>) se reportan la siguiente tabla de contingencia (tabla 17, pág. 52) donde se estudia la relación entre la información sobre la Diabetes Mellitus 2 y el estado nutricional de enfermos de diabetes adscriptos al Programa Nacional de Diabetes y que retiran medicamentos de la farmacia del Hospital Dr. Nanzer de Sto. Tomé, Sta. Fe.

INFORMACIÓN SOBRE DM2	DOMINIO NUTRICIÓN	
	Saludable	No saludable
Conducta no deseada	123	26
Conducta deseada	24	2

Realice una prueba de hipótesis, utilizando un software, para decidir si hay asociación entre la información y el estado nutricional de los pacientes. Indique a qué nivel de significación tomó la decisión. De hallar asociación, dé una medida adecuada de la misma.

Es una prueba de hipótesis de independencia entre dos variables cualitativas (Chi-Cuadrado).

1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimos necesarios (si los hubiere) para llevar a cabo la prueba

X = Información sobre diabetes mellitus 2"

Y = Estado nutricional de enfermos de diabetes adscriptos al Programa Nacional de Diabetes y que retiran medicamentos de la farmacia del Hospital Dr. Nanzer de Sto. Tomé, Sta. Fe

Supuestos: las unidades sobre las cuales se hacen las observaciones son tomadas independientemente unas de las otras.

2) Plantear las hipótesis por contrastar

H_0 : Las variables Información sobre DM2 (X) y Dominio Nutrición (Y) son independientes.

H_1 : Las variables Información sobre DM2 (X) y Dominio Nutrición (Y) no son independientes.

3) Indicar el estadístico de prueba adecuado y explicitar su distribución bajo la hipótesis nula

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(f_{ij}^o - f_{ij}^e)^2}{f_{ij}^e} \sim \chi_{1 \times 1}^2 \sim \chi_1^2$$

4) Calcular el valor observado del estadístico de prueba

$$\chi_{obs}^2 = 1,57$$

5) Calcular el nivel crítico o el valor p

$$p = 0,2105$$

6) Tomar una decisión

Como el valor $p = 0,2105$ no se rechaza la H_0 ni siquiera al 10%, ya que $0,2105 > 0,10$.

7) Dar una conclusión en términos del problema planteado

Se concluye que la Información sobre DM2 es independiente del estado nutricional de los enfermos de diabetes adscriptos al Programa Nacional de Diabetes y que retiran medicamentos de la farmacia del Hospital Dr. Nanzer de Sto. Tomé, Sta. Fe.

Salida computacional (Statistix)

Statistix 8.0				
Chi-Square Test for Heterogeneity or Independence				
Variable				
Case		DomSal	DomNoSal	
1	Observed	123	26	149
	Expected	125.16	23.84	
	Cell Chi-Sq	0.04	0.20	
2	Observed	24	2	26
	Expected	21.84	4.16	
	Cell Chi-Sq	0.21	1.12	
		147	28	175
Overall Chi-Square		1.57		
P-Value		0.2105		
Degrees of Freedom		1		
CAUTION: 1 cell(s) have expected values less than 5.0 ⁴				
Cases Included 4		Missing Cases 0		

6.- Conceptualice

- i) ¿Qué se entiende por hipótesis estadística?
- ii) Dé un ejemplo de una población en la que tenga por objetivo estudiar:
 - a) una variable cuantitativa. Formule acerca de la misma una H_0 y H_1 . Indique qué significaría en ese contexto cometer error de tipo I y cometer error de tipo II.
 - b) dos variables cualitativas cuya asociación desea investigar. Formule las hipótesis nula y alternativa correspondientes. Interprete los dos tipos de errores posibles, siempre en el contexto de su ejemplo.
- iii) Sintetice la lógica del test de hipótesis.
- iv) Defina nivel de significación y cómo incide su valor en la potencia de la prueba.
- v) ¿Por qué el valor p es más informativo que el valor crítico para tomar la decisión?
- vi) ¿Por qué es necesario calcular un tamaño de efecto?
- vii) ¿A qué debe su nombre de “nula” la hipótesis H_0 ?

Para elaborar estas respuestas es recomendable que releas el Capítulo 11 y Apéndice del libro de Bologna (2018) y el material y la clase teórica 9.

7.- Ubique en el material de la cátedra el artículo de Novo, M. et al. (2009). Lea los apartados: Resumen, *Participants* y la primera parte de Results donde alude a Table 1.

a) En *Participants* se informa $\chi^2(1, N = 116) = 0.31, ns$. ¿Qué significa esto en la descripción de la muestra?

