FE DE ERRATAS DE LA UNIDAD DIDÁCTICA 1

TEMA 1:

El punto 1.4.6. es una primera introducción a la decisión estadística que se desarrolla en el tema 12 (pdf). En sentido estricto y técnico, sólo la H_0 es la que se rechaza o no se rechaza. La decisión estadística se realiza comparando la probabilidad asociada al estadístico de contraste con el valor fijado del nivel de significación. La comparación de los valores de los estadísticos de contraste (empírico y crítica) es una derivación de lo anterior y, en consecuencia, no es necesario conocer el valor crítico, sino ambas probabilidades.

TEMA 2:

Tanto en la impresión original (2009) como en la reimpresión de 2012, se ha detectado la siguiente errata en la *Tabla 2.1. Escalas de medida (página 38) y al final de la página 40*…

…ahí aparece la 'temperatura' como una variable con nivel de medida 'de razón'; lo cual es incorrecto…

…lo correcto es decir que la 'temperatura' es una variable con nivel de medida 'de intervalo', dado que sus escalas de medida están construidas internamente con intervalos iguales pero siendo el punto cero de cada una de ellas un punto arbitrario (de hecho, el cero en la escala Kelvin es distinto al cero en la escala Celsius o en la escala Fahrenheit).

Tanto en la impresión original (2009) como en la reimpresión de 2012, en el capítulo 2, página 29, apartado 2.5.1., en el 3º párrafo, donde dice…

*"Aunque algunos autores suelen recurrir al término 'atributo' cuando se identifica con lo cualitativo y 'variable' cuando se trata de lo cuantitativo, en la gran mayoría de los manuales se tiende a englobar todo bajo el epígrafe de variables, **en el primer caso se les denomina categóricas, atributivas o cualitativas, bien sean discretas o continuas**"*

…es más correcto y completo decir:

*"Aunque algunos autores suelen recurrir al término 'atributo' cuando se identifica con lo cualitativo y 'variable' cuando se trata de lo cuantitativo, en la gran mayoría de los manuales se tiende a englobar todo bajo el epígrafe de variables. **En el primer caso se les denomina categóricas, atributivas o cualitativas; en el segundo caso se les denomina cuantitativas, bien sean discretas o continuas**"*

FE DE ERRATAS – Temas 4, 5 y 6

TEMA 4:

Pág. 90. Segundo párrafo, en vez de "…f_a/N…", debe poner "(frecuencia absoluta entre el número total de puntuaciones, f_i/N)".

Pág. 91, primera línea pone:

"…suele incluir, en primer lugar,…" debe decir: "…suele incluir, como primera columna,…"

TEMA 5:

Pág. 98. Línea 7, Donde dice: "…frecuencias absolutas (f)…" debe decir: "…frecuencias absolutas (f_i)…". Asimismo, en el encabezado de la tabla 5.1., segunda columna, debe poner f_i en vez de f.

Pág. 103. En la fórmula de la varianza, debe poner: $s^2 = \dfrac{\Sigma\,(x_i - \overline{x})^2}{N - 1}$

Pág. 105. (Tanto en la impresión original (2009) como en la reimpresión de 2012). Al final del epígrafe 5.3.5. El coeficiente de variación, donde dice: "Hay que tener en cuenta, sin embargo, que este índice tiene unos límites fijos…"; debe decir: "Hay que tener en cuenta, sin embargo, que este índice **NO** tiene unos límites fijos…"

Pág. 107, penúltima frase del primer párrafo, cuando dice: "…sin embargo, podría darse asimetría negativa (por ejemplo la mayoría…" Debe decir: "…sin embargo, podría darse asimetría positiva (por ejemplo la mayoría…."

Pág. 112, último párrafo. Donde dice:

"El diagrama de caja, también conocido como caja y patillas o caja y bigotes, es un gráfico muy práctico porque permite hacerse una idea rápida de la distribución de las puntuaciones en la zona central (el espacio en rojo, que comprende desde el cuartil 1 o percentil 25 hasta el cuartil 3 o percentil 75) y en los extremos. Las patillas representan las puntuaciones hasta los extremos de la distribución.

