Compartido por Sara C.

**Modelo C. ENERO de 2018. No debe entregar los enunciados**
**Fórmula de corrección: Aciertos x 0,4 - Errores x 0,2**
**Material permitido**: Formulario sin anotaciones y cualquier tipo de calculadora en la que no se pueda introducir texto

**SITUACIÓN 1**. Ofman, Pereyra-Girardi, Cófreces y Stefani (2014) estudiaron las diferencias de género en el afrontamiento de la hipertensión arterial. Utilizaron para ello el Cuestionario sobre Modos de Afrontamiento (Ways of CopingQuestionnaire). En dicho cuestionario, entre otras estrategias de afrontamiento,se mide la resolución planificada de problemas (RPP), que consiste en desarrollar tácticas para enfrentarse al problema, intentando alterar la situación desde una planificación analítica. Cuanto más elevadas son las puntuaciones en RPP mayor es el uso de esta estrategia.Ofman et al. (2014) encontraron diferencias significativas en el uso de RPP en función del género, puntuando los varones más alto que las mujeres.

Imagine que usted dispone de una muestra de 200 personas con una edad media de 57 años y desviación típica de 8 años. La mitad de ellos son hombres, para los cuales la puntuación media en RPP es de 3,4 con una cuasidesviación típica de 0,8. El grupo de mujeres presenta la misma variabilidad que los varones y una puntuación media de 3,2. ($\alpha = 0{,}05$).

1 – El intervalo de confianza para la media aritmética de la edad poblacional en la muestra total, se encuentra, aproximadamente entre los valores:

**A) 55,89 y 58,11**
B) 55 y 60.
C) 54,81 y59,22

**Solución**:
Disponemos los datos proporcionados para facilitarnos la resolución del problema:

| Muestra total | Varones | Mujeres |
|---|---|---|
| N = 200 | N = 100 | N=100 |
| $\bar{X}_{Edad} = 57$ | $\overline{RPP} = 3{,}4$ | $\overline{RPP} = 3{,}2$ |
| $S_{Edad} = 8$ | $\hat{S}_{RPP} = 0{,}8$ | $\hat{S}_{RPP} = 0{,}8$ |

Aplicamos la fórmula del IC para el valor de la media de edad de la muestra total:

$$\bar{Y} \pm t_{n-1;1-\alpha/2}\left(\frac{S_n}{\sqrt{n-1}}\right)$$

$$57 \pm 1{,}96 \cdot \frac{8}{\sqrt{200-1}} \rightarrow 57 \pm 1{,}11 \rightarrow \begin{cases} 58{,}11 \\ 55{,}89 \end{cases}$$

2 – Al contrastar si las varianzas poblacionales entre hombres y mujeres son iguales para la variable RPP, se obtienen un estadístico de contraste y un nivel crítico, aproximadamente:

A) $F = 1{,}352\ (p = 0{,}05)$
**B)** $\boldsymbol{F = 1\ (p > 0{,}20)}$
C) $F = 1\ (p < 0{,}20)$

**Solución**:

Como nos han indicado que las cuasidesviaciones típicas son idénticas entre ambas muestras (hombres y mujeres), el estadístico F sería igual a la unidad.

$$F = \frac{\hat{S}_1^2}{\hat{S}_2^2} = \frac{0{,}8^2}{0{,}8^2} = 1$$

Para determinar qué nivel crítico es apropiado, debemos considerar que los grados de libertad son 99 y 99 (aunque tendremos que aproximar a los g.l. más aproximados, en este caso, 120). Luego consultando en las tablas de la F para diversos valores de P encontramos:

| | | |
|---|---|---|
| Tabla 0,90 (p = 0,10) | ➔ | 1,265 |
| Tabla 0,95 (p = 0,05) | ➔ | 1,352 |
| Tabla 0,975 (p = 0,025) | ➔ | 1,433 |
| Tabla 0,99 (p = 0,01) | ➔ | 1,533 |

Observamos que una F = 1,265 deja por encima de sí una p = 0,10, luego una F = 1, al ser inferior a 1,265 deja por encima de sí un valor superior a 0,10. Como en el caso del contraste de varianzas debemos utilizar un contraste bilateral tenemos que multiplicarlo por 2 (0,10 x 2 = 0,20). La única opción correcta es, entonces, la B.

