

17

Análisis de Varianza

En este primer capítulo dedicado al tema del Análisis de Varianza, se verán los conceptos fundamentales, supuestos básicos y forma de cálculo del modelo más elemental, llamado *unifactorial*. En los capítulos subsiguientes se verán modelos *bifactoriales* y los temas relacionados con los mismos. Paralelamente se mostraran modelos no paramétricos, para los casos donde uno o más de los supuestos básicos no se verifiquen. En adelante, se emplearán el nombre que le puso J.W. Tukey al tema: ANOVA, una manera resumida y practica de referirse a la materia en cuestión y de uso muy difundido en la literatura actual.

Los modelos de ANOVA son el capítulo fundamental de la Estadística porque son algo más que unas técnicas aplicadas. Son la herramienta fundamental para adentrarse en la naturaleza de la variación de los acontecimientos; permiten discernir mejor las causas de los fenómenos y los efectos de los factores involucrados. Es una herramienta indispensable para el bioquímico y farmacéutico moderno. No solo lo introduce en la misma Naturaleza de las cosas, sino que es la herramienta básica para el diseño de experimentos. Toda vez que necesite buscar las causas que hayan descontrolado sus técnicas de laboratorio o de producción, podrá usar la filosofía de estos modelos para realizar su investigación.

El ANOVA puede ser considerado como una manera de verificar si dos o más medias muestrales fueron extraídas de una misma población o de poblaciones con el mismo valor esperado, para una magnitud clínica dada. En consecuencia, cuando estas medias muestrales no sean coincidentes habrá que suponer que provienen de poblaciones diferentes por el efecto causado por un factor en estudio. Como por ejemplo, comparar las medias muestrales de un placebo *versus* las medias de muestras con diferentes dosis de un medicamento. O bien, la comparación entre sí de varias marcas comerciales, proveedoras de drogas o kits de medición, como además comparar varios operadores, o equipos, o pipetas entre sí, etc.

Cuando se trabaja con dos muestras se usa el modelo de Student para muestras independientes como se vio en los temas anteriores. El ANOVA es un método más general, que se extiende a más de dos muestras y se puede demostrar que coincide con Student si se aplica a solo dos muestras. Es decir, el modelo de Student es un caso particular del ANOVA.

Estos modelos desarrollados por R. A. Fisher a principios de este siglo, tienen una distribución teórica esperada: la función F, tabulada por G. W. Snedecor, sita en la Tabla 7 del Anexo.

Es decir, que los tests estadísticos a realizar se basan en comparar el valor muestral calculado con los datos medidos: F contra un valor crítico de tablas F_{α} . La idea básica del método es que si las muestras son *normales, independientes y aleatorias*; y se supone que todas tienen la misma varianza (*homocedásticas*), entonces, para que provengan de una misma población se necesita únicamente que las medias muestrales sean todas iguales. Esta será la hipótesis nula H_0 que se usará en todos los modelos de ANOVA, junto con los cuatro supuestos mencionados.

17.1 Introducción

Para comenzar, conviene unificar la terminología que se usará en todos los modelos de ANOVA que se irán desarrollando. En el Cuadro 17.1 siguiente, se presenta la manera de volcar los datos recopilados. Para simplificar los cálculos se supone que se ha seguido la regla de oro: *Todas las muestras tienen el mismo tamaño*. El lector interesado en desarrollar las ecuaciones para muestras de distinto tamaño, puede recurrir a los textos, como el de Sokal, para analizar los cambios. Pero en el presente trabajo, se intenta reducir al mínimo la terminología matemática para facilitar la comprensión del lector no avezado.

Cuadro 17.1 Presentación de datos en una ANOVA simple

N°	Factor A						
	Grupo 1	Grupo 2	Grupo 3	...	Grupo j	...	Grupo a
1	X11	X12	X13	...	X1j	...	X1a
2	X21	X22	X23	...	X2j	...	X2a
3	X31	X32	X33	...	X3j	...	X3a
.
i	Xi1	Xi2	Xi3	...	Xij	...	Xia
.
n	Xn1	Xn2	Xn3	...	Xnj	...	Xna
Total	T.1	T.2	T.3	...	T.j	...	T.a

Este modelo es de tipo unifactorial, es decir se emplea para estudiar el efecto de un cierto Factor A sobre las muestras tomadas. A cada muestra se la denomina grupo. Hay un total de **a** grupos. En cada grupo se han realizado **n** determinaciones de la magnitud clínica **X**. Además,

X_{ij} : es la determinación número *i* del grupo *j*

T_j : es la suma de las *n* observaciones correspondientes al grupo *j*

$\bar{X}_j = T_j / n$: es el valor promedio del grupo *j*.