Este gráfico corresponde a la misma variable presentada en el histograma anterior, y podemos ver que la mitad del grupo se sitúa entre las puntuaciones 65 y 84, estando la mediana en 75. Según se defina en el *programa, las puntuaciones que se alejen de la media del grupo el número de desviaciones típicas que especifiquemos, son consideradas valores atípicos, raros o extremos (outliers). Aunque en el gráfico vemos que la puntuación menor sería el outlier, por ser un caso raro se le dibuja fuera de la patilla inferior, indicando que dicho caso se podría descartar en posteriores análisis".*

Debe decir:

"El diagrama de caja, también conocido como caja y patillas o caja y bigotes, es un gráfico muy práctico porque permite hacerse una idea rápida de la distribución de las puntuaciones en la zona central (el espacio sombreado, que comprende desde el cuartil 1 o percentil 25 hasta el cuartil 3 o percentil 75) y en los extremos. Otra de sus utilidades es que nos informa gráficamente del grado de asimetría de la distribución y presenta los valores atípicos (outliers). Las patillas o bigotes tienen como extremos posibles los valores máximo y mínimo de la variable. Sin embargo, las patillas tienen un máximo de prolongación; para determinar los límites de las patillas suele utilizarse la fórmula de multiplicar el intervalo intercuartil –que es la simple diferencia entre el tercer y el primer cuartil– por ±1,5. Si hay valores que se alejan de los límites de este producto, se clasifican como atípicos. También puede especificarse la existencia de valores aún más extremos: Si hay puntuaciones que se alejan más del producto del intervalo intercuartil por ±3, entonces se habla de puntuaciones extremas. Esto nos ayuda a identificar valores que pueden distorsionar el comportamiento de la variable al calcular valores como la media aritmética, la desviación típica, una correlación, etc. En ciertos casos, puede estar justificado eliminar de algunos análisis estadísticos a los sujetos con puntuaciones atípicas o extremas.

*Este gráfico corresponde a la misma variable presentada en el histograma anterior, y podemos ver que la mitad del grupo se sitúa entre las puntuaciones 65 y 84, estando la mediana en 75. En este caso, el intervalo intercuartil sería 84-65=19. 19*1.5=28.5. Entonces, si le sumamos y restamos este valor a la mediana, obtenemos los extremos de las patillas: 75±28.5= (46.5, 103.5). En el gráfico vemos que la puntuación menor sería el outlier, y por ser un caso raro se le dibuja fuera de la patilla inferior, indicando que dicho caso debería ser analizado antes de incluirlo en posteriores análisis".*

Pág. 113, última línea, donde dice "La mediana más alta corresponde al centro 19", debe decir: "La mediana más alta corresponde al centro 42",

TEMA 6:

Pág.120, final penúltimo párrafo (sobre la tabla 6.1):

DONDE DICE "Estos percentiles, como vimos en el Capítulo 4, se obtienen directamente al realizar la distribución de frecuencias y multiplicar por 100 las frecuencias relativas a las frecuencias acumuladas (f_r x 100/N)."

DEBE decir:

"...y multiplicar por 100 las frecuencias acumuladas relativas (f_{ac}/N x 100); es decir, desde los porcentajes acumulados (redondeados, sin decimales)."

Pág. 121. Fórmula:

Donde dice,

$$P_{25} = 1{,}5 + \frac{\frac{25 \cdot 1012}{100} - 179}{93} \cdot 1 = 2{,}5 + 0{,}80 = 2{,}3$$

Debe decir:

$$P_{25} = 1{,}5 + \frac{\frac{25 \cdot 1012}{100} - 179}{93} \cdot 1 = 1{,}5 + 0{,}80 = 2{,}3$$

Pág. 123, penúltimo párrafo

Donde dice,

"Si la puntuación típica hubiera sido z= -1,25, ¿qué probabilidad existe de obtener una puntuación igual o inferior a dicha z [p (z ≤ 1,25)]? En este caso, puesto que la distribución normal es simétrica, consultaríamos el área menor de Z = 1,25 (p = 0,1056)."

Debe decir:

"Si la puntuación típica hubiera sido z= -1,25, ¿qué probabilidad existe de obtener una puntuación igual o inferior a dicha z [p (z ≤ −1,25)]? En este caso, puesto que la distribución normal es simétrica, consultaríamos el área menor de Z = 1,25 (p = 0,1056). Es decir, por la propiedad de simetría de la curva normal, p (z ≤ −1,25) = p (z ≥1,25)] = 0,1056)."

Página 124.

