



UNIVERSIDAD DE BUENOS AIRES

ESTADÍSTICA - CÁTEDRA I

CÓDIGO 60

PROF. ADJUNTA A CARGO: DRA. MARÍA SILVIA GALIBERT

PRÁCTICA DE REPASO

UNIDADES 4 y 5

AUTORES

Profesores y Docentes de la Cátedra I



PRÁCTICA DE REPASO PARA EL SEGUNDO PARCIAL

1.- La Lic. Julieta Díaz trabaja en el gabinete psicopedagógico de un centro de salud. Ella está especializada en Trastorno por déficit de atención e hiperactividad (TDHA). Los chicos con este tipo de trastorno constituyen el 40% de los que llegan al centro de salud por algún problema de aprendizaje. Su tiempo de atención diaria es tal que si un día llegaran más de 3 a su consulta debería derivarlos a otro profesional. Si un día llegan al centro de salud 6 consultas, calcular la probabilidad de

- a) que tenga que derivar pacientes a otros profesionales.
- b) al menos una de las consultas sea por TDHA.

2.- A una Asociación Mutual asisten trabajadores en calidad de activos o pasivos. El 60% de los asistentes a la Mutual son trabajadores pasivos. Sea X la variable que indica la cantidad de trabajadores pasivos que se presentan de un total de 10 trabajadores que acuden a dicha Asociación Mutual.

¿A qué modelo de distribución de probabilidad responde la variable X , qué parámetros lo caracterizan y qué condiciones deben cumplirse para dicho ajuste?

- a) ¿Cuál es la probabilidad de que a lo sumo 2 trabajadores sean pasivos?
- b) ¿Cuál es la probabilidad de que menos de la tercera parte de los trabajadores sean pasivos?
- c) ¿Cuál es la probabilidad de que por lo menos 7 trabajadores sean activos?
- d) ¿Cuál es la probabilidad de que por lo menos 1 y a lo sumo 3 trabajadores sean activos?

3.- Los puntajes en un test de habilidades matemáticas se distribuyen normalmente con media 50 y desviación estándar 10. Stefanía tiene un rango percentilar de 80 y Luis un puntaje de 62.

- a) Hallar el puntaje de Stefanía.
- b) Hallar el rango percentilar de Luis.
- c) Comparar el desempeño de ambos con un criterio adecuado.

4.- Suponiendo que las puntuaciones correspondientes a la escala de *Factores Organizacionales* que integra el Cuestionario de Acoso Psicológico (Moreno Jiménez, B. et al., 2004) para un grupo normativo de trabajadores del sector de transportes y de comunicaciones de la Comunidad de Madrid pueden ser modelados con una distribución normal con media $\mu = 27$ y desvío a $\sigma = 3$.

- a) En relación a este modelo ¿cuál es la probabilidad de que el puntaje correspondiente a la escala de Factores Organizacionales supere el valor 30?
- b) ¿Cuál es el puntaje en dicha escala que es superado por el 30% de las observaciones más altas?
- c) Se elige una muestra aleatoria simple de 36 trabajadores de dicho sector laboral. ¿Cuál es la probabilidad de que la media muestral de los puntajes correspondiente a la escala de *Factores Organizacionales* supere el valor 30?
- d) Explique con la ayuda de un gráfico o esquema por qué las respuestas de los incisos a) y c) no son las mismas.

5.- Se cree que el tiempo medio para resolver una prueba de matemática en un operativo de evaluación en cierto establecimiento educativo es de 40 minutos. Se tomó una muestra aleatoria de 9 alumnos y se le administró la prueba a modo piloto. Los tiempos (en minutos) que necesitaron para resolverla fueron:

20 – 25 – 30 – 32 – 35 – 38 – 40 – 45 – 50

¿Qué puede concluir al 10% de significación?

Desarrolle el procedimiento paso por paso pero no haga cálculos; en su lugar utilice e interprete la salida computacional correspondiente.

Statistix 8.0

26/10/2019, 22:18:34

One-Sample T TestNull Hypothesis: $\mu = 40$ Alternative Hyp: $\mu \dots\dots 40$ (Complete la línea de puntos)

Variable	Mean	SE	95% Conf Interval		T	DF	P
			Lower	Upper			
TIEMPO	35.000	3.1579	27.718	42.282	-1.58	8	0.1520

6.- Cierta empresa aplicó 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. Para tomar la decisión, seleccionó al azar una muestra de 80 empleados y les administró una encuesta que, entre otras cosas, preguntaba: ¿considera que esta técnica lo motiva significativamente 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.

7.- Se leyó una lista de 12 palabras a una muestra de 9 personas y las cantidades de palabras recordadas fueron: 4 – 4 – 2 – 5 – 7 – 9 – 10 – 5 – 6.

A otra muestra de 9 personas se les leyeron las mismas palabras acompañadas por imágenes y las cantidades de palabras recordadas fueron: 2 – 5 – 5 – 7 – 12 – 11 – 9 – 8 – 7

a) A partir de la salida de Statistix decida si las poblaciones de la cantidad de palabras recordadas por los sujetos que sólo oyen las palabras y la de las que las oyen acompañadas por imágenes tienen o no igual varianza con un nivel de significación del 10%. (Indique el lugar de la salida donde se halla esta información).

a) Reemplace la línea de puntos en la hipótesis alternativa por lo que corresponda.

b) Indique cuál es la diferencia registrada entre las medias de palabras recordadas por los dos grupos y a favor de quién es dicha diferencia. ¿Es esa diferencia estadísticamente significativa al 5%? ¿En qué parte de la salida se encuentra esta información? ¿Qué puede concluir con respecto al efecto de acompañar las palabras con imágenes sobre la recordación?

Two-Sample T Tests for PALABRAS by IMAGENES

IMAGENES	Mean	N	SD	SE
Sin imagen	5.7778	9	2.5386	0.8462
Con imagen	7.3333	9	3.1225	1.0408
Difference	-1.5556			

Null Hypothesis: difference = 0

Alternative Hyp: difference ... 0

Assumption	T	DF	P	95% CI for Difference	
				Lower	Upper
Equal Variances	-1.16	16	0.2632	-4.3992	1.2881
Unequal Variances	-1.16	15.4	0.2639	-4.4089	1.2978

Test for Equality of Variances	F	DF	P
	1.51	8,8	0.2858

Cases Included 18 Missing Cases 0

8.- Se aplicó un programa de entrenamiento psicoprofiláctico sobre un grupo de 8 dirigentes. Se analizaron autovaloraciones e imágenes que sobre ellos tenían sus subordinados a través de una técnica (Bachmann y Strumper, 1982) acerca de la "Aceptación del interlocutor" compuesta por ítems positivos a los que el encuestado debe responder en una escala de 1 a 3 (1: No es cierto, 2: Algo cierto, 3: Absolutamente cierto). Se comparó al grupo antes mencionado con otro grupo de 10 dirigentes que no recibieron entrenamiento. Una de las variables que se obtiene con este instrumento refleja la imagen del dirigente mediante un puntaje X obtenido promediando puntajes de encuestas que sobre cada dirigente hicieron sus subordinados, entendiendo que a mayor puntaje corresponde una mayor valoración del mismo. Luego de la aplicación del programa se obtuvieron los siguientes resultados.

Grupo de dirigentes entrenados: 56 57 52 61 72 74 72 66
 Grupo de dirigentes no entrenados: 38 48 36 42 57 46 60 35 55 59

- a) ¿Son estos resultados congruentes con la hipótesis que el entrenamiento contribuirá a que los sujetos que lo reciben sean mejor valorados por sus subordinados? Usar un nivel de significación del 5%.
 b) ¿Qué supuestos deberían cumplirse para que el proceso de prueba de hipótesis sea válido?