$T = T.1 + T.2 + T.3 + \dots + T.j + \dots + T.a$: es el total de todas las observaciones X_{ij} hechas.

$N = n \cdot a$: es el número total de datos medidos.

$$\bar{X} = T/N = \sum_1^a \bar{X}_j / a = \sum_1^n \sum_1^a X_{ij} / N : \text{ es el promedio general de todos los datos.}$$

Para los totales, se usa un punto para denotar que se ha sumado a lo largo del subíndice i. Por ejemplo, T.2 es el total de los datos de la muestra número 2, o sea: T.2 = X12 + X22 + ... + Xn2. En estas condiciones se puede calcular la varianza total con :

$$DS^2 = SS_T / v_T = \sum_1^n \sum_1^a (X_{ij} - \bar{X})^2 / (N - 1)$$

La varianza total de los datos es el cociente entre la suma de los cuadrados totales SS_T y los grados de libertad totales $v_T = N - 1$. Mientras que la suma de cuadrados totales SS_T , es la sumatoria para N datos, del cuadrado de las diferencias, entre cada valor y su promedio general.

El motivo para escribir la varianza como un cociente entre la suma de los cuadrados y los grados de libertad es que ambos términos pueden ser divididos en dos partes. Una de esas partes se explica por el efecto del factor analizado y la otra parte es la inexplicada, debida al error aleatorio de las mediciones. En efecto :

$$v_T = N - 1 = N - a + a - 1 = (N - a) + (a - 1) = v_D + v_E$$

$v_E = a - 1$: Son los grados de libertad *entre* las muestras.

$v_D = N - a$: Son los grados de libertad *dentro* de las muestras.

Por su parte se puede describir la suma de cuadrados como:

$$SS_T = \sum \sum (X_{ij} - \bar{X})^2 = \sum \sum [(X_{ij} - \bar{X}_j) + (\bar{X}_j - \bar{X})]^2$$

Efectuando el desarrollo del cuadrado y simplificando queda :

$$SS_T = \sum \sum (X_{ij} - \bar{X}_j)^2 + \sum \sum (\bar{X}_j - \bar{X})^2$$

$$SS_T = SS_D + SS_E$$

La suma de cuadrados totales, se ha particionado en dos términos, la suma de cuadrados dentro de las muestras SS_D , calculada como la sumatoria del cuadrado de las diferencias entre cada observación y el promedio del grupo al cual pertenece y la suma de cuadrados entre las muestras SS_E , calculada como la sumatoria del cuadrado de las diferencias entre los promedios de cada grupo y el promedio general o media muestral.

Como se puede ver la SS_D , tiene un total de N términos cuadráticos libres, pero hay una cantidad de relaciones en el calculo de las medias grupales que le restan a grados de libertad, por lo tanto los grados de libertad de esa suma de cuadrados será $v_D = (N - a)$. Por su parte, como hay a medias grupales, la SS_E tendrá a términos cuadráticos libres, a los que se le debe restar un término, por la relación de vinculo entre ellos con la media general. O sea, $v_E = (a - 1)$.

Si se divide las sumas de los cuadrados por sus grados de libertad respectivos, se obtiene una nueva cantidad con todas las características de una varianza, denominada cuadrados medios, que se denota: **MS** (del término inglés: **Mean of Squares**).

Luego, los nuevos estadígrafos hallados son :

$MS_E = SS_E / v_E$: Son los cuadrados medios *entre* las muestras.

$MS_D = SS_D / v_D$: Son los cuadrados medios *dentro* de las muestras.

Notar que, la suma de ambos términos, **no es igual** al cuadrado medio total. Los que se particionan en dos partes son las SS y los grados de libertad, pero no los MS. Y allí reside el principal hecho que permite el estudio con los modelos de ANOVA; una medida cuantitativa de la variabilidad de las mediciones se divide en dos partes: Una atribuible al efecto del factor que se desea investigar y la otra, es el remanente, inexplicada, atribuida al error de medición.

R. A. Fisher demostró que el cociente entre estos dos estadígrafos tiene una distribución F y por lo tanto se puede plantear un test de hipótesis de la manera siguiente:

Se calcula: $F = MS_E / MS_D$ y se lo compara con un valor crítico de tablas $F_{\alpha; (a-1); (N-a)}$

Cuando $F > F_{\alpha}$ se rechaza la H_0 planteada, de que todas las muestras provienen de la misma población. En la forma habitual vista en los capítulos anteriores para testear la hipótesis nula, de que todas las medias son iguales. Si se rechaza la hipótesis nula, entonces se debe continuar para averiguar en cuál, o cuales, muestras hay diferencias significativas. Esto se verá en el Tema 18.