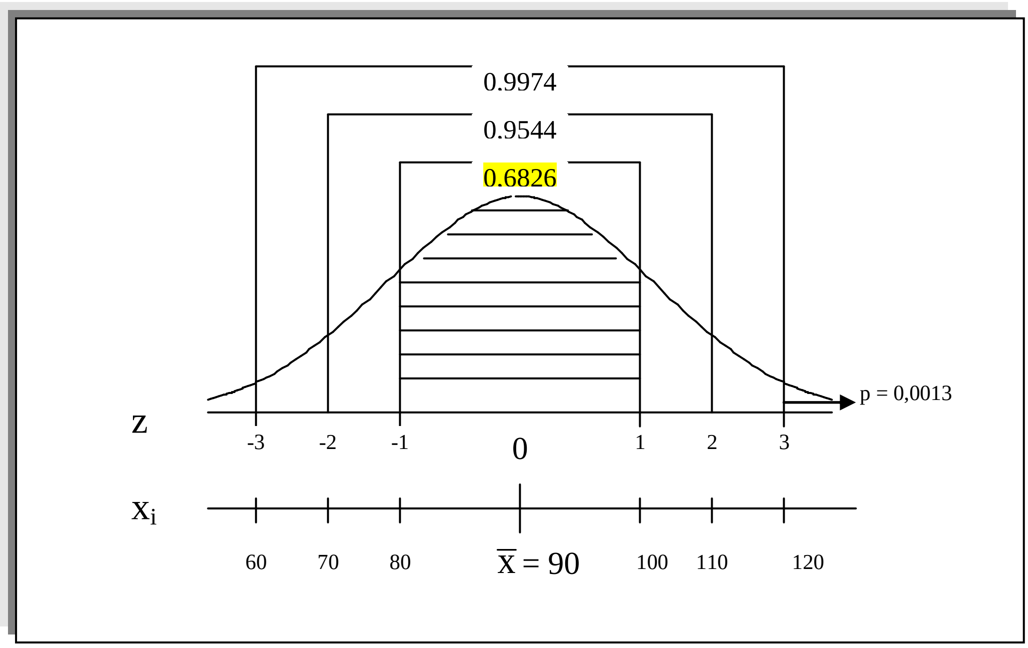
Donde dice,

Evidentemente, la probabilidad de obtener una puntuación superior a la media es de p= 0,5, es decir, en la curva normal el 50 % de los sujetos se encuentran por encima de la media y el otro 50 % por debajo (recordemos que en las puntuaciones Z la media siempre es igual a cero). Por poner otro ejemplo, el 84,14 % (área de la parte mayor; p= 0,8414)...

Debe decir:

Evidentemente, la probabilidad de obtener una puntuación superior a la media es de p= 0,5, es decir, en la curva normal el 50 % de los sujetos se encuentran por encima de la media y el otro 50 % por debajo (recordemos que en las puntuaciones Z la media siempre es igual a cero). Por poner otro ejemplo, el 84,13 % (área de la parte mayor; p= 0,8413)...

Pág. 124, el gráfico inferior debe ser:



Pág. 125. Final primer párrafo:

Donde dice,

"...encuentra el 64 % de los sujetos, entre ±2 el 95 % y entre ±3 casi el 100 % (el 99,7 %)..."

Debe decir:

"...encuentra el 68 % de los sujetos, entre ±2 el 95 % y entre ±3 casi el 100 % (el 99,7 %)..."

Pág. 125. Último párrafo:

Donde dice,

"Por ejemplo, el P_{10} es aquella puntuación que deja por debajo de su límite superior al 10% de los sujetos, por tanto, a una proporción de 0,01 sujetos."... Ahora basta con aplicar la fórmula para saber la puntuación directa equivalente al P_{10}....

Por tanto, una puntuación directa de 65 equivale al percentil normalizado P_{10}.

Debe decir:

"Por ejemplo, el P_{1} es aquella puntuación que deja por debajo de su límite superior al 1% de los sujetos, por tanto, a una proporción de 0,01 sujetos."... Ahora basta con aplicar la fórmula para saber la puntuación directa equivalente al P_{1}....

Por tanto, una puntuación directa de 65 equivale al percentil normalizado P_{1}.