El apartado comienza describiendo la muestra, la cual dice que fue “equalised in gender” y a continuación informa el valor $\chi^2(1, N = 116) = 0.31, ns$. Eso da un indicio de que corresponde a una prueba de independencia entre la composición por sexo en la población normativa (varón / mujer) y en la muestra

⁴ Cuando el valor esperado es inferior a 5 en una proporción alta de celdas, se requiere hacer ciertas correcciones al método que exceden los objetivos de este curso. Los autores del paper en cuestión realizaron el test de hipótesis tal cual se desarrolla en este ejercicio (sin las correcciones propuestas por la teoría estadística).

(varón / mujer). Por eso es una tabla de 2x2 y los grados de libertad son $(2-1) \times (2-1) = 1$. El valor del estadístico 0,31 indica la discrepancia entre la composición por sexos en la población normativa y en la muestra, la que resultó no significativa (ns). Es una forma de decir que la muestra no está sesgada por sexo sino que representa lo que acontece en la población normativa.

b) i) ¿De cuántas pruebas de hipótesis se informan resultados en *Table 1*?

De tantas como cantidad de valores del estadístico t y valores p hay; es decir, 12.

ii) ¿Qué tipo de pruebas de hipótesis son? ¿Dónde se informan los grados de libertad dentro de la tabla?

Como afirma el mismo título de la tabla, son pruebas de hipótesis t para una media. Los grados de libertad se informan al pie de la tabla, donde dice "Note: $df(115)$ ".

iii) Elija tres de ellas, explicita las hipótesis nula y alternativa, y verifique los valores p .

1) X = Puntaje en la dimensión de Obsesivo-Compulsivo del Brief Symptom Inventory (BSI), de los progenitores separados que asisten al Programa de "Ruptura de Pareja, no de Familia".

$H_0: \mu = 0,43$ No difieren en promedio del grupo normativo. (0,43 es la media el grupo normativo).

$H_1: \mu > 0,43$ Puntúan en promedio más alto que en el grupo normativo.

$t_{obs} = 7,92$

valor $p = P(t_{115} > 7,92) \approx 0$

2) X = Puntaje en la dimensión de Depresión del Brief Symptom Inventory (BSI), de los progenitores separados que asisten al Programa de "Ruptura de Pareja, no de Familia".

$H_0: \mu = 0,28$ No difieren en promedio del grupo normativo. (0,28 es la media el grupo normativo).

$H_1: \mu > 0,28$ Puntúan en promedio más alto que en el grupo normativo.

$t_{obs} = 8,28$

valor $p = P(t_{115} > 8,28) \approx 0$

3) X = Puntaje en la dimensión de Ansiedad del Brief Symptom Inventory (BSI), de los progenitores separados que asisten al Programa de "Ruptura de Pareja, no de Familia".

$H_0: \mu = 0,35$ No difieren en promedio del grupo normativo. (0,28 es la media el grupo normativo).

$H_1: \mu > 0,35$ Puntúan en promedio más alto que en el grupo normativo.

$t_{obs} = 8,63$

valor $p = P(t_{115} > 8,63) \approx 0$

iv) ¿Qué frase del resumen alude a los resultados de estas pruebas de hipótesis?

Resultados: los resultados mostraron, en contraste con la población normativa, que los progenitores separados puntuaban significativamente más alto en los PSMs.

8.- Mariano es publicista y quiere demostrarle a su cliente que, a partir de la campaña que ideó y le vendió, logró aumentar la proporción de consumidores de la marca de su cliente. Para ello va a realizar un muestreo y una prueba de hipótesis. ¿Cuál de los siguientes niveles de significación elegiría Mariano para favorecer arribar a la conclusión que desea?

a. 0,01

b. 0,05

c. 0,10

Explicación: Un nivel de significación más alto amplía la zona de rechazo de H_0 , lo que favorece su rechazo y, por consiguiente, la aceptación de la hipótesis alternativa que se quiere probar.

9.- En un estudio sobre el efecto de un entrenamiento en meditación (mindfulness) sobre las habilidades sociales se administró la *Escala de Habilidades Sociales (EHS)*, de Gismero (2000) a un grupo de 25 estudiantes antes (Pre) y después (Post) de dicho entrenamiento.

Parte I

¿Qué test de hipótesis corresponde realizar?