17.2 Formas cortas de cálculo

Las ecuaciones anteriores parecen complicadas, sin embargo hay maneras de simplificar los cálculos para obtener el estadígrafo principal F, como se muestra a continuación:

$$SS_T = \sum \sum (X_{ij} - \bar{X})^2 = \sum \sum (X_{ij} - (T/N))^2 = \sum \sum (X_{ij})^2 - (T^2 / N)$$

Así, los cuadrados medios totales, se pueden calcular como la sumatoria de los cuadrados de todas las observaciones realizadas, menos el *término de corrección*: T^2 / N , que se calcula como el cuadrado del total de las observaciones, dividido el tamaño muestral del experimento. Por su parte, se puede simplificar el:

$$SS_E = \sum \sum (\bar{X}_{.j} - \bar{X})^2 = \sum \sum [a.T.j - (T/N)]^2 = (1/n) \sum_1^a (T.j)^2 - (T^2 / N)$$

Así, los cuadrados medios entre las muestras, se puede calcular como la sumatoria de los cuadrados de los totales muestrales dividido su tamaño muestral, menos el término de corrección. Finalmente, la suma de cuadrados entre las muestras se obtiene como la diferencia de los dos términos anteriores.

Entonces, ya se está en condiciones de armar el denominado: “Cuadro de ANOVA”, volcando en el mismo los datos principales:

Cuadro 17.3 : Cuadro de ANOVA.

Variabilidad	Suma de cuadrados	Grados de Libertad	Cuadrados medios	$F = MS_E / MS_D$
Entre los Grupos	SS_E	v_E	MS_E	
Dentro de los grupos	SS_D	v_D	MS_D	
Total	SS_T	v_T		

Los diez pasos a seguir para la simplificación de los cálculos usando las relaciones anteriores son las siguientes:

Paso 1) Se calcula el gran total T como la suma de todas las observaciones.

Paso 2) Se obtiene el llamado término de corrección T^2 / N

Paso 3) Se calcula la suma de los cuadrados de todas las observaciones realizadas:

$$T_x^2 = \sum_1^n \sum_1^a (X_{ij})^2$$

Paso 4) Se calcula la suma de los cuadrados de los totales grupales $T.j$ y se divide por n:

$$T_G^2 = (1/n) \sum_1^a (T.j)^2$$

Paso 5) Se obtiene el valor de SS_E restando el término de corrección al valor calculado en el caso anterior. O sea: $SS_E = Paso 4 - Paso 2$.

Paso 6) Se obtiene el valor de SS_T restando a la cantidad obtenida en el Paso 3 el término de corrección. O sea: $SS_T = Paso 3 - Paso 2$.

Paso 7) Se calcula el valor de $SS_D = SS_T - SS_E$.
 O sea: $SS_D = Paso 6 - Paso 5$.

Paso 8) Se calculan los grados de libertad respectivos y se completa el cuadro de ANOVA.

Paso 9) Se calcula el estadígrafo $F = MS_E / MS_D$

Paso 10) Se obtienen de las tablas los valores críticos para los niveles de significación buscados (95%, 99% y 99,9%) y se decide si se acepta o rechaza la hipótesis nula, como es usual.

En estos diez pasos se resume el método de cálculo de F, a cuentas sencillas y rápidas de efectuar. Sin embargo, existen programas de computadoras específicos para resolver un ANOVA, tales como el Statistics, Stat-Graphics, Mat-Acad, etc. Por su parte, una planilla de cálculo cualquiera como Lotus o Excel, permiten programar los pasos anteriores a fin de simplificar el trabajo de cálculo.

17.3 Aplicación en Control de Calidad

Para ilustrar los conceptos anteriores se presentan ejemplos de aplicación, relacionados con el tema del Control de Calidad. La idea es practicar las fórmulas anteriores y comparar con las técnicas relacionadas presentadas en los capítulos precedentes.

Ejemplo 1) En un sistema Regional de Control de Calidad, con 15 laboratorios afiliados se quiere investigar las fluctuaciones entre 3 diferentes maneras de medir RGR (Recuento de Glóbulos Rojos). La primera forma es usando equipos automatizados de recuento, o contadores hematológicos, tales como el Technicon H301 y similares. De entre todos los afiliados que usen ese método se eligen cinco de ellos al azar para conformar el Grupo 1. La segunda manera es usando el método microhematocrito y para ello, se eligen al azar, otros cinco afiliados que usan tal método para conformar el Grupo 2. Finalmente, el Grupo 3 se conforma con otros cinco laboratorios seleccionados al azar, de entre los que usan otros métodos, como por ejemplo el macrohematocrito, recuento en cámara, etc. Los 15 laboratorios siguen un programa de Control de Calidad interno y se suponen calibrados. Se envía a cada laboratorio una muestra ciega, con una sangre calibrada en el laboratorio de referencia de: $\mu \in (2,9467 \pm 0,0004) \cdot 10^6$ gl/ml.