3 – En función de los resultados de la pregunta anterior, se puede asumir que:

A) Las varianzas poblacionales para RPP son diferentes para hombres y mujeres.
**B) Las varianzas poblacionales para RPP son iguales para hombres y mujeres.**
C) Las varianzas poblaciones para RPP son normales para hombres y mujeres.

Solución:

tal y como se indica directamente en el enunciado y el resultado inferencial del estadístico F obtenido en la pregunta anterior, la opción correcta es la B. La opción C es irrelevante ya que se refiere al tipo de distribución de la variable RPP.

4 – Para contrastar si los resultados de su muestra replican los resultados de Ofman et al. (2014), la hipótesis nula que se debe plantear es:

A) $H_0: \mu_{Hombres} = \mu_{Mujeres}$
B) $H_0: \mu_{Hombres} \geq \mu_{Mujeres}$
**C)** $\boldsymbol{H_0: \mu_{Hombres} \leq \mu_{Mujeres}}$

**Solución**:

Ofman et al. (2014) encontraron que los varones puntuaban más alto que las mujeres ($\mu_{Hombres} > \mu_{Mujeres}$). Como nuestro objetivo es replicar este resultado de Ofman, la hipótesis que tenemos que poner a prueba sin favorecer la hipótesis en la que creemos (que los resultados que vamos a obtener en nuestra muestra van a replicar los obtenidos por Ofman, esta es nuestra hipótesis) es aquella que plantea su complementaria: que la media de RPP en varones será inferior o igual al de las mujeres.

5 – El estadístico de contraste para contrastar la hipótesis de la pregunta anterior, vale, aproximadamente:

A) $T = 1{,}12$
B) $T = 1{,}37$
**C)** $\boldsymbol{T = 1{,}77}$

**Solución**:
aplicamos un contraste de medias para grupos independientes asumiendo varianzas iguales y desconocidas:

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{(n_1-1)\hat{S}_1^2+(n_2-1)\hat{S}_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{3{,}4-3{,}2}{\sqrt{\frac{99\cdot 0{,}8^2+99\cdot 0{,}8^2}{100+100-2}\left(\frac{1}{100}+\frac{1}{100}\right)}} = 1{,}77$$

6 – El nivel crítico asociado al estadístico de contraste de la pregunta anterior, se encuentra entre los valores:

A) $\mathbf{0{,}025} < p < 0{,}05$
B) $0{,}10 < p < 0{,}15$
C) $0{,}05 < p < 0{,}10$

**Solución***:*

*Acudimos a la tabla de curva normal dado que contamos con más de 100 g.l., donde observamos que para una z = 1,77, la p correspondiente es* $p = 1 - 0{,}9616 = \ 0{,}0384$que es un valor que se encuentra entre 0,05 y 0,025.

7 – Los resultados obtenidos en su muestra:

**A) Replican los resultados de Ofman et al. (2014), si bien hay que tener en cuenta que el tamaño del efecto es pequeño según la clasificación de Cohen.**
B) Replican los resultados de Ofman et al. (2014), siendo, además muy grande el tamaño del efecto obtenido según la clasificación de Cohen.
C) No replican los resultados de Ofman et al. (2014).

Solución: debido a que hemos rechazado la hipótesis de que la media en hombres sea superior a la de las mujeres, nuestros resultados van en la línea de Ofman et al. 82014) y calculando el tamaño del efecto se muestra que es pequeño en la clasificación de Cohen.

$$d = \frac{|\bar{Y}_1 - \bar{Y}_2|}{\sqrt{\frac{(n_1-1)\hat{S}_1^2+(n_2-1)\hat{S}_2^2}{n_1+n_2-2}}} = \frac{3{,}4-3{,}2}{\sqrt{\frac{99\cdot 0{,}8^2+99\cdot 0{,}8^2}{100+100-2}}} = 0{,}25$$

*La respuesta correcta es A. Los resultados replican los obtenidos por Ofman et al. (2014) el tamaño del efecto es pequeño.*

8 – Para aplicar un contraste paramétrico a los datos de su muestra:

A) Tanto la variable dependiente como la independiente han de ser cuantitativas (de intervalo o razón).
**B) La variable dependiente ha de ser cuantitativa (de intervalo o razón).**
C) La variable independiente ha de ser cuantitativa (de intervalo o razón).