Nota. Este ejercicio es una adaptación con fines didácticos del artículo de Castillo, N. et al. (1994).

9.- Un encuesta fue realizada en base a una muestra aleatoria de 400 trabajadores del sector transportes a los que se los indagó acerca de dos variables: presencia o ausencia de Acoso Psicológico (detectado al superar cierto umbral en el Cuestionario de Acoso Psicológico) y la variable Tipo de Contrato alcanzado por el trabajador (eventual, a plazo fijo, planta permanente). A continuación se presenta la distribución conjunta de estas dos variables en la muestra, así como los resultados del test de independencia efectuado con el programa Statistix.

Chi-Square Test for Heterogeneity or Independence

Acoso Psicológico		Tipo de Contrato			
		Eventual	Plazo Fijo	Permanente	
Presencia	Observed	92	47	37	176
	Expected	80.52	44.88	50.60	
	Cell Chi-Sq	1.64	0.10	3.66	
Ausencia	Observed	91	55	78	224
	Expected	102.48	57.12	64.40	
	Cell Chi-Sq	1.29	0.08	2.87	
		183	102	115	400
Overall Chi-Square		9.63			
P-Value		0.0081			
Degrees of Freedom		2			
Cases Included		6			
Missing Cases		0			

- a) ¿Cuál es la hipótesis nula de interés?
 b) Si H_0 fuera verdadera, cuántos trabajadores con contrato eventual piensa Usted que habrían sido clasificados con presencia de acoso psicológico?
 c) Si H_0 fuera verdadera, cuántos trabajadores con contrato de planta permanente piensa Usted que no hubieran presentado características de acoso psicológico?
 d) Ponga a prueba la hipótesis nula del inciso a) con un nivel de significación del 1%.

10.- Ubique el artículo de Martín-Ramos, R. et al (2015) en el material de la Cátedra o en la web: <https://www.medigraphic.com/pdfs/arcneu/ane-2015/ane154a.pdf>. Lea los párrafos de Objetivo (p.234), *Instrumentos de evaluación psicológica* (p.235), Tabla 2 (p.235) y Tabla 4 (p. 236).

- ¿De cuántas pruebas de hipótesis son los resultados de la Tabla 2, qué tipo de pruebas son y qué objetivo tienen?
- ¿En cuántas de ellas se obtuvieron resultados estadísticamente significativos? ¿A qué conclusión se llega en cada caso?
- Sabiendo que los grupos control sano y con esclerosis múltiple estaban conformados por 30 personas cada uno, halle el tamaño del efecto para las pruebas de Diseño con cubos (Waiss-III), Fluidez verbal fonológica (COWA) e Inventario de depresión de Beck (IDB). Exprese una conclusión en cada caso.
- Según la Tabla 3 ¿Qué correlaciones resultaron significativas?
- ¿Cómo se relacionan estas correlaciones con los resultados de la Tabla 4?
- Considere los resultados de los análisis de regresión de la Tabla 4. ¿Cuáles resultaron estadísticamente significativos? En esos casos, ¿cuál fue la proporción de varianza explicada?

11.- Ubique el artículo de Huertas, E. et al (2019) en el material de la Cátedra o en la web: <http://www.psicothema.com/pdf/4538.pdf>. Lea el resumen y los párrafos de *Participants* (pág.240) y *Genetic data* (pág. 241)

- i) ¿Se realizó algún tipo de muestreo aleatorio?
- ii) ¿Cuál es la población considerada?
- Al describir los resultados de las pruebas de asociación se mencionan algunos valores del estadístico χ^2 y sus correspondientes valores p. Utilizando los datos de *Table 1* verifique ambas cosas en las pruebas de asociación entre el genotipo y el sexo. ¿Qué se concluyó al respecto?

En los siguientes ítems elija la única opción correcta.

12.- En una distribución binomial de parámetros n y p , la media y la mediana coinciden cuando

- $n \geq 30$.
- la media es np .
- $p = 0,5$.

13.- Se cree que el 60% de los enfermeros de cierto hospital tiene la opinión de que se deben hacer modificaciones substanciales en el protocolo sanitario. Si así fuera, ¿cuál es la probabilidad de que, entre 16 enfermeros elegidos al azar de dicho hospital, más del 75% opine que hay que hacer tales modificaciones?

- 0,10142
- 0,16657
- 0,06515

14.- El puntaje en un test de depresión se distribuye normalmente en cierta población. Un psicólogo atiende problemas de depresión moderados pero deriva los casos que considera muy severos a otro especialista. Considera severos los casos de puntuaciones mayores a la media en más de una desviación estándar. ¿Cuál es la probabilidad de que tenga que derivar a lo sumo 3 pacientes entre los próximos 6 que le solicitan tratamiento?

- 0,05482
- 0,99275
- 0,04757

15.- Para estudiar el poder de discriminación de un nuevo test de ansiedad quieren compararse los puntajes que obtienen en dicho test dos grupos que están en los polos de la distribución de los puntajes de ansiedad que se obtienen con el inventario STAI; para ello se consideran los individuos que están en la primera y en la última intercuartila. Suponiendo que el puntaje en STAI se distribuye normalmente con media 25 y desviación estándar 3, los límites (redondeados) de dichas intercuartilas son:

- a. 23 y 27
- b. 25 y 75
- c. 22 y 28

16.- Un test de evaluación cognitiva se distribuye normalmente con media 27 y desviación estándar 4. Se considera que manifiestan problemas cognitivos puntajes menores o iguales a 21. ¿Qué cantidad de personas con problemas cognitivos cabe esperar en una población de 10000 personas?

- a. 1336
- b. 9332
- c. 668

17.- Para llevar a cabo los métodos de la inferencia estadística es necesario conocer la distribución de probabilidades de los estadísticos para

- a. calcular el verdadero valor de los parámetros.
- b. controlar los errores en las pruebas de hipótesis.
- c. obtener el error estándar de los estimadores.

18.- Los puntajes en una escala de distracción al realizar tareas escolares on-line se distribuyen normalmente con media 12 en un grupo normativo. Puntajes superiores representan mayores niveles de distracción. Una maestra ensayó un programa para mejorar la atención en este tipo de tareas, en una muestra piloto de 9 estudiantes. Finalizado el programa se les administró el test de distracción. Se obtuvo una desviación estándar de 4 puntos. Para decidir sobre la eficacia del programa se hará la correspondiente prueba de hipótesis con un nivel de significación del 5%. Se concluirá que el programa es eficaz sólo si en la muestra se observara una media

- a. inferior a 9,52.
- b. superior a 14,48.
- c. inferior a 9,81.

19.- En cierta escala de amabilidad los puntajes se distribuyen normalmente con una media de 46 puntos y una desviación estándar de 8 puntos en el grupo normativo. Se consideran extremadamente amables las personas cuyo puntaje supera los 56 puntos. Se administró la escala en una localidad para corroborar, mediante una prueba de hipótesis, si la proporción de gente extremadamente amable coincide con la del grupo normativo. En una muestra de 200 personas, 28 resultaron extremadamente amables. El valor p es:

- a. 0,0570
- b. 0,1141
- c. 0,1057

20.- Dos proyectos de investigación compiten por un subsidio en un concurso. Los puntajes en una escala de 1 a 5 obtenidos por cada uno de ellos a partir de la opinión de 7 jueces son los siguientes:

Juez	1	2	3	4	5	6	7
Proyecto A	4	5	4	4	3	5	3
Proyecto B	2	4	4	5	3	4	1

Considérense a estos jueces como una muestra representativa de una población de expertos en el tema. Para decidir si los puntajes asignados a los proyectos difieren significativamente en promedio se debe realizar una prueba de hipótesis para

- a. la diferencia de medias de muestras independientes.
- b. la homogeneidad de varianzas de muestras independientes.
- c. la media de la diferencia de datos pareados.