FE DE ERRATAS DEL TEMA 7

TEMA 7:

Tanto en la impresión original (2009) como en la reimpresión de 2012 se ha detectado en la **página 134**, en el recuadro inferior, donde dice: "…Lo que se interpreta como que el tiempo…"

Debería decir: "…Lo que se interpreta como que ambas variables comparten un 49% de su varianza".

(No se puede decir, en el contexto de la correlación, varianza de una variable explicada por la otra, sino que hay que hablar de varianza compartida. En una correlación nunca podemos hablar de relaciones causa-efecto, sino de variación concomitante).

En la **página 142**, penúltimo párrafo, donde dice: "…una relación entre las variables nivel socioeconómico y elección de estudios de tipo medio"

Debería decir: "…una relación entre las variables nivel socioeconómico y elección de estudios de tipo bajo"

En la **página 145**, para una mejor interpretación del coeficiente r_{bp}, se puede eliminar el valor absoluto del numerador. Así quedaría $r_{bp} = -0,20$ y la interpretación sería: correlación de intensidad muy baja negativa: al valor bajo en la dicotómica *q=0=hombres*, se asocia (con muy poca intensidad) un nivel más alto en la variable cuantitativa.

En la **página 149**, en el cálculo de $\overline{X}_p$, $\overline{X}_q$ y $\overline{X}_t$ falta incluir el símbolo Σ (sumatorio) en el numerador. Así, donde pone $\overline{X}_p = \frac{X_i * f_p}{n_p}$ debería poner $\overline{X}_p = \frac{\Sigma(X_i * f_p)}{n_p}$. Análogamente para $\overline{X}_q$ y $\overline{X}_t$

FE DE ERRATAS DEL TEMA 8

TEMA 8:

En el **ejemplo 8.1 (páginas 153 y 154)** se obtiene erróneamente un r_{xy} = 0,96.

El r_{xy} correcto es 0,923 cuyo cálculo se detalla a continuación:

X	Y	X^2	Y^2	X*Y
12	13	144	169	156
18	17	324	289	306
15	15	225	225	225
11	10	121	100	110
9	10	81	100	90
17	16	289	256	272
13	15	169	225	195
8	9	64	81	72
18	17	324	289	306
12	12	144	144	144
7	6	49	36	42
11	10	121	100	110
15	16	225	256	240
20	14	400	196	280
14	14	196	196	196
8	8	64	64	64
10	11	100	121	110
17	16	289	256	272
14	13	196	169	182
11	12	121	144	132
7	8	49	64	56
17	15	289	225	255
12	12	144	144	144
12	13	144	169	156
18	19	324	361	342
15	16	225	256	240
13	13	169	169	169
8	7	64	49	56
11	13	121	169	143
18	17	324	289	306
10	10	100	100	100
17	16	289	256	272
8	9	64	81	72
12	13	144	169	156
14	15	196	225	210
16	17	256	289	272
8	9	64	81	72
11	10	121	100	110
18	19	324	361	342
12	12	144	144	144
Σ = 517	Σ = 517	Σ = 7201	Σ = 7117	Σ = 7121

$$r_{xy} = \frac{(40 \cdot 7121) - (517 \cdot 517)}{\sqrt{[40 \cdot 7201 - (517)^2][40 \cdot 7117 - (517)^2]}} = \frac{284840 - 267289}{\sqrt{20751 \cdot 17391}} = \frac{17551}{18996,86} = 0,923$$

En la **página 155, punto 8.2.1.3., segundo párrafo**. Donde dice: "En efecto: por un lado está la puntuación que corresponde a los ítems impares y por otro la de los impares". Debe decir "pares".

En la **página 160, punto 8.2.2.1, apartado a), segundo párrafo**. Donde dice:

$$n = \frac{\text{nº de elementos finales}}{\text{nº elementos sin iniciales}}$$

Debe decir:

$$n = \frac{\text{nº de elementos finales}}{\text{nº elementos iniciales}}$$

En el **ejemplo 8.7 de la página 164** hay varios errores en la **Tabla 8.3.**, que se arrastran al calcular el coeficiente de correlación r_{xy} de validez predictiva.