**Solución**:

Uno de los criterios para aplicar los tests paramétricos es que la VD sea cuantitativa. En el caso de la VI no hay nada específico para la distinción entre paramétrico y no paramétrico. La opción correcta es la B.

**SITUACIÓN 2**. Un profesor está interesado en comparar la eficacia de tres métodos de enseñanza (M1, M2 y M3) sobre el aprendizaje académico. Como sospecha que el Cociente Intelectual (CI) también afecta a este aprendizaje, también manipula a los participantes en función del CI (Alto y Bajo). Dispone de una muestra de 9 sujetos con un CI "Alto" y 9 sujetos con CI "Bajo". Para cada grupo de CI, asigna, aleatoriamente, a 3 sujetos a cada uno de los métodos de enseñanza. Sabiendo que:$[Y] = 811; [AB] = 801$ y conociendo los valores de la matriz AB:

| | M1 | M2 | M3 | |
|---|---|---|---|---|
| Alto | 15 | 24 | 30 | 69 |
| Bajo | 6 | 15 | 21 | 42 |
| | 21 | 39 | 51 | 111 |

***Solución**:*
*La muestra está formada por 18 sujetos, 9 con CI alto y otros 9 con CI bajo. Cada uno de estos 9 sujetos se reparten entre los tres métodos de enseñanza, con lo que n= 3. Por tanto, el diseño es de dos factores con muestras independientes: Método de enseñanza ($a = 3$) y CI ($b = 2$), y cada una de las condiciones experimentales con $n = 3$.*

*Las razones básicas y sumas de cuadrados son:*

| *Razones básicas* | *Sumas de cuadrados* |
|---|---|
| $[A] = \dfrac{21^2 + 39^2 + 51^2}{2 \cdot 3} = 760{,}5$ | $SC_A = [A] - [T] = 760{,}5 - 684{,}5 = 76$ |
| $[B] = \dfrac{69^2 + 42^2}{3 \cdot 3} = 725$ | $SC_B = [B] - [T] = 725 - 684{,}5 = 40{,}5$ |
| $[AB] = 801$ | $SC_{AB} = [AB] - [B] - [A] + [T] = 801 - 760{,}5 - 725 + 684{,}5 = 0$ |
| $[Y] = 811$ | $SC_{Intra} = [Y] - [AB] = 811 - 801 = 10$ |
| $[T] = \dfrac{111^2}{18} = 684{,}5$ | $SC_{Total} = [Y] - [T] = 811 - 684{,}5 = 126{,}50$ |

A continuación, se construye la tabla de ANOVA con esta información, lo cual nos permitirá contestar casi directamente las cuestiones posteriores.

| *FV* | *SC* | *g.l.* | *MC* | *F* | *p* | *F tabla* |
|---|---|---|---|---|---|---|
| A | 76,00 | 2 | 38,00 | 45,600 | 0,0000 | 3,885 |
| B | 40,50 | 1 | 40,50 | 48,600 | 0,0000 | 4,747 |
| A×B | 0,00 | 2 | 0,00 | 0,000 | 1,0000 | 3,885 |
| Intra | 10,00 | 12 | 0,83 | | | |
| Total | 126,50 | 17 | 7,44 | | | |

9–Se trata de un diseño:
- A) De un factor con muestras independientes.
- B) De un factor con muestras relacionadas.
- **C) De dos factores con muestras independientes.**

Dos factores: los tres métodos de enseñanza y el CI de dos niveles, alto y bajo.