21.- En una prueba de hipótesis la potencia disminuye cuanto

- a. menor es el nivel de significación.
- b. mayor es el tamaño de muestra.
- c. menor es el valor p .

22.- La cantidad promedio de errores gramaticales en una muestra de niños que recibieron capacitación específica en lengua resultó significativamente inferior a la cantidad promedio de dichos errores en la población objetivo sin la capacitación. Se utilizó un nivel de significación del 5%. De ello se concluye, con una probabilidad de error de 0,05, que:

- a. La capacitación tiene el efecto de disminuir el promedio de los errores en una medida considerada muy importante.
- b. La muestra de niños resultó sesgada, por lo que se obtuvieron puntajes inferiores a los que podría esperarse.
- c. la media de errores en una población de niños capacitados sería inferior a la de la población sin la capacitación.

REFERENCIAS

- Bachmann, W., y Strumper, R. (1982). Programa para el entrenamiento psicoprofiláctico. Berlín: Instituto Central de Medicina del Trabajo de la República Democrática Alemana.
- Castillo, N., Almirall, P., Casacó, A., Román, J. (1994). Salud ocupacional: efectos de un programa de entrenamiento psicoprofilactico sobre un grupo de dirigentes *Revista Latinoamericana de Psicología* 26, 2, 271-282.
- Huertas, E., López-Moreno, J., Fernández, V., Echeverry-Alzate, V. y Bühler, K. (2019). Associations between experimental substance use, FAAH-gene variations, impulsivity and sensation seeking. *Psicothema* 31, 3, 239-245.
- Moreno Jiménez, B., Rodríguez Muñoz, A., Garrosa Hernández, E. y Morante Benadero, M. (2005). Antecedentes organizacionales del acoso psicológico en el trabajo: un estudio exploratorio. *Psicothema* 17, 4, 627-632.
- Martin-Ramos, R., Jiménez-Morales, R., Herrera-Jiménez, L., Guerra-Pérez, L., Martin-Ramos, R. y Pérez-Medinilla, Y. (2015). Influencia de la ansiedad y depresión sobre las funciones ejecutivas en pacientes con esclerosis múltiple. *Archivos de Neurociencias* 20, 4, 233-239.

RESOLUCIÓN

1.- La Lic. Julieta Díaz trabaja en el gabinete psicopedagógico de un centro de salud. Ella está especializada en Trastorno por déficit de atención e hiperactividad (TDHA). Los chicos con este tipo de trastorno constituyen el 40% de los que llegan al centro de salud por algún problema de aprendizaje. Su tiempo de atención diaria es tal que si un día llegaran más de 3 a su consulta debería derivarlos a otro profesional. Si un día llegan al centro de salud 6 consultas, calcular la probabilidad de

a) que tenga que derivar pacientes a otros profesionales.

X = Cantidad de pacientes que requieren atención por TDHA entre 6 que se presentan en el centro de salud.

Supondremos que los pacientes que llegan al centro de salud requieren o no tratamiento por TDHA independientemente unos de otros (Independencia. Requiere suponer que la población de los pacientes que llegan es mucho mayor que 6). Por otra parte, como la única fuente de información para asignar probabilidad de que se requiera tratamiento para TDHA es el dato del 40% histórico, la probabilidad que se asigna a requerir atención por TDHA es la misma para todos: 0,4 (probabilidad de éxito constante). Por tanto podemos decir que la variable X se distribuye según el modelo Binomial de parámetros $n = 6$ y $p = 0,4$. En símbolos: $X \sim B(6 ; 0,4)$

Derivar pacientes a otro profesional equivale a que X exceda a 3 ($X > 3 = X \geq 4$), cuya probabilidad se pide calcular; esto es:

$$P(X \geq 4) = 0,1792 \text{ (con Probability Distribution)}$$

Respuesta: La probabilidad de que la Lic. Julieta Díaz tenga que derivar pacientes a otro profesional es 0,1792. *(Nota del profesor: Lo que se interpreta como que debe hacerlo el 17,92% de las veces que llegan 6 pacientes al consultorio).*

b) al menos una de las consultas sea por TDHA.

$$P(X \geq 1) = 0,95334 \text{ (con Probability Distribution)}$$

Si se utiliza una aplicación que no tiene probabilidades acumuladas a derecho, se calcula así:

$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0,04666 = 0,95334$$

Respuesta: La probabilidad de que al menos una consulta entre 6 que llegan sea por TDHA es 0,95334.

2.- A una Asociación Mutual asisten trabajadores en calidad de activos o pasivos. El 60% de los asistentes a la Mutual son trabajadores pasivos. Sea X la variable que indica la cantidad de trabajadores pasivos que se presentan de un total de 10 trabajadores que acuden a dicha Asociación Mutual.

¿A qué modelo de distribución de probabilidad responde la variable X , qué parámetros lo caracterizan y qué condiciones deben cumplirse para dicho ajuste?

X = Cantidad de trabajadores pasivos que se presentan de un total de 10 trabajadores que acuden a dicha Asociación Mutual.

Se supone que los trabajadores que asisten tienen la condición de ser activos o pasivos independientemente unos de otros (Independencia – Como el muestreo es sin reposición, en rigor hay que suponer que la muestra de 10 es pequeña en relación a la población de todos los trabajadores que asisten a tal mutual). Además la probabilidad de que un trabajador que acude a la mutual sea pasivo proviene de la composición porcentual

informada, 60%, por la que la probabilidad de que un asistente elegido al azar entre los que van a la mutual sea pasivo es la misma para cualquiera: 0,6.

Bajo esas condiciones $X \sim B(10 ; 0,6)$

a) ¿Cuál es la probabilidad de que a lo sumo 2 trabajadores sean pasivos?

$$P(X \leq 2) = 0,01229$$

b) ¿Cuál es la probabilidad de que menos de la tercera parte de los trabajadores sean pasivos?

La tercera parte de 10 es 3,33. Menos de la tercera parte es hasta 3 inclusive.

$$P(X \leq 3) = 0,05476$$

c) ¿Cuál es la probabilidad de que por lo menos 7 trabajadores sean activos?

Por lo menos 7 activos significa 7 o más activos. Como la variable X está definida en función de los pasivos, eso significa que hay hasta 3 pasivos, lo cual se calculó en b); es decir, la respuesta es 0,05476.

d) ¿Cuál es la probabilidad de que por lo menos 1 y a lo sumo 3 trabajadores sean activos?

Un activo equivale a 9 pasivos y 3 activos a 7 pasivos. Entonces, en términos de X se pide calcular $P(7 \leq X \leq 9) = P(X \leq 9) - P(X \leq 6) = 0,99395 - 0,61772 = 0,37623$

También puede calcularse sumando las probabilidades puntuales:

$$P(7 \leq X \leq 9) = P(X = 7) + P(X = 8) + P(X = 9) = 0,21499 + 0,12093 + 0,04031 = 0,37623$$

3.- Los puntajes en un test de habilidades matemáticas se distribuyen normalmente con media 50 y desviación estándar 10. Stefanía tiene un rango percentilar de 80 y Luis un puntaje de 62.

X = Puntaje en un test de habilidades matemáticas. $X \sim N(50 ; 10)$

a) Hallar el puntaje de Stefanía.

