DISEÑOS DE INVESTIGACIÓN Y ANÁLISIS DE DATOS **Enero/febrero 2014. Código asignatura: 62012054** **MODELO A DURACION: 2 HORAS**
Material: **Formulario, Tablas y calculadora no programable** **Calificación= (0,4 x Aciertos) - (0,2 x Errores)**
No debe entregar los enunciados

KIBBUTZ

SITUACIÓN 1. Un psicólogo educacional interesado en estudiar el estilo educativo utilizado por los padres, dispone de una muestra de 200 familias, observando que 40 de ellas aplican un estilo autoritario que prioriza el cumplimiento de las normas, 100 emplean un estilo equilibrado, que a pesar de mantener valores altos en exigencias y control, adapta las estrategias educativas a cada situación concreta, y el resto de las familias utilizan un estilo permisivo, con poca consistencia en las normas y un elevado grado de sobreprotección. Nivel de confianza 95%.

1- El error típico para la proporción de familias que utilizan un estilo permisivo, es: A) 0,05; **B) 0,0324**; C) 0,01.

La proporción de familias que utilizan un sistema permisivo es: $p = \frac{60}{200} = 0{,}30$*, luego:*

$$\sigma_P = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0{,}30 \cdot 0{,}70}{200}} = 0{,}0324$$

2- A partir de los datos obtenidos en la muestra, la proporción poblacional de familias que utilizan un sistema permisivo es un valor comprendido entre: A) 0,2164; 0,3836; **B) 0,2365; 0,3635**; C) 0,1934; 0,4066

$$p \pm E_{max} \rightarrow 0{,}30 \pm 1{,}96 \cdot 0{,}0324 \rightarrow \begin{cases} 0{,}3635 \\ 0{,}2365 \end{cases}$$

3- Si el psicólogo quiere comprobar que la proporción de familias que utilizan un sistema permisivo es distinta de 0,25, ¿Cuál es la hipótesis nula? A) $\pi_{Permisivo} \leq 0{,}25$; B) $\pi_{Permisivo} \geq 0{,}25$; **C)** $\boldsymbol{\pi_{Permisivo} = 0{,}25}$

Las hipótesis que se han de plantear son:

$$H_0: \pi_{Permisivo} = 0{,}25$$
$$H_1: \pi_{Permisivo} \neq 0{,}25$$

4- El psicólogo quiere comprobar que la proporción de familias que utilizan un sistema permisivo es distinta de 0,25. La conclusión a la que llega es: **A) mantener la hipótesis nula**; B) rechazar la hipótesis nula; C) rechazar la hipótesis nula porque la proporción en la muestra se aleja mucho de la proporción postulada por la hipótesis nula.

Para comprobar esta hipótesis el estadístico de contraste es:

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0{,}30 - 0{,}25}{\sqrt{\frac{0{,}25 \cdot 0{,}75}{200}}} = 1{,}63$$

Dado que el estadístico de contrate se encuentra entre los valores críticos, $-1{,}96 < 1{,}63 < +1{,}96$*, se mantiene la hipótesis nula.*

5- El psicólogo quiere comprobar que la proporción de familias que utilizan un sistema equilibrado es distinta de 0,5. Con los datos de la muestra, la conclusión a la que llega es: A) rechazar la hipótesis nula; **B) mantener la hipótesis nula**; C) no puede llegar a una conclusión, porque el estadístico de contraste es igual a cero.

Las hipótesis a plantear son:

$$H_o: \pi_{Equilibrado} = 0{,}5$$
$$H_1: \pi_{Equilibrado} \neq 0{,}5$$

La proporción de familias que utilizan un sistema equilibrado es: $p = \frac{100}{200} = 0{,}50$*, luego no sería ni siquiera necesario realizar el contraste de hipótesis, dado que la proporción de la muestra es igual que la hipótesis nula. Efectivamente:*

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0{,}5 - 0{,}5}{\sqrt{\frac{0{,}5 \cdot 0{,}5}{200}}} = 0$$

6- Si el psicólogo quiere comprobar que la proporción de familias que utilizan un sistema equilibrado es distinta de 0,5 ¿Cuál es el nivel crítico?: **A) uno**; B) cero; C) no se puede calcular.

