Bioestadística: Inferencia Estadística. Análisis de Varias Muestras

M. González

Departamento de Matemáticas. Universidad de Extremadura

2. Para evaluar la influencia del tipo de acidosis del recién nacido en los niveles de glucemia medidos en el cordón umbilical del mismo, se obtuvieron los datos de la siguiente tabla:

Controles	51	56	58	60	62	63	65	68	72	73
Acidosis Respiratoria	60	65	66	68	68	69	73	75	78	80
Acidosis Metabólica	69	73	74	78	79	79	82	85	87	88
Acidosis Mixta	70	75	76	77	79	80	82	86	88	89

Suponiendo normalidad de la variable e igualdad de varianzas, analiza los resultados obtenidos.

$$X_1 = \text{nivel glucemia control} \sim N(\mu_1, \sigma^2)$$

$$X_2 = \text{nivel glucemia acidosis respiratoria} \sim N(\mu_2, \sigma^2)$$

$$X_3 = \text{nivel glucemia acidosis metabólica} \sim N(\mu_3, \sigma^2)$$

$$X_4 = \text{nivel glucemia acidosis mixta} \sim N(\mu_4, \sigma^2)$$

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 \quad \text{vs.} \quad H_1 : \mu_i \neq \mu_j \text{ para algún } i \neq j$$

PARAMÉTRICO. RESUMEN XV: ANOVA DE 1 VÍA

2.

$$\begin{array}{lll} n_1 = 10 & T_{1\cdot} = 628 & \bar{x}_{1\cdot} = 62.8 \\ n_2 = 10 & T_{2\cdot} = 702 & \bar{x}_{2\cdot} = 70.2 \\ n_3 = 10 & T_{3\cdot} = 794 & \bar{x}_{3\cdot} = 79.4 \\ n_4 = 10 & T_{4\cdot} = 802 & \bar{x}_{4\cdot} = 80.2 \\ N = 40 & T_{\cdot\cdot} = 2926 & \bar{x}_{\cdot\cdot} = 73.15 \end{array}$$

$$V_T = V_e + V_d$$

- $V_e = \sum_{i=1}^{k} n_i(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2 = \sum_{i=1}^{k} \frac{T_{i\cdot}^2}{n_i} - \frac{T_{\cdot\cdot}^2}{N}$

$$V_e = \frac{628^2}{10} + \frac{702^2}{10} + \frac{794^2}{10} + \frac{802^2}{10} - \frac{2926^2}{40} = 2045.9$$

- $V_T = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{\cdot\cdot})^2 \sum_{i=1}^{k} \sum_{j=1}^{n_i} X_{ij}^2 - \frac{T_{\cdot\cdot}^2}{N}$

$$V_T = (51^2 + 56^2 + \ldots + 88^2 + 89^2) - \frac{2926^2}{40} = 3517.1$$

- $V_d = V_T - V_e = 1471.2$

2. TABLA DE ANÁLISIS DE LA VARIANZA

F.V.	S. C.	G. L.	M. C.	F
ENTRE	$V_e = 2045.9$	3	$mc_e = 681.97$	$F_{exp} = 16.69$
DENTRO	$V_d = 1471.2$	36	$mc_d = s_d^2 = 40.87$	
TOTAL	$V_T = 3517.1$	39		

Rechazamos H_0 al nivel de significación α si $F_{exp} > F_\alpha(k-1, N-k)$

		$F_\alpha(3,40)$	$F_\alpha(3,30)$	
$\alpha = 0.05$	$F_{0.05}(3,36) \in$	(2.84 ,	2.92)	H_1
$\alpha = 0.01$	$F_{0.01}(3,36) \in$	(4.31 ,	4.51)	H_1
$\alpha = 0.001$	$F_{0.001}(3,36) \in$	(6.59 ,	7.05)	H_1

$$p < 0.001$$

2. COMPARACIONES MÚLTIPLES. RESUMEN XVII.
MÉTODO DE TUKEY-KRAMER

$$\begin{array}{cccc} m_1 = \bar{x}_{1.} = 62.8 & m_2 = \bar{x}_{2.} = 70.2 & m_3 = \bar{x}_{3.} = 79.4 & m_4 = \bar{x}_{4.} = 80.2 \\ \mu_1' = \mu_1 & \mu_2' = \mu_2 & \mu_3' = \mu_3 & \mu_4' = \mu_4 \end{array}$$

μ_i' y μ_j' se declaran no homogéneas si $q_{ij} = \dfrac{m_i - m_j}{\sqrt{\frac{s_d^2}{2}\left(\frac{1}{n_i'} + \frac{1}{n_j'}\right)}} > q_\alpha(N-k, k)$

$$\begin{array}{ll} \alpha = 0.1 & q_{0.1}(36,4) \in (3.35, 3.39) \\ \alpha = 0.05 & q_{0.05}(36,4) \in (3.79, 3.85) \\ \alpha = 0.01 & q_{0.01}(36,4) \in (4.70, 4.80) \end{array}$$

	$\mu_1 - \mu_4$	$\mu_2 - \mu_4$	$\mu_3 - \mu_4$	$\mu_1 - \mu_3$	$\mu_2 - \mu_3$	$\mu_1 - \mu_2$
q_{ij}	$q_{14} = 8.61$	$q_{24} = 4.95$	$q_{34} = 0.40$	$q_{13} = 8.21$	$q_{23} = 4.55$	$q_{12} = 3.66$
$\alpha = 0.1$			H_0			H_1
$\alpha = 0.05$					H_1	H_0
$\alpha = 0.01$	H_1	H_1		H_1	H_0	

Comparación Medias Poblacionales

2. ¿Qué ocurre si no tenemos normalidad o igualdad de varianzas?