El dato de Stefanía es su rango percentilar 80, lo que corresponde a una probabilidad de 0,8 acumulada a la izquierda del puntaje de Stefanía, que es lo que queremos calcular. En otras palabras, queremos calcular el percentil 80 de la distribución. Lo llamaremos p_{80} . Para ello entramos a la aplicación que utilizemos e informamos de los valores de los parámetros $\mu = 50$ y $\sigma = 10$. Informamos también el valor de probabilidad acumulada a derecha: 0,80. Si estamos utilizando Probability Distribution volcamos ese valor en la ventana de la derecha y miramos la devolución del percentil en la ventana de la izquierda. Si estamos usando EXCEL, debemos buscar la Normal Inversa.

$$P(X < p_{80}) = 0,80 \Rightarrow p_{80} = 58,41621$$

El puntaje (redondeado) de Stefanía es 58,42.

b) Hallar el rango percentilar de Luis. El rango percentilar de 62 es la probabilidad acumulada a la izquierda de 62 expresada en unidades de 100.

$$P(X < 62) = 0,88493$$

El rango percentilar que le corresponde a Luis es 88,49.

c) Comparar el desempeño de ambos con un criterio adecuado.

Pueden compararse sus puntajes o equivalentemente sus rangos percentilares.

Criterio del puntaje:

Como el puntaje de Luis (62) es mayor que el de Stefanía (58,42), el desempeño de Luis en las habilidades matemáticas medidas por el test es superior que el de Stefanía.

Criterio del rango percentilar: Luis puntúa por encima del 88,49% de la población mientras que Stefanía por encima del 80%; por lo que Luis tiene mejor desempeño que Stefanía.

4.- Suponiendo que las puntuaciones correspondientes a la escala de *Factores Organizacionales* que integra el Cuestionario de Acoso Psicológico (Moreno Jiménez, B. et al., 2004) para un grupo normativo de trabajadores del sector de transportes y de comunicaciones de la Comunidad de Madrid pueden ser modelados con una distribución normal con media $\mu = 27$ y desvío a $\sigma = 3$.

X = Puntaje en la escala de Factores Organizacionales del Cuestionario de Acoso Psicológico.

$$X \sim N(27; 3)$$

a) En relación a este modelo ¿cuál es la probabilidad de que el puntaje correspondiente a la escala de Factores Organizacionales supere el valor 30?

$$P(X > 30) = 0,15866$$

b) ¿Cuál es el puntaje en dicha escala que es superado por el 30% de las observaciones más altas?

Se pide calcular el puntaje que acumula a su derecha una probabilidad de 0,3 (o equivalentemente 0,7 a izquierda); en otras palabras se pide calcular el percentil 70 de la distribución: p_{70} .

$$P(X > p_{70}) = 0,30 \Rightarrow p_{70} = 28,5732$$

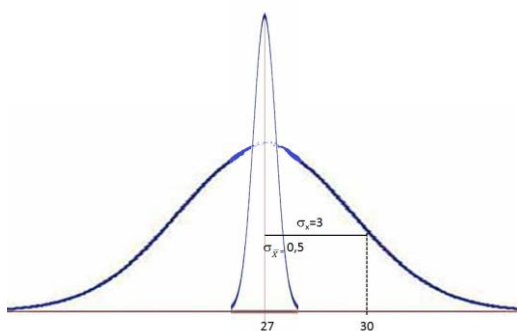
c) Se elige una muestra aleatoria simple de 36 trabajadores de dicho sector laboral. ¿Cuál es la probabilidad de que la media muestral de los puntajes correspondiente a la escala de *Factores Organizacionales* supere el valor 30?

La media muestral $\bar{X}$ es normal con la misma media de X (27) y la desviación estándar de X dividida la raíz del tamaño de muestra. En símbolos:

$$\bar{X} \sim N\left(27, \frac{3}{\sqrt{36}}\right) = N(27; 0,5)$$

$$P(\bar{X} > 30) \approx 0$$

d) Explique con la ayuda de un gráfico o esquema por qué las respuestas de los incisos a) y c) no son las mismas.



La campana ancha representa la distribución de la variable X cuya desviación estándar es 3. La campana delgada corresponde a la distribución de la media muestral $\bar{X}$ que también es normal con la misma media de X (por eso están centradas en el mismo eje) pero cuya desviación estándar es la sexta parte de la misma: 0,5. El diagrama es sólo un bosquejo; esta última campana debería dibujarse mucho más larga porque debe encerrar la misma área: 1, que la anterior. Puede observarse que a la derecha de 30 hay una buena porción de área por debajo de la campana de X , la que corresponde a la probabilidad de 0,15866 calculada en a). Pero, a la derecha de 30, debajo de la curva delgada de $\bar{X}$, el área es prácticamente nula, despreciable; por eso la probabilidad da aproximadamente cero. Recordemos que en cualquier distribución

normal, la probabilidad a más de tres unidades de desviación estándar es casi nula y 30 está a 6 unidades de desviación estándar ($6 \times 0,5 = 3$) por encima de 27.

5.- Se cree que el tiempo medio para resolver una prueba de matemática en un operativo de evaluación en cierto establecimiento educativo es de 40 minutos. Se tomó una muestra aleatoria de 9 alumnos y se les administró la prueba a modo piloto. Los tiempos (en minutos) que necesitaron para resolverla fueron:
20 – 25 – 30 – 32 – 35 – 38 – 40 – 45 – 50

¿Qué puede concluir al 10% de significación?

Desarrolle el procedimiento paso por paso pero no haga cálculos; en su lugar utilice e interprete la salida computacional correspondiente.

Statistix 8.0 26/10/2019, 22:18:34

One-Sample T Test

Null Hypothesis: $\mu = 40$

Alternative Hyp: $\mu \dots\dots 40$ (Complete la línea de puntos)

Variable	Mean	SE	95% Conf Interval		T	DF	P
			Lower	Upper			
TIEMPO	35.000	3.1579	27.718	42.282	-1.58	8	0.1520

1) Nombrar la/s variable/s y los supuestos necesarios.

X : tiempo (en minutos) que demora en resolver una prueba de matemática un alumno de cierto establecimiento educativo.

Como es una muestra pequeña, se necesita hacer el supuesto de que X se distribuye normalmente.

2) Plantear las hipótesis simbólica y coloquialmente.

$H_0: \mu = 40$ el tiempo medio que tardan los alumnos del establecimiento en resolver la prueba de matemática es 40 minutos.

$H_1: \mu \neq 40$ el tiempo medio que tardan los alumnos del establecimiento en resolver la prueba de matemática difiere de 40 minutos.

3) Indicar el nivel de significación de la prueba.

$\alpha = 0,10$

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

$$t = \frac{\bar{X} - 40}{\frac{S}{\sqrt{9}}} \sim t_8 \quad \text{bajo } H_0$$

5) Calcular el valor observado del estadístico de prueba y el valor p .

Estos valores están en la salida computacional:

$t_{\text{obs}} = -1,58$

valor $p = 0,1520$

6) Establecer la regla de decisión y tomar la decisión.

Se rechaza la hipótesis nula si y sólo si el valor $p < 0,10$

Como $0,1520 > 0,1$ no se rechaza la hipótesis nula.

7) Interpretar la decisión en términos del problema.

Los estudiantes del establecimiento necesitan un promedio de 40 minutos para realizar la prueba de matemática.

6.- Cierta empresa aplicó 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. Para tomar la decisión, seleccionó al azar una muestra de 80 empleados y les administró una encuesta que, entre otras cosas, preguntaba: ¿considera que esta técnica lo motiva significativamente 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 sobre una proporción poblacional.