10 – El estadístico de contraste para el factor “Método de enseñanza” vale:
- A) 76
- B) 38
- **C) 45,6**

11 – El estadístico de contraste para el factor “CI”, vale:

**A) 48,60**
B) 40,50
C) 4,747

12 – El estadístico de contraste para la interacción, vale:

**A) Cero.**
B) Uno.
C) 3,885

13 – Para $\alpha = 0{,}05$, el valor crítico para comprobar si existe interacción es igual a:

A) 5,096
**B) 3,885**
C) 6,927

**Solución**: para resolver esta cuestión hemos de ver en las tablas de 0,95, el valor correspondiente de F para 2 y 12 grados de libertad. Este valor resulta valer 3,885.

14 – Los resultados revelan diferencias significativas:

A) **Para los factores “Método de enseñanza” y “CI” pero no para la interacción**.
B) Para la interacción pero no para los factores “Método de enseñanza” y “CI”.
C) Observaremos diferencias significativas o no, dependiendo del nivel de confianza adoptado.

15 – Lo más apropiado en este caso, teniendo en cuenta el resultado obtenido para la interacción, es buscar en las tablas el valor crítico para el factor “Método de enseñanza” con los siguientes grados de libertad:

A) 2 y 12
**B) 2 y 14**
C) 1 y 12

**Solución:** como la interacción no ha resultado significativa, lo que haremos será incluir la interacción en el término de error, incrementando de esta forma los grados de libertad del término “intra” en 2 unidades. Lo cual hace que los niveles de libertad sean 2 y 14.

**SITUACIÓN 3.**García-Jiménez, Alvarado-Izquierdo y Jiménez-Blanco (2000) estudiaron la capacidad de la regresión lineal múltiple en la predicción del rendimiento académico, encontrando que el rendimiento previo es un buen predictor del rendimiento futuro, teniendo también un peso importante el grado de participación del alumno en clase. Se desea estudiar en qué medida estas variables son capaces de predecir el rendimiento en alumnos de la UNED en la asignatura de Diseños de Investigación. Para ello se dispone de las puntuaciones de una muestra de 18 estudiantes en las siguientes variables: $X_1$ (nota en la asignatura de primero Introducción al Análisis de Datos), $X_2$ (participación en los foros) e Y (nota en el examen presencial en Diseños de Investigación). Conociendo los siguientes datos y tomando $\alpha = 0{,}05$:

| | Matriz de correlaciones | | |
|---|---|---|---|
| | $X_1$ | $X_2$ | Y |
| $X_1$ | - | 0,658 | 0,361 |
| $X_2$ | | - | 0,028 |
| Y | | | - |

| | Varianzas |
|---|---|
| $X_1$ | 2,917 |
| $X_2$ | 1,930 |
| Y | 3,090 |

16 – La pendiente de la ecuación de regresión lineal simple de Y sobre X1 vale, aproximadamente:

A) 0,36
B) 3,49

**C) 0,37**

**Solución:** En esta situación no disponemos de los datos directos, pero sí disponemos de los valores de correlación y desviaciones típicas. Por consiguiente, calculamos la pendiente mediante:

$$B = r_{xy}\frac{S_y}{S_x} = 0{,}361 \cdot \frac{\sqrt{3{,}090}}{\sqrt{2{,}917}} = 0{,}37$$

17 – El estadístico de contraste para comprobar si es significativo el coeficiente de correlación de Pearson entre X1 e Y, vale, aproximadamente:

A) 1,10
B) 1,84
**C) 1,55**

**Solución:**

$$T = \frac{r_{xy}\sqrt{n-2}}{\sqrt{1-r_{xy}^2}} = \frac{0{,}361\sqrt{18-2}}{\sqrt{1-0{,}361^2}} = 1{,}55$$

18 – El nivel crítico asociado al estadístico de contraste de la pregunta anterior se encuentra, aproximadamente, entre los valores:

**A)** $\mathbf{0,10} < p < 0{,}20$
B) $0{,}20 < p < 0{,}30$
C) $0{,}05 < p < 0{,}10$

**Solución**:
Buscamos en las tablas de la T con 18-2 = 16 grados de libertad. Encontramos que el valor T = 1,55 se encuentra entre los valores 1,337 y 1,746, los cuales dejan por encima de sí 1-0,900 y 1-0,95 respectivamente (es decir, 0,10 y 0,05). Como el contraste de la correlación debe considerarse bilateral, estos valores debemos multiplicarlos por dos. Esto nos deja que la p de T = 1,5 se encuentra entre 0,20 y 0,10.