La distribución normal tipificada tiene por media cero, luego la probabilidad de obtener un valor más extremo es 0,5, y como el contraste es bilateral, multiplicamos por dos esta probabilidad.

7- Al aumentar el tamaño de la muestra, el intervalo de confianza: A) permanece inalterado; B) es más ancho; **C) es más estrecho**.

8- Suponiendo que H_0 es cierta, ¿Cuál de las siguientes probabilidades está asociada al estadístico de contraste?: A) la probabilidad de rechazarla; B) la probabilidad de encontrar valores más extremos que el valor crítico**; C) la probabilidad de encontrar valores iguales a los obtenidos en la muestra, o más extremos**.

SITUACIÓN 2. "*El Programa Internacional para la Evaluación de la Competencia de los Adultos del 2013 (PIAAC, por su sigla en Inglés) sitúa a España en los últimos puestos entre los países de la OCDE, solo por encima de Italia, en comprensión lectora. Los españoles examinados de entre 15 y 65 años obtuvieron en comprensión lectora una media igual a 252 puntos, 21 por debajo de la media de la OCDE y 19 por debajo de la media de la UE. Entre sus conclusiones destaca la estrecha relación entre los niveles de comprensión lectora de los adultos y sus posibilidades de encontrar trabajo".* Imagine que, ante estos resultados, usted desea corroborar si el nivel de comprensión lectora de los jóvenes con empleo es significativamente mayor que la de los jóvenes desempleados. Para ello, selecciona una muestra de 182 jóvenes de su localidad con

KIBBUTZ

edades comprendidas entre 25 y 34 años, de los que el 33,5% están en situación de desempleo, y les aplica la misma prueba. La media y desviación típica insesgada de los desempleados es de 251 y 115, respectivamente, frente a 289 y 92 de los que tienen empleo. Establezca un nivel de significación de 0,05.

Los datos del problema son:

	España	*OCDE*	*UE*
Media	*252*	*252 + 21 = 273*	*252 + 19 = 271*

Para la muestra en la localidad del investigador:

	Media	*Cuasidesviación típica*	*n*
Con empleo	289	92	$182 \cdot 0{,}665 \approx 121$
Desempleados	*251*	*115*	$182 \cdot 0{,}335 \approx 61$

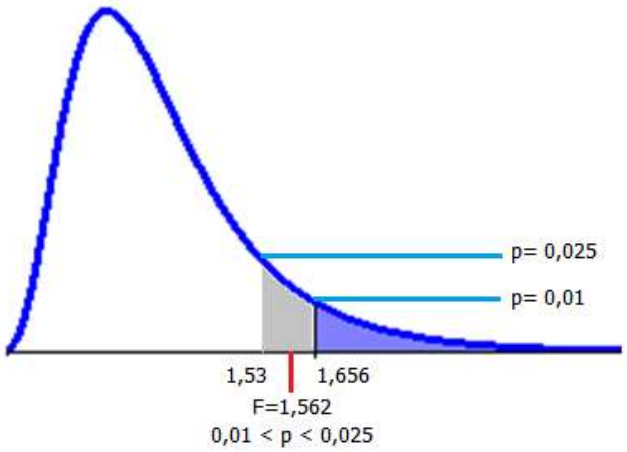
9- En función de sus objetivos, ¿cuál de las siguientes hipótesis está correctamente formulada?: A) $H_0: \mu_{Empl} \geq \mu_{Desemp}$; B) $H_1: \mu_{Empl} < \mu_{Desemp}$; C) $\boldsymbol{H_1: \mu_{Empl} > \mu_{Desemp}}$

10- El estadístico de contraste para comprobar si las dos muestras provienen de poblaciones con la misma varianza es: **A) 1,562**; B) 1,391; C) 2,031.

$$F = \frac{\hat{S}^2_{Desempleados}}{\hat{S}^2_{Con\ empleo}} = \frac{115^2}{92^2} = 1{,}562$$

11- El nivel crítico que corresponde al estadístico de contraste sobre la igualdad de las varianzas poblacionales es: A) p < 0,01; B) p < 0,0025; **C) p < 0,05**.