1) Variables y supuestos mínimos necesarios para llevar a cabo la prueba

X : Opinión de los empleados de una empresa sobre si una técnica recibida para hacer uso del sentido del humor como reforzador motivacional es beneficiosa o no.

Supuestos. Las opiniones de los empleados son independientes unas de otras con la misma probabilidad de responder afirmativamente.

2) Hipótesis estadísticas

$H_0: p \leq 0,75$ A lo sumo el 75% de los empleados consideran que la técnica es beneficiosa $\Rightarrow$ No se continúa con su implementación.

$H_1: p > 0,75$ Más del 75% de los empleados consideran que la técnica sería beneficiosa $\Rightarrow$ Se se continúa con su implementación.

3) Nivel de significación: $\alpha = 0,10$

4) Estadístico de prueba y su distribución bajo H_0

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) Cálculo del valor observado del estadístico de prueba

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

6) Cálculo del nivel crítico o valor p

La prueba es unilateral a derecha; por lo que el valor p es $P(\hat{p} > 0,85 \text{ bajo } H_0) = 0,01941$.

7) Conclusión

Como $0,01941 < 0,10$ se rechaza H_0 ; es decir que se aconsejaría a la empresa que continúe con la aplicación de la técnica.

7.- Se leyó una lista de 12 palabras a una muestra de 9 personas y las cantidades de palabras recordadas fueron: 4 – 4 – 2 – 5 – 7 – 9 – 10 – 5 – 6.

A otra muestra de 9 personas se les leyeron las mismas palabras acompañadas por imágenes y las cantidades de palabras recordadas fueron: 2 – 5 – 5 – 7 – 12 – 11 – 9 – 8 – 7

a) A partir de la salida de Statistix decida si las poblaciones de la cantidad de palabras recordadas por los sujetos que sólo oyen las palabras y la de las que las oyen acompañadas por imágenes tienen o no igual varianza con un nivel de significación del 10%. (Indique el lugar de la salida donde se halla esta información).

El valor p para decidir si las varianzas son no iguales entre las poblaciones se halla en la siguiente parte de la salida de Statistix:

Test for Equality	F	DF	P
of Variances	1.51	8, 8	0.2858

Como $0,2858 > 0,10$ no se rechaza la igualdad de varianzas.

b) Reemplace la línea de puntos en la hipótesis alternativa por lo que corresponda.

Como en el enunciado no se sugiere una hipótesis en alguna dirección (que recuerdan más los que sólo escuchan las palabras o bien los que escuchan palabras acompañadas por imágenes), lo que se desea es verificar si hay diferencias entre los grupos, si hay efecto en agregar imágenes en algún u otro sentido. Por tanto la hipótesis alternativa es que la diferencia entre las medias poblacionales difiere de cero.

Alternative Hyp: difference ... 0

c) Indique cuál es la diferencia registrada entre las medias de palabras recordadas por los dos grupos y a favor de quién es dicha diferencia. ¿Es esa diferencia estadísticamente significativa al 5%? ¿En qué parte de la salida se encuentra esta información? ¿Qué puede concluir con respecto al efecto de acompañar las palabras con imágenes sobre la recordación?

La diferencia es -1,5556 como indica la salida en la línea encabezada por "Difference". La misma se realizó según el orden en que fueron ingresadas las variables. El título de la salida PALABRAS by IMÁGENES indica que primero se ingresó la variable correspondiente a sólo palabras (sin imágenes) y la diferencia es media del grupo "Sin imagen": 5,7778 menos la media del grupo "Con imagen": 7,3333; por eso da negativa. En otras palabras, el grupo "Con imagen" puntuó en promedio 1,5556 por encima del otro grupo, por lo que la diferencia es a favor de éste, como lo muestra el hecho de que su media es mayor.

La parte de la tabla que contiene toda esa información es:

Two-Sample T Tests for PALABRAS by IMAGENES				
IMAGENES	Mean	N	SD	SE
Sin imagen	5.7778	9	2.5386	0.8462
Con imagen	7.3333	9	3.1225	1.0408
Difference	-1.5556			

La diferencia muestral observada no es estadísticamente significativa, como lo indica el hecho de que el valor p para la prueba de diferencia de medias en poblaciones de igual varianza es mayor que el nivel de significación. En efecto: valor $p = 0,2632 > 0,05$; por lo que no se rechaza la hipótesis nula de igualdad de medias a nivel poblacional.

La información sobre el valor p para decidir sobre la igualdad de medias está en el fragmento:

Assumption	T	DF	P
Equal Variances	-1.16	16	0.2632

Conclusión. La diferencia de 1,5556 a favor del grupo con imágenes es atribuida al azar y no a un efecto de la presencia de imágenes acompañando las palabras. No hay evidencia de que la presencia de imágenes acompañando el listado de palabras produzca algún efecto sobre la recordación.

8.- Se aplicó un programa de entrenamiento psicoprofiláctico sobre un grupo de 8 dirigentes. Se analizaron autovaloraciones e imágenes que sobre ellos tenían sus subordinados a través de una técnica (Bachmann y Strumper, 1982) acerca de la “Aceptación del interlocutor” compuesta por ítems positivos a los que el encuestado debe responder en una escala de 1 a 3 (1: No es cierto, 2: Algo cierto, 3: Absolutamente cierto). Se comparó al grupo antes mencionado con otro grupo de 10 dirigentes que no recibieron entrenamiento. Una de las variables que se obtiene con este instrumento refleja la imagen del dirigente mediante un puntaje X obtenido promediando puntajes de encuestas que sobre cada dirigente hicieron sus subordinados, entendiendo que a mayor puntaje corresponde una mayor valoración del mismo. Luego de la aplicación del programa se obtuvieron los siguientes resultados.

Grupo de dirigentes entrenados: 56 57 52 61 72 74 72 66

Grupo de dirigentes no entrenados: 38 48 36 42 57 46 60 35 55 59

a) ¿Son estos resultados congruentes con la hipótesis que el entrenamiento contribuirá a que los sujetos que lo reciben sean mejor valorados por sus subordinados? Usar un nivel de significación del 5%.

Paso 1

X_1 = Puntaje que asignarían los subordinados a los dirigentes que recibieran el entrenamiento.

X_2 = Puntaje que asignarían los subordinados a los dirigentes que no recibieran el entrenamiento.

Supuestos: X_1 y X_2 son independientes y normalmente distribuidas.

Observación: Tanto los dirigentes entrenados como los subordinados encuestados constituyen muestras de sus correspondientes poblaciones hipotéticas.

Paso 2

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 > 0$

Resolveremos con Statistix

Statistix 8.0

28/06/2020, 17:38:02

Two-Sample T Tests for PUNTAJE by GRUPO

GRUPO	Mean	N	SD	SE
DE	63.750	8	8.4304	2.9806
DNE	47.600	10	9.6977	3.0667
Difference	16.150			

Null Hypothesis: difference = 0

Alternative Hyp: difference > 0

Assumption	T	DF	P	95% CI for Difference	
				Lower	Upper
Equal Variances	3.71	16	0.0009	6.9342	25.366
Unequal Variances	3.78	15.8	0.0008	7.0773	25.223

Test for Equality of Variances	F	DF	P
	1.32	9,7	0.3639

Cases Included 18 Missing Cases 0

Como $0,3639 > 0,10$ no se rechaza la hipótesis de igualdad de varianzas y para decidir sobre las medias se considera el valor p de la línea "Equal Variances". Como dicho valor $p = 0,0009 < 0,05$ se rechaza la hipótesis nula de igualdad de medias y se concluye que el puntaje medio de los subordinados de los dirigentes entrenados es mayor que el que asignarían los subordinados a los dirigentes no entrenados. Por tanto, efectivamente estos resultados son congruentes con la hipótesis de que el entrenamiento contribuirá a que los sujetos que lo reciban sean mejor valorados por sus subordinados. (No sólo al 5% sino a niveles mucho más bajos como por ej. 0,001 en vistas de cómo resultó el valor p)

b) ¿Qué supuestos deberían cumplirse para que el proceso de prueba de hipótesis sea válido?