La tabla y el coeficiente correctos son:

	A	B	C	D	E	F
1		**X**	**Y**	**X^2**	**Y^2**	**X*Y**
2		20	8	400	64	160
3		12	3	144	9	36
4		17	5	289	25	85
5		25	10	625	100	250
6		8	2	64	4	16
7		15	5	225	25	75
8		21	7	441	49	147
9		13	4	169	16	52
10		16	6	256	36	96
11		30	9	900	81	270
12		9	2	81	4	18
13		18	5	324	25	90
14		14	3	196	9	42
15		28	10	784	100	280
16		22	6	484	36	132
17		11	2	121	4	22
18		15	6	225	36	90
19		22	5	484	25	110
20		26	9	676	81	234
21		13	3	169	9	39
22	Suma	355	110	7057	738	2244
23						
24	Correlación r_{xy}	**0,91944057**				

En la **página 167, punto 8.3.6.3**. El resultado del cálculo de "E": donde dice 52,50, debe decir 0,5250

FE DE ERRATAS – Temas 9, 10 y 11

TEMA 9

Modelos estadísticos y probabilidad. La curva normal de probabilidades

LAS ERRATAS CORREGIDAS SE MARCAN EN ROJO

Pág. 196

$$\chi^2 = \Sigma\, [(f_o\ -\ f_e)^2 / f_e]$$

Este estadístico se distribuye según la distribución χ^2 para **un valor igual al de filas** menos 1 cuando μ y σ son conocidas, y con -3 en caso de ser estimadas. Vale la pena recordar aquí que χ^2 es un modelo estadístico.

Pág. 197. Cuadro 9.2

I	f_o	L_i	z_i	$p(z_i)$	p_i	f_e	$(f_o - f_e)$	$(f_o - f_e)^2$	$(f_o - f_e)^2 / f_e$
					0,0055	15,48	15,48	239,54	15,48
10	74	10,5	2,54	0.9945	0.0157	44,18	29,82	889,23	20,13
9	175	9,5	2,026	0.9788	0.0443	124,66	50,34	2534,12	20,33
8	219	8,5	1,51	0.9345	0.0956	269,02	50,02	2502,00	9,3
7	340	7,5	0,99	0.8389	0,1563	439,83	99,83	9966,03	22,66
6	528	6,5	0,475	**0.6808**	**0.1968**	**553,79**	**25,79**	**665,12**	**1,2**
5	750	5,5	-0,04	0.4840	0,1963	552,39	197,61	39049,71	70,69
4	370	4,5	-0,56	0.2877	0,1454	409,16	39,16	1533,51	3,75
3	210	3,5	-1,07	0.1423	0,0864	243,13	33,13	1097,6	4,51
2	96	2,5	-1,59	0.0559	0.0385	108,34	12,34	152,27	1,41
1	43	1,5	-2,11	0.0174	0,0132	**37,14**	**5,86**	**34,34**	**0,92**
0	9	0,5	-2,625	0.0043	0,0043	12,10	3,10	9,61	0,79
N = 2814						2814			**171,17**

Página 197, último párrafo:

"El caso de la primera **fila** es diferente..."

Pág. 198

$$\chi^2 = \Sigma\, [(f_o\ -\ f_e)^2 / f_e] = \mathbf{171{,}17}$$

Si nuestro valor empírico (171,17) fuera igual o mayor que el de las tablas para los g.l. correspondientes, rechazaríamos H_0 y afirmaríamos que nuestros datos empíricos no son compatibles con el modelo de la curva normal de probabilidades con una probabilidad de tomar una decisión errónea $\leq \alpha$. Pues bien, las tablas de ji cuadrado, para un nivel de confianza del 99 % (α = 0,01), y (11- 1) g.l. nos dan un valor de 23,209. Para 11 – 3 g.l., el valor es de 20,090.

TEMA 10

Los baremos. Muestreo y aplicaciones

LAS ERRATAS CORREGIDAS SE MARCAN EN ROJO

Pág. 203

En el primer caso, bastaría decir que la mediana del grupo de niños del que ese alumno forma parte es de 22 puntos para tener ya una primera idea del valor de esos 35. Si, además, supiéramos que ~~**de**~~ la puntuación...

Pág. 207, nota 14 a pie de página

El intervalo de confianza es el conjunto de puntuaciones entre cuyos límites se considera que estará la verdadera puntuación de la población, si bien no con seguridad sino para un determinado nivel de probabilidad (por lo general el 0,95 o 0,99, esto es, 95 o el 99%).