En la distribución F con $61 - 1 = 60$ *y* $121 - 1 = 120$ *grados de libertad figuran los valores 1,53 que deja por debajo una probabilidad de 0,975 (y por encima de 0,025) y el valor 1,656 que deja por debajo una probabilidad de 0,99 y por encima de 0,01.*



KIBBUTZ

Dado que el contraste es bilateral, hay que multiplicar por dos las probabilidades obtenidas en las tablas, luego, finalmente:

$$0{,}02 < p < 0{,}05$$

12- El análisis de datos que utilizará en su estudio es el que corresponde a: A) una muestra con varianza poblacional desconocida; B) dos muestras independientes extraídas de dos poblaciones con la misma varianza; **C) dos muestras independientes extraídas de dos poblaciones con varianzas distintas**.

13- El error típico de la distribución muestral de la diferencia de medias es: A) 17,8; B) 15,74; **C) 16,94**.

Habiendo rechazado la H_0 de igualdad de varianzas, el error típico de la distribución muestral, es:

$$\sigma_{\bar{Y}_1-\bar{Y}_2} = \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}} = \sqrt{\frac{115^2}{61} + \frac{92^2}{121}} = 16{,}94$$

14- El valor absoluto del estadístico de contraste y el nivel crítico p para tomar una decisión respecto a su Ho, son, aproximadamente: A) 2,24 (p > 0,05); B) 2,41 (p < 0,005); **C) 2,24 (p < 0,025)**.

La distribución muestral de la diferencia de medias en muestras extraídas de poblaciones con varianzas distintas es la distribución t de Student:

$$T = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{Y}_1-\bar{Y}_2}} = \frac{(289 - 251) - 0}{16{,}94} = 2{,}24$$

Los grados de libertad, aplicando la corrección de Welch, son:

$$g.l. = \frac{\left(\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}\right)^2}{\frac{\left(\hat{S}_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(\hat{S}_2^2/n_2\right)^2}{n_2 - 1}} = \frac{\left(\frac{115^2}{61} + \frac{92^2}{121}\right)^2}{\frac{\left(115^2/61\right)^2}{61 - 1} + \frac{\left(92^2/121\right)^2}{121 - 1}} = 99{,}77 \approx 99$$

En la distribución t con 100 gl (los más próximos a 99), el estadístico $t = 2{,}24$ se encuentra entre los valores 1,984 y 2,364 que dejan por encima una probabilidad de 0,025 y 0,01, respectivamente. Por tanto, al valor $t = 2{,}24$ le corresponde un nivel mayor que 0,01 y menor que 0,025.

15- Si quiere contrastar la hipótesis de que la proporción de empleados entre los jóvenes de su localidad supera en más de quince puntos porcentuales a la proporción de desempleados, el estadístico de contraste es, en valor absoluto: **A) 2,43**; B) 4,22; C) 3,47.

$$\sigma_{p_1-p_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = \sqrt{\frac{0{,}335\,(1-0{,}335)}{61} + \frac{0{,}665\,(1-0{,}665)}{121}} = 0{,}074$$

$$Z = \frac{(p_1 - p_2) - D}{\sigma_{p_1-p_2}} = \frac{(0{,}665 - 0{,}335) - 0{,}15}{0{,}074} = 2{,}43$$

16- Si optase por utilizar muestras más grandes y suponiendo que se mantienen los mismos estadísticos obtenidos en estas muestras, ¿cuál de las siguientes consecuencias se producirían?: A) aumentaría el

KIBBUTZ

tamaño del efecto; **B) aumentaría la probabilidad de rechazar la hipótesis nula**; C) disminuiría la potencia del contraste.