Los supuestos son que las variables X_1 y X_2 son independientes y normalmente distribuidas.

9.- Un encuesta fue realizada en base a una muestra aleatoria de 400 trabajadores del sector transportes a los que se los indagó acerca de dos variables: presencia o ausencia de Acoso Psicológico (detectado al superar cierto umbral en el Cuestionario de Acoso Psicológico) y la variable Tipo de Contrato alcanzado por el trabajador (eventual, a plazo fijo, planta permanente). A continuación se presenta la distribución conjunta de estas dos variables en la muestra, así como los resultados del test de independencia efectuado con el programa Statistix.

Chi-Square Test for Heterogeneity or Independence

Acoso Psicológico		Tipo de Contrato			
		Eventual	Plazo Fijo	Permanente	
Presencia	Observed	92	47	37	176
	Expected	80.52	44.88	50.60	
	Cell Chi-Sq	1.64	0.10	3.66	
Ausencia	Observed	91	55	78	224
	Expected	102.48	57.12	64.40	
	Cell Chi-Sq	1.29	0.08	2.87	
		183	102	115	400
Overall Chi-Square		9.63			
P-Value		0.0081			
Degrees of Freedom		2			
Cases Included		6			
Missing Cases		0			

a) ¿Cuál es la hipótesis nula de interés?

La hipótesis nula de interés es que no hay asociación entre el tipo de contrato y la presencia de acoso psicológico.

b) Si H_0 fuera verdadera, cuántos trabajadores con contrato eventual piensa Usted que habrían sido clasificados con presencia de acoso psicológico?

Si las variables fueran independientes, cabría esperar la misma proporción dentro de este grupo de contrato eventual como en cualquier otro o en la muestra general. En la muestra general dicha proporción es $176/400$. Luego, llamando f_{11}^e a la cantidad de trabajadores con contrato eventual y acoso psicológico esperada bajo la hipótesis de independencia, debería ocurrir que $f_{11}^e/183 = 176/400$. Despejando se obtiene $f_{11}^e = 183 \times 176/400 = 80,52 \approx 81$.

Cabría esperar 81 trabajadores con presencia de acoso psicológico en el grupo de los trabajadores eventuales si las variables fueran independientes.

c) Si H_0 fuera verdadera, cuántos trabajadores con contrato de planta permanente piensa Usted que no hubieran presentado características de acoso psicológico?

Siguiendo el mismo razonamiento, $f_{23}^e/115 = 224/400 \Rightarrow f_{23}^e = 115 \times 224/400 = 64,4 \approx 64$.

d) Ponga a prueba la hipótesis nula del inciso a) con un nivel de significación del 1%.

Como el valor $p = 0,0081 < 0,01$, se rechaza la hipótesis de independencia. Las variables están asociadas. Según se calculó en b) la cantidad de trabajadores con presencia de acoso psicológico (92) fue superior a la esperada (81) si las variables fueran independientes. Por otra parte, por lo visto en c) la cantidad de trabajadores de planta permanente sin acoso psicológico fue mayor a la que cabría esperar si fueran independientes. Por lo cual se observa que el tipo de contrato afecta, cuanto más estable, mayor bienestar psicológico.

10.- Ubique el artículo de Martín-Ramos, R. et al (2015) en el material de la Cátedra o en la web: <https://www.medigraphic.com/pdfs/arcneu/ane-2015/ane154a.pdf>. Lea los párrafos de Objetivo (p.234), *Instrumentos de evaluación psicológica* (p.235), Tabla 2 (p.235) y Tabla 4 (p. 236).

a) ¿De cuántas pruebas de hipótesis son los resultados de la Tabla 2, qué tipo de pruebas son y qué objetivo tienen?

Los resultados de la tabla 2 corresponden a 11 pruebas de hipótesis. Son pruebas de hipótesis para comparación de medias de dos muestras independientes (pacientes con esclerosis múltiple y grupo control). El objetivo que tienen es analizar la influencia de la ansiedad y la depresión en las funciones ejecutivas de los pacientes con esclerosis múltiple. Para ello se utilizaron distintos instrumentos de evaluación neuropsicológica y mediante las distintas pruebas de hipótesis se analizó la existencia de diferencias significativas entre la media del grupo con esclerosis múltiple y la media del grupo control para cada una de las variables analizadas.

b) ¿En cuántas de ellas se obtuvieron resultados estadísticamente significativos? ¿A qué conclusión se llega en cada caso?

En todas las pruebas de hipótesis se obtuvieron resultados estadísticamente significativos, demostrando que el grupo de pacientes con esclerosis múltiple rindieron significativamente peor que el grupo de sujetos sanos (grupo control) en la ejecución de todos los test cognitivos aplicados. Entre ellos: velocidad en el procesamiento de la información y memoria de trabajo, control inhibitorio, flexibilidad cognitiva y razonamiento abstracto, atención alternante, solución de problemas prácticos constructivos y fluencia verbal.

c) Sabiendo que los grupos control sano y con esclerosis múltiple estaban conformados por 30 personas cada uno, halle el tamaño del efecto para las pruebas de Diseño con cubos (Waiss-III), Fluidez verbal fonológica (COWA) e Inventario de depresión de Beck (IDB). Expresar una conclusión en cada caso.

Para cada caso hay que reemplazar los datos en la fórmula:

$$g = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}}$$

$$n_1 = n_2 = 30$$

	$\bar{x}_1$	$\bar{x}_2$	s_1	s_2	Numerador	Denominador	g	
WAISS-III	30,83	45,6	15,448	12,159	14,77	13,901	1,06	GRANDE
COWA-F	25,73	53,1	10,336	5,536	27,37	8,291	3,30	MUY GRANDE
IDB	20,43	6,17	5,418	1,895	14,26	4,059	3,51	MUY GRANDE

La condición de esclerosis múltiple tiene afecta de manera importante disminuyendo las habilidades cognitivas que miden los tres test en cuestión.

d) Según la Tabla 3 ¿Qué correlaciones resultaron significativas?

Resultaron significativas las correlaciones, en los pacientes con esclerosis múltiple, de la variable ansiedad rasgo con los puntajes en los tests: de colores y palabras de Stroop, diseño con cubos, fluidez verbal fonológica y fluidez verbal semántica. También fueron significativas las correlaciones entre los

puntajes en el IDB con los obtenidos con los tests de Diseño con cubos y fluidez verbal semántica. Todas estas correlaciones fueron negativas, es decir que a medida que aumentan los puntajes de una variable disminuyen los puntajes en la otra (y viceversa). Hubo una correlación significativa y positiva el puntaje en IDB y los errores perseverativos. Dicho sintéticamente, la ansiedad correlaciona negativamente con ciertas habilidades cognitivas y positivamente con cierto tipo de errores.

e) ¿Cómo se relacionan estas correlaciones con los resultados de la Tabla 4?

La tabla 4 muestra el análisis de regresión lineal de la ansiedad (como estado y como rasgo) y la depresión con los componentes de las funciones ejecutivas en pacientes con esclerosis múltiple. Comparando los resultados de la tabla 3 y la tabla 4, puede observarse como las mismas correlaciones que resultaban significativas estadísticamente al analizar el r de Pearson, lo continúan siendo cuando se analiza la regresión lineal y su bondad de ajuste con R^2 .