NO PARAMÉTRICO: RESUMEN XVI TEST DE KRUSKAL-WALLIS

C	(1) 51	(2) 56	(3) 58	(4.5) 60	(6) 62	(7) 63	(8.5) 65	(12) 68	(17) 72	(19) 73
AR	(4.5) 60	(8.5) 65	(10) 66	(12) 68	(12) 68	(14.5) 69	73	75	78	80
AM	(14.5) 69	(19) 73	74	78	79	79	82	85	87	88
AMi	(16) 70	75	76	77	79	80	82	86	88	89

C	AR	AM	AMi
$R_1 = 80$	$R_2 = 161$	$R_3 = 283$	$R_4 = 296$

EMPATES:

$60 \rightarrow 2 = t_1$ $\quad 65 \rightarrow 2 = t_2$ $\quad 68 \rightarrow 3 = t_3$ $\quad 69 \rightarrow 2 = t_4$ $\quad 73 \rightarrow 3 = t_5$

$75 \rightarrow 2 = t_6$ $\quad 78 \rightarrow 2 = t_7$ $\quad 79 \rightarrow 3 = t_8$ $\quad 80 \rightarrow 2 = t_9$ $\quad 82 \rightarrow 2 = t_{10}$

$88 \rightarrow 2 = t_{11}$

$T_i = 2^3 - 2 = 6, \; i = 1,2,4,6,7,9,10,11\,; \quad T_i = 3^3 - 3 = 24, \; i = 3,5,8;$

$\sum_i T_i = 6 \times 8 + 24 \times 3 = 120 \rightarrow C = 1 - \frac{1}{N^3 - N} \sum_i T_i = 1 - \frac{1}{40^3 - 40} 120 = 0.998$

2. $R_1 = 80, R_2 = 161, R_3 = 283, R_4 = 296 \qquad C = 0.998$

$$\begin{aligned} H &= \frac{12}{CN(N+1)}\left(\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right) - \frac{3(N+1)}{C} \\ &= \frac{12}{0.998 \times 40 \times 41}\left(\frac{80^2}{10} + \frac{161^2}{10} + \frac{283^2}{10} + \frac{296^2}{10}\right) - \frac{3 \times 41}{0.998} \\ &= 23.407 \end{aligned}$$

Se rechaza H_0 al nivel α si $H > \chi^2_\alpha(k-1)$

$\alpha = 0.05 \qquad \chi^2_{0.05}(3) = 7.817 \qquad H > \chi^2_{0.05}(3) \qquad H_1$

$\alpha = 0.001 \qquad \chi^2_{0.001}(3) = 16.291 \qquad H > \chi^2_{0.001}(3) \qquad H_1$

$\mathrm{p} < 0.001$

7. El control del estrés es uno de los factores que más influyen en los resultados de los deportistas de élite, pero su nivel no es el mismo en los distintos deportes (aún cuando se trate de competiciones del máximo nivel). Para comprobar esto, se tomaron atletas de tres deportes (fútbol, atletismo y natación) en los que se anotó su nivel de estrés (muy fuerte, fuerte, moderado y leve). Los resultados se dan en la siguiente tabla:

	Muy fuerte	Fuerte	Moderado	Leve	Total
Fútbol	22	48	120	30	220
Atletismo	41	38	20	11	110
Natación	48	39	10	18	115
Total	111	125	150	59	445

¿Tenemos razones para pensar que el grado de estrés no es el mismo para cada uno de estos tres deportes?

RESUMEN XVIII: TEST DE HOMOGENEIDAD DE VARIAS MUESTRAS - CASO DISCRETO

7. VALORES OBSERVADOS - VALORES ESPERADOS:

	Muy fuerte	Fuerte	Moderado	Leve	Total
Fútbol	22 54.9	48 61.8	120 74.2	30 29.2	220
Atletismo	41 27.4	38 30.9	20 37.1	11 14.6	110
Natación	48 28.7	39 32.3	10 38.8	18 15.2	115
Total	111	125	150	59	445

CONDICIONES DE VALIDEZ: Todos $E_{ij} > 5$.

$$\chi^2_{exp} = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \sum_i \sum_j \frac{O_{ij}^2}{E_{ij}} - T$$

$$= \frac{22^2}{54.9} + \frac{48^2}{61.8} + \ldots + \frac{18^2}{15.2} - 445 = 104.46$$

Se rechaza H_0 al nivel α si $\chi^2_{exp} > \chi^2_{\alpha}((r-1)(s-1))$

$\alpha = 0.05$	$\chi^2_{0.05}(6) = 12.59$	$\chi^2_{exp} > \chi^2_{0.05}(6)$	H_1
$\alpha = 0.001$	$\chi^2_{0.001}(6) = 22.676$	$\chi^2_{exp} > \chi^2_{0.001}(6)$	H_1

$p < 0.001$

7. TABLA DE PORCENTAJES:

	Muy fuerte	Fuerte	Moderado	Leve	Total
Fútbol	22	48	120	30	220
Atletismo	41	38	20	11	110
Natación	48	39	10	18	115
Total	111	125	150	59	445

%	Muy fuerte	Fuerte	Moderado	Leve	Total
Fútbol	10	21.8	54.5	13.6	100
Atletismo	37.3	34.5	18.2	10	100
Natación	41.7	33.9	8.7	15.7	100