19 – En función de los resultados obtenidos podemos concluir que la relación entre $X_1$ e Y:

A) Es significativa para NC = 99%
B) Es significativa para NC =95% pero no para NC = 99%
**C) No es significativa para NC = 95%**

**Solución:** al 0,05 el valor de % = 1,55 no resulta significativo. Por consiguiente, tampoco lo es para 0,01.

20 –La proporción de varianza del rendimiento en Diseños de Investigación explicada por el rendimiento en Introducción al Análisis de Datos y la participación en los foros conjuntamente, vale, aproximadamente:

A) 0,13
**B) 0,21**
C) 0,46

Solución: Nos están preguntando por el coeficiente de determinación múltiple de X1 y X2 para predecir Y. Aplicamos la fórmula:

$$R_{y.12}^2 = \frac{r_{y1}^2 + r_{y2}^2 - 2r_{y1}r_{y2}r_{12}}{1-r_{12}^2} = \frac{0{,}361^2 + 0{,}028^2 - 2\cdot 0{,}361 \cdot 0{,}028 \cdot 0{,}658}{1-0{,}658^2} = 0{,}208$$

21 – El coeficiente de correlación entre las variables $X_2$ e Y eliminando de ambas el influjo de $X_1$ vale, aproximadamente:

A) 0,028
**B) - 0,30**
C) 0,30

Solución: Nos están preguntando por el coeficiente de correlación parcial (eliminamos el influjo de X1 tanto de X2 como de Y, luego es parcial):

$$pr_2 = \frac{r_{y2} - r_{y1}r_{12}}{\sqrt{(1 - r_{y1}^2)(1 - r_{12}^2)}} = \frac{0{,}028 - 0{,}361 \cdot 0{,}658}{\sqrt{(1 - 0{,}361^2)(1 - 0{,}658^2)}} = -0{,}30$$

**Preguntas teóricas**

22 – Cuál de los siguientes estimadores de su correspondiente parámetro poblacionalNO es insesgado:

A) La media aritmética.
**B) La varianza.**
C) La proporción.

**Solución**:
La varianza muestral no es un estadístico insesgado de la varianza poblacional mientras que la media aritmética y la proporción sí lo son.

23 – La probabilidad de mantener una hipótesis nula falsa se denomina:

A) $\alpha$
**B)** $\boldsymbol{\beta}$
C) $1 - \beta$

**Solución**:
Nos están preguntado por el error tipo II que es Beta.

24 – Cuando al realizar un ANOVA un investigador no está interesado en comprobar las diferencias entre todas las medias y sabe de antemano que comparaciones le interesan, se denomina a las comparaciones múltiples que ha de realizar:

**A) "a priori".**
B) "a posteriori".
C) "post hoc".

**Solución**: las comparaciones a posteriori se utilizan cuando el investigador no establece a priori las comparaciones a realizar y solo lo hace después de ejecutado el experimento. Por el contrario, cuando sí conoce estos contrastes antes de realizar el estudio, se denominan "a priori".

25 – Al realizar inferencias sobre correlación y regresión, el supuesto de homocedasticidad se refiere a:

A) La normalidad de las distribuciones condicionadas en Y para cada valor de X.
B) La independencia en Y de los valores estimados condicionados para cada valor de X.
**C) La igualdad de varianzas de las distribuciones de los errores condicionadas a cada valor de X.**

**Solución:** el supuesto de homocedasticidad hace referencia a la igualdad de varianzas en las distribuciones de los errores condicionadas a cada valor de X (no a la normalidad ni a la independencia entre valores de Y).