Entre los mismos pares de variables donde se estudió la correlación (tabla 3), ahora se calcula la recta de regresión. La bondad de ajuste de tales rectas está expresada por el coeficiente de determinación R^2 que es el cuadrado del coeficiente de correlación r . Hay un valor p que, en cada caso, expresa si el modelo lineal elegido es significativo; es decir, si explica algo de la variabilidad de la ansiedad / depresión, más allá de lo que se obtiene muestralmente. De ahí que, en los pares de variables donde los valores de correlación eran significativos (tabla 3); es decir, donde había cierta relación lineal, también la regresión resulta significativa.

f) Considere los resultados de los análisis de regresión de la Tabla 4. ¿Cuáles resultaron estadísticamente significativos? En esos casos, ¿cuál fue la proporción de varianza explicada?

Los resultados expuestos en la tabla 4 muestran que la variación en los puntajes del Test de colores y palabras de Stroop, diseño con cubos, fluidez verbal fonológica y fluidez verbal semántica, son explicados por la relación lineal que dichas variables tienen con la variable ansiedad rasgo y resultaron estadísticamente significativos al 5% (en tanto el p value es $< 0,05$).

Por otro lado, la variabilidad en los puntajes de los errores perseverativos (Test de clasificación de tarjetas de Wisconsin), en el diseño con cubos y en la fluidez verbal semántica, son explicados por su relación con los puntajes en depresión y resultaron también estadísticamente significativos.

En cuanto a la proporción de varianza explicada, a continuación se informa el porcentaje habiendo multiplicado por 100 el coeficiente de determinación R^2 .

Test	Ansiedad rasgo (IDARE)	Depresión (BDI)
Test de colores y palabras de Stroop	14,8%	
Test de clasificación de tarjetas (errores perseverativos)		15,8%
Diseño con cubos	13,4%	29,9%
Fluidez verbal fonológica	17,1%	
Fluidez verbal semántica	17,1%	24,4%

11.- Ubique el artículo de Huertas, E. et al (2019) en el material de la Cátedra o en la web: <http://www.psicothema.com/pdf/4538.pdf>. Lea el resumen y los párrafos de *Participants* (pág.240) y *Genetic data* (pág. 241)

a) i) ¿Se realizó algún tipo de muestreo aleatorio?

No, no se realizó muestreo aleatorio, los estudiantes seleccionados fueron voluntarios que se propusieron para hacer el estudio.

ii) ¿Cuál es la población considerada?

La población está compuesta por los estudiantes de pregrado de la Universidad Complutense con antepasados europeos y disposición a participar a cambio de créditos en este tipo de estudios.

b) Al describir los resultados de las pruebas de asociación se mencionan algunos valores del estadístico χ^2 y sus correspondientes valores p. Utilizando los datos de *Table 1* verifique ambas cosas en las pruebas de asociación entre el genotipo y el sexo. ¿Qué se concluyó al respecto?

Se pide realizar el test de independencia cuando se cruza 1) el Sexo con el Genotipo C385A y 2) el Sexo con el Genotipo rs12075550. Para ello hay que considerar sendas tablas de contingencia de 3x2, a saber: Sexo x C385A y Sexo x rs12075550. Se deben volcar las frecuencias pertinentes (*n*) de la tabla 1 en estas dos tablas de contingencia y hacer el test de Independencia. Utilizaremos Statistix.

Chi-Square Test for Heterogeneity or Independence

sexo by C385A

sexo		C385A			
		AA	AC	CC	
male	Observed	12	57	138	207
	Expected	8.41	63.47	135.11	
	Cell Chi-Sq	1.53	0.66	0.06	
female	Obs	23	207	424	654
	Expected	26.59	200.53	426.89	
	Cell Chi-Sq	0.48	0.21	0.02	
		35	264	562	861

Overall Chi-Square 2.96

P-Value 0.2276

Degrees of Freedom 2

Cases Included 861 Missing Cases 0

Chi-Square Test for Heterogeneity or Independence

sexo by rs12075550

Case		Variable			
		CC	CT	TT	
male	Observed	43	95	69	207
	Expected	41.59	96.89	68.52	
	Cell Chi-Sq	0.05	0.04	0.00	
female	Observed	130	308	216	654
	Expected	131.41	306.11	216.48	
	Cell Chi-Sq	0.02	0.01	0.00	
		173	403	285	861

Overall Chi-Square 0.12
P-Value 0.9438
Degrees of Freedom 2

Cases Included 6 Missing Cases 0

En ambos casos se concluyó que la distribución de los genotipos en cuestión no está asociada al sexo; textualmente: "The genotype distributions of both SNPs did not deviate from Hardy-Weinberg equilibrium (FAAH C385A $\chi^2 = 0.32$, $p = 0.57$, rs12075550 $\chi^2 = 1.96$, $p = 0.16$) ***neither were there significant differences between genotypes in the gender distribution (C385A $\chi^2 = 2.96$, $p = 0.23$, rs12075550 $\chi^2 = 0.12$, $p = 0.94$).***"

En los siguientes ítems elija la única opción correcta.

12.- En una distribución binomial de parámetros n y p , la media y la mediana coinciden cuando

- a. $n \geq 30$.
- b. la media es np .
- ☒ c. $p = 0,5$.

Explicación. La distribución binomial es simétrica cuando $p = 0,5$; por tanto, en ese caso, coinciden media y mediana. Cuando n es impar, también coincide con la moda. Cuando n es par, los dos valores centrales son la moda y, el promedio de los mismos, también coincide con la media y la mediana.

13.- Se cree que el 60% de los enfermeros de cierto hospital tiene la opinión de que se deben hacer modificaciones substanciales en el protocolo sanitario. Si así fuera, ¿cuál es la probabilidad de que, entre 16 enfermeros elegidos al azar de dicho hospital, más del 75% opine que hay que hacer tales modificaciones?

- a. 0,10142
- b. 0,16657
- ☒ c. 0,06515

Explicación. La variable X = "Cantidad de enfermeros entre 16 encuestados que opina que hay que hacer modificaciones substanciales en el protocolo sanitario" es una variable binomial de parámetros $n = 16$, $p = 0,6$. Más del 75% de la muestra, significa más de 12, es decir a partir de 13; por lo que hay que calcular $P(X \geq 13)$.

14.- El puntaje en un test de depresión se distribuye normalmente en cierta población. Un psicólogo atiende problemas de depresión moderados pero deriva los casos que considera muy severos a otro especialista. Considera severos los casos de puntuaciones mayores a la media en más de una desviación estándar. ¿Cuál es la probabilidad de que tenga que derivar a lo sumo 3 pacientes entre los próximos 6 que le solicitan tratamiento?

- a. 0,05482
- ☒ b. 0,99275
- c. 0,04757

Explicación. La probabilidad de que una persona tenga un puntaje de depresión superior a una desviación estándar por encima de la media, es un puntaje $Z > 1$ en una distribución normal estándar. Dicha probabilidad es 0,15866. La variable X = "Cantidad de pacientes por derivar entre los próximos 6 que solicitan tratamiento" puede considerarse Binomial de parámetros $n = 6$ y $p = 0,15866$. Por tanto hay que calcular $P(X \leq 3)$.