Pág. 210. Tabla 10.1

I	X_i	f_i	f_a	$\%_a$	z_i	z_{norm}
37-40	38,5	1	70	100	2,75	
33-36	34,4	2	69	98,57	2,12	2,19
29-32	30,5	6	67	95,71	1,48	1,72
25-28	26,5	11	61	**87,14**	0,84	**1,13**
21-24	22,5	15	50	71,43	0,21	0,57
17-20	18,5	20	35	50	-0,425	0
13-16	14,5	10	15	21,43	-1,06	-0,79
9-12	10,5	4	5	7,14	- 1,70	**-1,465**
5- 8	6,5	1	1	1,43	- 2,33	-2,19
		N = 70 Media: 1482,8 : 70 = 21.18 s = 6,296				

Pág. 215

En el caso de otro sujeto con puntuación directa de 13, su z = **-1,35**. Si deseamos obtener su z normalizada deberemos comprobar qué % de casos deja por debajo de sí. Como 13 deja 6 casos, el % es de 8,57. En la tabla de áreas de la curva normal, columna "área de la parte menor", encontramos que z = **-1,37** supera al 8,53 %. Esta sería, por tanto, su z normalizada[1].

(...)

[1] Podemos comprobar que, mientras en esta ocasión la z empírica y la normalizada son muy próximas, en el caso anterior son muy diferentes. Este hecho se explica porque la distribución empírica se aleja claramente de la normal en la parte superior de los datos, como puede comprobar el lector en el histograma correspondiente

La figura 10.3 recoge la equivalencia de una serie de puntuaciones normalizadas. Vale la pena señalar que las normas o baremos utilizados en PISA toman como media = 500 y s = 100. Por tanto, a una alumna que obtenga 600 puntos se corresponde una puntuación normalizada en PISA igual a +1; si fuera -1 su puntuación directa sería de 400 puntos y si hubiera obtenido 950 su puntuación normalizada seria de **+4,5.**

Pág. 216

Estaninas y pentas

En los EE.UU se utiliza frecuentemente una escala de diez rangos, creados a partir de 9 puntos –estanina = contracción de *standard nine-* cuya media es de 5 y su desviación típica de 2. En nuestro país se utiliza con cierta frecuencia una escala de cinco rangos, denominada *pentas,* que permite dividir la serie en cinco grandes bloques, cuyos límites en puntuaciones *z* se aprecian en la Tabla 10.3. La escala de pentas tiene como media 3 y como desviación típica **1.**

Pág. 218

Para el cálculo del tamaño de la muestra contamos con dos fórmulas diferentes, según que el tamaño de la población de origen sea **infinita** (10.3) o **finita** (10.4):

(...)

Donde *N* es el tamaño de la población, *n* el de la muestra; *z* es el valor que corresponde al nivel de confianza elegido (número de desviaciones típicas precisas para que la curva normal deje en su interior el 95, 99, 99,9 % etc.) y *p* y *q* el correspondiente a la proporción de la característica en la población. *E* representa el error de estimación admitido por el investigador. Tanto *z* como *E* están elevados al cuadrado.

Pág. 220

El muestreo sistemático es una modalidad del anterior que nos permite fijar el primero de los sujetos de la muestra y, a partir de él, seleccionar sistemáticamente el resto sumándole un valor constante, en concreto el denominado *coeficiente de elevación,* esto es, el cociente entre el tamaño de la población y el de la muestra. Por ejemplo, en nuestro primer caso, tal coeficiente sería: 108000 / 1849 = 58,41. Pues bien: seleccionado al azar el primer caso, seguiríamos eligiendo en las tablas de 58 en 58, hasta llegar a los 1849 que integran la muestra.

Pág. 221

Así, en el caso de las dos muestras anteriores, para poblaciones de 108.000 y 35.600 casos respectivamente, los errores muestrales, para el caso de que *p* = *q* = 50 en el primer caso, y de que *p* = 35 y *q* = 65 en el segundo, tendríamos:

- E = √ ($2{,}56^2$ * 50 * 50) / 1849 = **3**
- E = √ [($2{,}56^2$ * 65 * 35 / 1607) x (35.600 – 1607) / (35.599)] = √ 9,42 * 0.955 = √ 8,996 = **3**
- En caso de que las proporciones de *p* y *q* fueran, como en el caso anterior, de 50 * 50, el valor resultante sería de **3,14. (**√ 10,35 * 0.955 = √ 9,88 = **3,14)**

(...)