Los valores observados se muestran en el cuadro 19.3. Se necesita sacar algunas conclusiones de los datos obtenidos.

Cuadro 17.4 : Valores de RGR expresados en 10^6 gl/ml

N°	Grupo1	Grupo2	Grupo3
1	2,94	3,01	3,30
2	2,96	3,10	3,20
3	2,95	3,04	3,20
4	2,96	3,06	3,10
5	2,94	3,04	3,30
Total	14,75	15,25	16,10
Media	2,95	3,05	3,22
$\sum x^2$	43,5129	46,5169	51,87
DS	0,01	0,03317	0,08367

Para usar este modelo, se comienza con los diez pasos descriptos más arriba hasta conseguir el resultado buscado:

Paso 1) Se calcula el gran total $T = 14,75 + 15,25 + 16,10 = 46,1$

Paso 2) Se obtiene el llamado término de corrección $T^2 / N = (46,1)^2 / 15 = 141,68066$

Notar que *no se ha redondeado* el resultado. Como se verá a continuación, si se realiza una cosa así, se pierden las sutiles diferencias con la que trabaja el modelo.

Paso 3) Se calcula la suma de los cuadrados de todas las observaciones realizadas. O sea:

$$T_x^2 = \sum_1^n \sum_1^a (X_{ij})^2 = (2,94)^2 + (2,96)^2 + \dots + (3,30)^2 = 141,8998$$

Paso 4) Se calcula la suma de los cuadrados de los totales grupales T_j y se divide por n :

$$T_G^2 = \sum_1^a (T_j)^2 / n = (1/5) [(14,75)^2 + (15,25)^2 + (16,1)^2] = 141,867$$

Paso 5) Se obtiene el valor de SS_E restando a la cantidad obtenida en el Paso 4 el término de corrección.

$$SS_E = 141,867 - 141,68066 = 0,18634$$

Paso 6) Se obtiene el valor de SS_T restando a la cantidad obtenida en el Paso 3 el término de corrección.

$$SS_T = 141,8998 - 141,68066 = 0,21914$$

Paso 7) Se calcula el valor de $SS_D = SS_T - SS_E$.

$$SS_D = 0,21914 - 0,18634 = 0,0328$$

Paso 8) Se calculan los grados de libertad respectivos y se completa el cuadro de ANOVA

$$v_T = N - 1 = 14 ; v_D = (N - a) = 12 ; v_E = (a - 1) = 2$$

Variabilidad	SS	v	MS
Entre grupos	0,18634	2	0,09317
Dentro grupos	0,0328	12	0,0027333
Total	0,21914	14	

$$F = 34,1***$$

Paso 9) Se calcula el estadígrafo de comparación con $F = MS_E / MS_D = 34,1***$

Paso 10) Se obtiene de tablas los valores críticos $F_{0,95;2;12} = 3,89$ $F_{0,999;2;12} = 12,967$

La conclusión es que debe rechazarse la hipótesis nula. La muestra no proviene de la misma población. Para los enunciados se concluye que las tres técnicas de laboratorios arrojan valores diferentes. Pero para poder comparar entre sí, los grupos testeados se necesitan más conceptos del ANOVA, que permiten avanzar un poco más en la investigación. Con este análisis, solo se sabe que hay una diferencia entre los tres grupos testeados, pero no se puede determinar entre quienes.

Como a estas alturas no se dispone de los modelos adecuados para poder efectuar las comparaciones entre grupos, que se verán en el próximo capítulo, todo lo que se puede aplicar es lo visto en pequeñas muestras. Esto es, se puede proceder a realizar comparaciones de dos muestras entre sí, con el modelo de Student para muestras independientes. Entonces, habrá tres comparaciones posibles: Grupo 1 vs. Grupo 2; Grupo 2 vs. Grupo 3 y Grupo 1 vs. Grupo 3.

Cuadro 17.5 : Comparaciones de muestras entre sí.

Caso 1) Comparación del Grupo 1 versus el Grupo 2:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{DS_1^2}{n_1}\right) + \left(\frac{DS_2^2}{n_2}\right)}} = \frac{(2,95 - 3,05) - 0}{\sqrt{\left(\frac{(0,01)^2}{5}\right) + \left(\frac{(0,0332)^2}{5}\right)}} = -6,45***$$

Pues $t_{\alpha;v} = t_{0,999;8} = 5,041$: La conclusión es que ambas técnicas arrojan diferentes resultados y sus diferencias son altamente significativas.