15.- Para estudiar el poder de discriminación de un nuevo test de ansiedad quieren compararse los puntajes que obtienen en dicho test dos grupos que están en los polos de la distribución de los puntajes de ansiedad que se obtienen con el inventario STAI; para ello se consideran los individuos que están en la primera y en la última intercuartila. Suponiendo que el puntaje en STAI se distribuye normalmente con media 25 y desviación estándar 3, los límites (redondeados) de dichas intercuartilas son:

- ☒ a. 23 y 27
- b. 25 y 75
- c. 22 y 28

Explicación. Hay que hallar los cuartiles 1 y 3 utilizando la normal inversa para las probabilidades a izquierda y derecha de 0,25.

16.- Un test de evaluación cognitiva se distribuye normalmente con media 27 y desviación estándar 4. Se considera que manifiestan problemas cognitivos puntajes menores o iguales a 21. ¿Qué cantidad de personas con problemas cognitivos cabe esperar en una población de 10000 personas?

- a. 1336
- b. 9332
- ☒ c. 668

Explicación. La probabilidad se interpreta como una proporción, la frecuencia relativa en el modelo; por lo que la cantidad esperada es la probabilidad por 10000. Luego, hay que buscar la probabilidad correspondiente a valores inferiores o iguales a 21 en la distribución normal de media 27 y desviación estándar 4. Dicha probabilidad es 0,06681, multiplicarla por 10000 y redondearla.

17.- Para llevar a cabo los métodos de la inferencia estadística es necesario conocer la distribución de probabilidades de los estadísticos para

- a. calcular el verdadero valor de los parámetros.
- ☒ b. controlar los errores en las pruebas de hipótesis.
- c. obtener el error estándar de los estimadores.

Explicación. Para determinar la zona de rechazo de H_0 se necesita conocer la distribución del estadístico de prueba; dicha zona de rechazo se halla prefijando baja la probabilidad de error de tipo I. La distribución del estadístico permite también calcular la potencia del test en diversos casos.

18.- Los puntajes en una escala de distracción al realizar tareas escolares on-line se distribuyen normalmente con media 12 en un grupo normativo. Puntajes superiores representan mayores niveles de distracción. Una maestra ensayó un programa para mejorar la atención en este tipo de tareas, en una muestra piloto de 9 estudiantes. Finalizado el programa se les administró el test de distracción. Se obtuvo una desviación estándar de 4 puntos. Para decidir sobre la eficacia del programa se hará la correspondiente prueba de hipótesis con un nivel de significación del 5%. Se concluirá que el programa es eficaz sólo si en la muestra se observara una media

- ☒ a. inferior a 9,52.
- b. superior a 14,48.
- c. inferior a 9,81.

Explicación. Como el programa se propone lograr mejor atención, los puntajes en distracción deberían disminuir; por tanto corresponde un test unilateral a izquierda para la media. Dada la hipótesis de normalidad y que la desviación estándar es la de la muestra, se trata de un test t de Student con 8 grados de libertad. El valor que proporciona la aplicación es $t_{8;0,05} = -1,85955$. Se rechaza H_0 si y sólo si

$$\frac{\bar{X}-12}{4/\sqrt{9}} < -1,85955 \Leftrightarrow \bar{X} < 12 - 1,85955 \cdot \frac{4}{3} = 9,5206$$

19.- En cierta escala de amabilidad los puntajes se distribuyen normalmente con una media de 46 puntos y una desviación estándar de 8 puntos en el grupo normativo. Se consideran extremadamente amables las personas cuyo puntaje supera los 56 puntos. Se administró la escala en una localidad para corroborar, mediante una prueba de hipótesis, si la proporción de gente extremadamente amable coincide con la del grupo normativo. En una muestra de 200 personas, 28 resultaron extremadamente amables. El valor p es:

- a. 0,0573
- ☒ b. 0,1146
- c. 0,1057

Explicación. La proporción de gente extremadamente amable en el grupo normativo es la probabilidad de los valores mayores a 56 en la distribución normal con media 46 y desviación estándar 8. Se busca ese valor en la aplicación, el cual es 0,1057. Para corroborar si en la localidad se mantiene esta proporción se hace el correspondiente test de hipótesis bilateral donde $H_0: p = 0,1057$. La proporción muestral es $28/200 = 0,14$.

Bajo H_0 el error estándar del estimador de p es $\sqrt{\frac{0,1057 \times (1-0,1057)}{200}} \approx 0,02174$. Siendo una prueba bilateral, el valor p es el doble de la probabilidad a la derecha de 0,14 en una distribución normal con media 0,1057 y desviación estándar aproximada a 0,02174. La probabilidad a derecha de 14 es 0,0573 y, el doble de la misma es el valor $p = 0,1146$. (Puede haber diferencias en el cuarto decimal según la precisión con la que se hayan hecho los redondeos).

20.- Dos proyectos de investigación compiten por un subsidio en un concurso. Los puntajes en una escala de 1 a 5 obtenidos por cada uno de ellos a partir de la opinión de 7 jueces son los siguientes:

Juez	1	2	3	4	5	6	7
Proyecto A	4	5	4	4	3	5	3
Proyecto B	2	4	4	5	3	4	1

Considérense a estos jueces como una muestra representativa de una población de expertos en el tema. Para decidir si los puntajes asignados a los proyectos difieren significativamente en promedio se debe realizar una prueba de hipótesis para

- a. la diferencia de medias de muestras independientes.
- b. la homogeneidad de varianzas de muestras independientes.
- ☒ c. a media de la diferencia de datos pareados.

Explicación. Cada juez evalúa dos proyectos; por eso los datos son pareados. Se quiere decidir si los puntajes difieren en la media; por eso corresponde el test de la diferencia de medias de datos pareados.

21.- En una prueba de hipótesis la potencia disminuye cuanto

- a. menor es el nivel de significación.
- b. mayor es el tamaño de muestra.
- c. menor es el valor p.

Explicación. El nivel de significación es la probabilidad de cometer error de tipo I: Rechazar H_0 cuando es verdadera. Cuanto menor es el nivel de significación, más evidencia se pide para rechazar H_0 , por lo cual se la rechaza menos con el riesgo de mantenerla aun cuando es falsa, lo cual significa que disminuye la potencia. Cuando mayor es el tamaño de muestra la potencia es mayor porque con más información, el test de hipótesis es más sensible a captar las diferencias con respecto a lo que postula H_0 . El valor p depende del valor observado del estadístico de prueba mientras que la potencia depende del valor del parámetro bajo la hipótesis alternativa y puede ser conocida antes de tomar los datos.

22.- La cantidad promedio de errores gramaticales en una muestra de niños que recibieron capacitación específica en lengua resultó significativamente inferior a la cantidad promedio de dichos errores en la población objetivo sin la capacitación. Se utilizó un nivel de significación del 5%. De ello se concluye, con una probabilidad de error de 0,05, que:

- a. La capacitación tiene el efecto de disminuir el promedio de los errores en una medida considerada muy importante.
- b. La muestra de niños resultó sesgada, por lo que se obtuvieron puntajes inferiores a los que podría esperarse.
- c. la media de errores en una población de niños capacitados sería inferior a la de la población sin la capacitación.

Explicación. En una prueba de hipótesis, que la media muestral es significativamente menor a la postulada en H_0 lleva a concluir que la media de la población de la que se sacó la muestra (en este caso la población hipotética de niños capacitados) es inferior a la postulada en H_0 , que corresponde a la media de la población de niños sin la capacitación. Esta afirmación está sujeta a una probabilidad de error dada por el nivel de significación. La significación estadística no indica nada sobre la magnitud de tal diferencia entre los parámetros (la media de toda la población de niños capacitados con la media de la población de todos los niños sin la capacitación). Para aseverar algo así como que el efecto de la capacitación es importante habría que calcular el tamaño del efecto.