Así, asumiendo que en una muestra de 1.849 adultos, de una población de 108.000 que no obtuvieron el graduado en Educación Secundaria Obligatoria, distribuida normalmente, el 46 % fueron mujeres, podemos crear un intervalo de confianza, para una probabilidad del 99 %, sumando y restando a ese 46 % el valor del error muestral, esto es:

$$46 \pm 3 = 43 \text{ y } 49.$$

Pág. 222

Ahora sí podemos afirmar que, en toda la población, con una probabilidad del 0.99 (nivel de confianza del 99 %) el número de **sujetos** que no obtuvieron el graduado es superior entre los varones que entre las mujeres.

Página 219:

La primera fórmula dice: N = (2,58... debe decir **n** = (2,58...

TEMA 11

Estimación de parámetros. Errores de estimación.

Pág. 229, tercer párrafo.

Donde dice,

“Recordamos que en el denominador aparecerá N-1 cuando se haya calculado en la muestra la desviación típica insesgada, o sólo N cuando se haya calculado la desviación típica sesgada.”

Debe decir:

“Recordamos que en el denominador aparecerá N-1 cuando se haya calculado en la muestra la desviación típica sesgada, o sólo N cuando se haya calculado la desviación típica insesgada.”

Pág. 229, quinto párrafo.

Donde dice,

“Para ello hemos tenido que definir el nivel de significación con el que vamos a realizar la estimación, calcular la puntuación *z* correspondiente a dicho nivel (puntuación conocida gracias a que la distribución muestral es normal) y multiplicar dicha *z* por el ***error típico***, esto es, la desviación típica de la distribución muestral.”

Debe decir:

Para ello hemos tenido que definir el nivel de significación con el que vamos a realizar la estimación. Si σ es conocido (o también si la muestra es grande), debemos entonces recurrir a la distribución normal y escoger la puntuación *z* correspondiente a dicho nivel de significación, calcular la puntuación *z* correspondiente a dicho nivel (puntuación conocida gracias a que la distribución muestral es normal) y multiplicar dicha *z* por el *error típico*, esto es, la desviación típica de la distribución muestral.”

Pág. 230, penúltimo párrafo.

Donde dice,

“En definitiva, I.C. = 105 ± 0,82, con lo que obtenemos ambos límites confidenciales, inferior y superior: $L.C._I$ = 104,18 y $L.C._S$ = 105,82. En conclusión, podemos decir con un nivel de confianza del 99 % que la media en inteligencia de los adolescentes de la Comunidad de Madrid se encuentra entre los valores 104,18 y 105,82.”

Debe decir:

“En definitiva, I.C. = 105 ± 0,82, con lo que obtenemos ambos límites confidenciales, inferior y superior: $L.C._I$ = 104,18 y $L.C._S$ = 105,82. En conclusión, podemos decir con un nivel de confianza del 99 % que, si repitiésemos la estimación un número muy elevado de ocasiones, en el 99% de esas muestras, la media en inteligencia de los adolescentes de la Comunidad de Madrid se encontraría entre los valores 104,18 y 105,82."

Pág. 230, título 11.4.1.

Donde dice,

11.4.1. Estimación del parámetro media aritmética para muestras pequeñas

Debe decir:

11.4.1. Estimación del parámetro media aritmética (σ desconocido y para muestras pequeñas).

Pág. 232, segundo párrafo.

Donde dice,

"En definitiva, I.C. = 105 ± 5,71, con lo que obtenemos ambos límites confidenciales, inferior y superior: $L.C._I$ = 99,29 y $L.C._S$ = 110,71. En conclusión, podemos decir con un nivel de confianza del 99 % que la media en inteligencia de los adolescentes de la Comunidad de Madrid se encuentra entre los valores 99,29 y 110,71. Como vemos, hemos perdido un gran nivel de precisión al disminuir el tamaño de la muestra, de modo que ahora la horquilla del intervalo de confianza es superior a 11 puntos."