- Un test t de diferencia de medias de muestras independientes.
- Un test t para la media de la diferencia de datos pareados.
- Un test χ^2 de independencia entre el entrenamiento y las habilidades sociales.

Explicación: al mismo grupo de 25 personas se las mide antes y después del entrenamiento. Para decidir si hay efecto se comparan las medias obtenidas en cada caso, lo que coincide con considerar la media de las diferencias entre los puntajes obtenidos entre las dos administraciones.

Parte II

Los datos se procesaron con Statistix y se obtuvo la salida computacional siguiente:

Paired T Test for Post - Pre

Null Hypothesis: difference = 0

Alternative Hyp: difference > 0

Mean	11.640
Std Error	0.8080
Mean - H0	11.640
Lower 95% CI	9.9723
Upper 95% CI	13.308
T	14.41
DF	24
P	0.0000

Cases Included 25 Missing Cases 0

Del análisis de la salida se concluye que:

- el entrenamiento tuvo efecto estadísticamente significativo mejorando las habilidades sociales.
- hubo un aumento de 11,64 puntos en la media que no resultó estadísticamente significativo.
- el tamaño del efecto del entrenamiento en Mindfulness fue de 14,41 puntos.

Explicación. El valor $p = 0$ es menor que cualquier nivel de significación, por lo que se rechaza H_0 y eso es equivalente a decir que se observó una diferencia (efecto) estadísticamente significativa. En el título de la tabla (Paired T Test for Post – Pre) se menciona que la diferencia se hace en el sentido: puntaje en el Postest (Post) – puntaje en el Pretest (Pre); por tanto la diferencia señalada en Mean (11.64) es a favor del Post test y por eso se concluye que tiene efecto mejorando las habilidades.

10.- García Jiménez, M. et al (2020) estudian la relación entre abandonar un proceso judicial abierto por violencia de género y el contacto de las mujeres que abrieron dichos procesos con sus parejas. Consideraron si las mujeres renunciaron a continuar con el proceso al inicio del mismo o cuando estuviera avanzado y, por otra parte, si el contacto con sus parejas estaba siendo, en ese momento de renuncia, nulo, ocasional o frecuente. Es decir que tomaron en cuenta 2 momentos (inicial o avanzado) y 3 tipos de contacto (nulo,

ocasional o frecuente) y generaron así una tabla de contingencia de 2 x 3. El valor observado del estadístico de χ^2 fue 7,19. El valor p fue

- a. 0,027 y se concluyó al 5% que las variables estaban asociadas.
- b. 0,007 y se concluyó al 1% que las variables estaban asociadas.
- c. 0,973 y se concluyó que las variables eran independientes.

Explicación. Al ser una tabla de 2 x 3, los grados de libertad son $1 \times 2 = 2$. El test de independencia es siempre unilateral a derecha. Utilizando la distribución χ^2 con 2 grados de libertad, la probabilidad a derecha del valor del estadístico 7,19 es 0,027 (redondeada), la cual implica un rechazo de H_0 (independencia) a cualquier nivel de significación mayor; por ejemplo, 0,05.

11.- El *valor p* para testear si un tratamiento tiene efecto beneficioso fue 0,0025.

¿Cuál de las siguientes interpretaciones es correcta?

- a. El tratamiento tiene un efecto importante.
- b. El tratamiento tiene un efecto cuyo tamaño debe ser calculado.
- c. La diferencia entre recibir o no recibir el tratamiento no es significativa.

Explicación. Este *valor p* de 0,0025 es inferior a todos los niveles de significación habituales, lo que llevaría a rechazar la hipótesis nula y admitir el efecto del tratamiento. Pero el *valor p* no es medida del efecto, la cual debe ser computada por un cálculo aparte como, por ejemplo, la d de Cohen.

12.- En cierto experimento de sensibilidad auditiva los investigadores quieren validar la afirmación de que el 80% de las personas es capaz de reconocer la diferencia entre dos frecuencias de sonido. Para ello realizan la prueba con 64 personas y obtienen que 44 de ellas pueden reconocer la diferencia. Con un nivel de significación del 1% resulta que

- a. se valida la afirmación, ya que el valor p es mayor que 0,01.
- b. se concluye que el porcentaje es menor, ya que el valor $p = 0,0122$.
- c. se concluye que hay una diferencia significativa al 1%, ya que el valor $p = 0,02445$.