SITUACIÓN 3. En un estudio de Psiconeuroinmunología, se investigó si suplementar la dieta con ácidos Omega-3 tenía efectos beneficiosos en las personas que ya disfrutaban de una dieta equilibrada (Daruna, J. H., 2004). Como se sospechaba que la edad podría afectar a los resultados, se escogieron dos grupos de adultos y dos grupos de personas mayores. En cada edad, la mitad de las personas sirvieron de control (no se les proporcionaba Omega-3) y la otra mitad tomaba una cápsula de aceite Omega-3 tres veces al día durante 2 meses. El total de sujetos fueron 64 y el diseño fue equilibrado. Se realizó un ANOVA con dos factores (Edad y Grupo) utilizando como variable dependiente la diferencia en la concentración en sangre de la proteína c-reactiva (CRP) antes y después de aplicar el tratamiento [$CRP_{antes} - CRP_{después}$] (también para el grupo control). Los resultados indicaron que sólo la interacción entre los factores fue significativa ($F(1; 60) = 7{,}302$, $MC_{S/AB} = 0{,}321$, $p < 0{,}05$) mientras que las Fs de Edad (1,802) y del Grupo (1,390) no lo fueron.

En primer lugar, completamos la tabla de ANOVA con los datos del enunciado (en rojo y negrita):

FV	SC	gl	MC	F
A (Edad)	0,578	**1**	0,578	**1,802**
B (Grupo)	0,446	**1**	0,446	**1,390**
AxB	2,344	**1**	2,344	**7,302**
Intra	19,26	**60**	**0,321**	
Total	22,629	**63**		

17- Como tenemos un diseño equilibrado, esto significa que: A) los niveles elegidos son los únicos que nos interesan; **B) el número de sujetos en cada combinación de Edad y Grupo es de 16**; C) los niveles elegidos se han escogido al azar de entre todos los posibles.

Son 64 personas las que forman parte del experimento, como tenemos cuatro combinaciones de tratamientos, en cada una de ellas hay 16 personas.

18- El número de razones básicas que tenemos que calcular en este diseño, para realizar la tabla de ANOVA, son: A) 2; B) 4; **C) 5**.

Las razones básicas que hay que calcular son: [A], [B], [AB], [Y] y [T]

19- Si asumimos que una mayor CRP indica un mejor sistema inmune, y que tomar Omega-3 aumenta la concentración de proteína c-reactiva entonces: A) no pueden producirse puntuaciones negativas en la variable dependiente; B) según H_0, las puntuaciones de la variable dependiente fluctuarán alrededor de la unidad; **C) para los sujetos que toman Omega-3, la mayor parte de las puntuaciones en la variable dependiente serán negativas.**

20- La H_1 para el factor Grupo (control *vs*. experimental) es: **A)** $\boldsymbol{H_1: \mu_{control} \neq \mu_{experimental}}$; B) $H_1: \mu_{control} = \mu_{experimental}$; C) $H_1: \mu_{adultos} \neq \mu_{mayores}$.

KIBBUTZ

21- Los grados de libertad de la F para el factor Edad serán: **A) (1, 60)**; B) (2, 60) C) 60.

22- El valor de la SC de los sujetos dentro de los tratamientos (S/AB o, lo que es lo mismo, del error) es (aproximadamente): **A) 19,26**; B) 2,34; C) 0,321.

23- El valor de la SC_{total} es aproximadamente: **A) 22,63**; B) 19,26; C) 41,89.

24- Como la interacción resultó significativa, se aplicó un test de efectos simples. Los resultados indicaron que sólo las personas mayores que tomaron Omega-3 obtuvieron puntuaciones menores en la variable dependiente respecto a los otros tres grupos. Esto significa que: A) sólo los jóvenes se beneficiaron de tomar Omega -3; **B) suplementar una dieta ya de por sí equilibrada es beneficioso para las personas mayores;** C) tomar Omega -3 no es beneficioso ni para jóvenes ni para mayores.

25- En esta situación deberíamos realizar el análisis de los efectos simples considerando: A) dos fuentes de variación; **B) cuatro fuentes de variación**; C) tres fuentes de variación.

XIBBUTZ