Debe decir:

"En definitiva, I.C. = 105 ± 5,71, con lo que obtenemos ambos límites confidenciales, inferior y superior: $L.C._I$ = 99,29 y $L.C._S$ = 110,71. En conclusión, podríamos afirmar que el intervalo de confianza del 99% alrededor de la medida de la muestra cae entre los valores de 99,29 y 110,71. También podemos decir que si utilizásemos este mismo procedimiento de estimación en cientos o miles de muestras, con un nivel de confianza del 99 %, en el 99% de esas muestras, la media en inteligencia de los adolescentes de la Comunidad de Madrid se encontraría entre los límites confidenciales establecidos (y la nuestra podría ser, o no, una de ellas). Sería un error decir que tenemos una probabilidad de 0.99 de que el valor del parámetro se encontrará entre esos límites confidenciales. Como vemos en nuestro ejemplo, hemos perdido un gran nivel de precisión al disminuir el tamaño de la muestra, de modo que ahora la horquilla del intervalo de confianza es superior a 11 puntos."

Pág. 232, última fórmula.

En el numerador de la fórmula del error típico de la proporción falta la raíz cuadrada.

Donde dice $\sigma_p = \frac{p \cdot q}{\sqrt{N-1}}$

Debe decir $\sigma_p = \frac{\sqrt{p \cdot q}}{\sqrt{N-1}}$

Pág. 232, último párrafo.

Donde dice:

"El intervalo de confianza se establecerá igual que en la media aritmética, pero partiendo de una proporción. En el típico caso del muestreo electoral, donde podemos obtener un estadístico de, por ejemplo, p= 0,36, es decir, la media es 0,36, o también, que el 36% de la muestra va a votar al partido X. Luego, calculamos los límites confidenciales para decir que la horquilla de votantes a ese partido se encuentra entre el 34% y el 38%."

Debe decir

"El intervalo de confianza se establecerá igual que en la media aritmética, pero partiendo de una proporción. Este es el típico caso del muestreo electoral. Supongamos una intención de voto del 36% para el partido X (p = 0,36) en una muestra de 300 sujetos, es decir, la media es 0,36. El numerador sería la raíz cuadrada de 0,64*0,36 (esto es, 0,48). Por tanto, el error típico sería 0,028. Tomando un α = 0,01 → $\alpha/2$ = 0,005 → $z_{(\alpha/2)}$ = 2,57; calculamos los límites confidenciales para decir que la horquilla de votantes a ese partido se encuentra en el intervalo IC = 0,36 ± EM = 0,36 ± 2,57* 0,028 = 0,36 ± 0,07; es decir entre el 29% y el 43% de intención de voto.

Pág. 233, en el primer párrafo (Tanto en la impresión original (2009) como en la reimpresión de 2012)**.**

Donde dice:

"Por tanto, el error típico sería de 0.028".

Debe decir:

"Por tanto, el error típico sería de 0.028 multiplicado por el correspondiente valor de t o z".

Pág. 235, en el segundo párrafo.

Donde dice:

"Aplicando la fórmula, obtenemos que el error típico será de $(1\text{-}0{,}35^2)/\sqrt{19}$ = 0,29 y la $z_{(\alpha/2)}$ sabemos que es 1,96, luego el error muestral es igual a 0,56.

En consecuencia, la correlación en la población estará entre los límites -0,21 y 0,91. Efectivamente, como pensará el estudiante, es un enorme intervalo de confianza, desde una correlación negativa baja hasta una correlación positiva y muy alta. ¿Por qué sucede…"

Debe decir:

"Aplicando la fórmula, obtenemos que el error típico será de $(1\text{-}0{,}35^2)/\sqrt{19}$ = 0,20 y la $z_{(\alpha/2)}$ sabemos que es 1,96, luego el error muestral es igual a 0,39.

En consecuencia, la correlación en la población estará entre los límites -0,04 y 0,74. Efectivamente, como pensará el estudiante, es un enorme intervalo de confianza, desde una correlación negativa (o nula) hasta una correlación positiva y alta. ¿Por qué sucede…”

Pág. 238, primera línea

Donde dice

y para muestras pequeñas e independientes

Debe decir

y para muestras pequeñas e independientes (también utilizable para muestras grandes)

Pág. 238, encima del gráfico

Donde dice

E.M. = $z_{(\alpha/2)} \cdot \sigma_{(\bar{x}1 - \bar{x}2)}$ = 1,96 · 1,57 = 3,08; por tanto,

Debe decir

E.M. = $t_{(\alpha/2)} \cdot \sigma_{(\bar{x}1 - \bar{x}2)}$ = 1,96 · 1,57 = 3,08; por tanto,