Explicación. Es una prueba de hipótesis sobre una proporción. Como se quiere validar la afirmación de la igualdad ($H_0: p = 0,8$), la alternativa es bilateral ($p \neq 0,8$). La afirmación se valida si el valor $p > 0,01$. El valor

observado del estadístico de prueba estandarizado es $\frac{\frac{44}{64} - 0,8}{\sqrt{\frac{0,8 \times 0,2}{64}}} = -2,25$. Como el estadístico tiene bajo H_0

distribución aproximadamente normal estándar, para hallar el valor p en una prueba bilateral, se puede utilizar, en el Probability Distribution, la tercera opción de la derecha que da la probabilidad de las colas simétricas: a la derecha de 2,25 y a izquierda de -2,25; probabilidad que es 0,02445. Como ésta es mayor que 0,01, se mantiene (valida) H_0 .

El valor p también puede buscarse sin estandarizar el estadístico, utilizando la normal con media 0,8 y

desviación estándar $0,05 = \sqrt{\frac{0,8 \times 0,2}{64}}$ para un valor del estadístico de $0,6875 = 44/64$.

13.- Las puntuaciones en el Inventario de Burnout de Maslach-Encuesta General (Schaufeli, Leiter, Maslach, & Jackson, 1996) de los transportistas de una ciudad española se distribuyen normalmente con media $\mu = 2,3$ y desviación estándar $\sigma = 1,5$. Se quiso analizar si la puntuación promedio de los transportistas de la ciudad de La Plata superaría a la media de la de los transportistas de la ciudad española en dicha escala. Para ello fueron seleccionados 36 transportistas de la ciudad de La Plata y se registró su puntaje en el mencionado inventario. Se obtuvo una media de 2,8. Suponiendo que la desviación estándar en la población bonaerense es similar a la de la población española, se realizó la prueba de hipótesis adecuada al 5%. Se concluyó que

- a. la diferencia observada no es estadísticamente significativa; es atribuible al azar.
- b. la diferencia de medias poblacionales entre los transportistas de La Plata y los españoles es de 0,5.
- c.** los transportistas de La Plata están más estresados, en promedio, que los españoles.

Explicación. Es una prueba de hipótesis sobre una media con desviación estándar poblacional conocida, unilateral a derecha. Por tanto hay que buscar el valor p en la Normal poniendo en la aplicación los parámetros $\mu = 2,3$ y desviación estándar $\sigma = 1,5/6 = 2,25$ y $x = 2,8$. El valor p es la probabilidad a derecha de $2,8 = 0,02275$. Como es menor que 0,05, se rechaza la hipótesis nula y se concluye que $\mu > 2,3$, con la consecuente interpretación.

14.- Se ha realizado un estudio acerca de los factores psicosociales que intervienen en el comportamiento tabáquico en adolescentes españoles. Entre otras variables de personalidad se ha registrado el puntaje de ansiedad rasgo obtenido mediante el cuestionario de Autoevaluación de Ansiedad Estado/Rasgo en niños (STAIC). Se clasificó a 20 jóvenes seleccionados al azar en dos grupos: 10 adolescentes fumadores (Grupo 1) y 10 adolescentes no fumadores (Grupo 2). Se realizó la prueba de hipótesis con el programa Statistix para estudiar si el nivel de ansiedad rasgo difiere significativamente al 10% entre estos dos grupos de adolescentes. Se supone que estos puntajes se distribuyen normalmente. Se obtuvo la siguiente salida:

Variable	Mean	N	SD	SE		
Grupo 1	35.200	10	3.0111	0.9522		
Grupo 2	34.100	10	2.3781	0.7520		
Difference	1.1000					
Null Hypothesis: difference = 0						
Alternative Hyp: difference <> 0						
				95% CI for Difference		
Assumption	T	DF	P	Lower	Upper	
Equal Variances	0.91	18	0.3766	-1.4492	3.6492	
Unequal Variances	0.91	17.1	0.3772	-1.4590	3.6590	
Test for Equality	F	DF	P			
of Variances	1.60	9,9	0.2465			

Según la salida precedente se concluye que, **en la población considerada**

- a. los fumadores son en promedio más ansiosos que los no fumadores.
- b.** fumadores y no fumadores no difieren en promedio en cuanto a su ansiedad.
- c. el tabaco tiene un efecto moderado elevando los niveles de ansiedad.

Explicación. Es un test de diferencia de medias de muestras independientes, bilateral. Como el valor p (0,3766) es superior al nivel de significación 0,10, no se rechaza la igualdad de medias.