Caso 2) Comparación del Grupo 3 versus el Grupo 2:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{DS_1^2}{n_1}\right) + \left(\frac{DS_2^2}{n_2}\right)}} = \frac{(3,22 - 3,05) - 0}{\sqrt{\left(\frac{(0,08367)^2}{5}\right) + \left(\frac{(0,0332)^2}{5}\right)}} = +4,25**$$

Pues $t_{\alpha;v} = t_{0,99;8} = 3,36$: La conclusión es que ambas técnicas arrojan diferentes resultados y sus diferencias son muy significativas.

Caso 3) Comparación del Grupo 1 versus el Grupo 3:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{DS_1^2}{n_1}\right) + \left(\frac{DS_2^2}{n_2}\right)}} = \frac{(2,95 - 3,22) - 0}{\sqrt{\left(\frac{(0,01)^2}{5}\right) + \left(\frac{(0,08367)^2}{5}\right)}} = -7,16***$$

Pues $t_{\alpha;v} = t_{0,999;8} = 5,041$: La conclusión es que ambas técnicas arrojan diferentes resultados y sus diferencias son altamente significativas.

Conclusión: Los 3 grupos difieren entre sí.

17.3.1 Control de exactitud

Para controlar la exactitud de una técnica clínica o industrial, se necesita de un patrón. Con este valor se pueden comparar los valores promedios de cada grupo analizado, para ver si cada uno está calibrado. El procedimiento es sencillo y se puede realizar de dos maneras. La primera es realizar un test de hipótesis con el modelo Student. La segunda es determinar el intervalo de confianza de cada valor promedio encontrado, para cada grupo, y ver si el valor patrón cae dentro o fuera del mismo. Cuando no caiga dentro hay evidencia como para pensar que el sistema está descalibrado, hay un error de tipo sistemático que también se puede calcular.

Ejemplo 2) Con los datos del ejemplo anterior determinar en forma analítica el control de exactitud que se le puede realizar a los tres grupos investigados, con un test y con un intervalo.

Método 1) Calibración con el modelo Student

H_0 : El método está calibrado : $\mu = \bar{X}$

H_1 : El método no está calibrado: $\mu \neq \bar{X}$

$$t = \frac{(\bar{x} - \mu)}{DS/\sqrt{n}} \text{ versus } t_{\alpha; v}$$

Grupo 1: Contador Hematológico.

$$t = \frac{(\bar{x} - \mu)}{DS/\sqrt{n}} = \frac{(2,95 - 2,9467)}{0,01/\sqrt{5}} = 0,738 \text{ versus } t_{\alpha; v} = t_{0,95; 4} = 2,776$$

No hay evidencia como para rechazar la H_0 . Como el resultado no fue significativo, se acepta la hipótesis nula que el Grupo 1 está calibrado, con respecto al laboratorio de referencia. A su vez, se puede proceder con cada dato individual. El programa regional, le envía esta información a cada uno de ellos, y le suele agregar un ranking. Así, el laboratorio que envió el dato número 3 resultara el primero de los 15, pues es el más cercano al valor control; mientras los que enviaron los datos número 1 y 5 salen en el segundo puesto y así sucesivamente. Para el investigador significa que su técnica de rutina debería ser la que utiliza el contador hematológico, por ser la más exacta de todas y estar calibrada. Esto se deduce, comparando las medias muestrales halladas.

Grupo 2: Microhematocrito.

$$t = \frac{(\bar{x} - \mu)}{DS/\sqrt{n}} = \frac{(3,05 - 2,9467)}{0,0332/\sqrt{5}} = 6,96^{**} \text{ versus } t_{\alpha; v} = t_{0,99; 4} = 4,604 \text{ y } t_{0,999; 4} = 8,610$$

Hay evidencia como para rechazar la H_0 al (99% de confianza). Como el resultado fue muy significativo, se rechaza la hipótesis de que el grupo 2 está calibrado. Se concluye que hay un error sistemático con respecto al laboratorio de referencia. Esto significa que el programa regional debe informar a los 5 laboratorios del grupo número 2 de esta situación, además de darles el ranking obtenido a cada uno de ellos, deberá arbitrar las soluciones del caso. Si no se usa una sangre calibrada, la única conclusión a sacar es que no es lo mismo una técnica que la otra, pero cuando se tiene el valor patrón, para el responsable de cada laboratorio de ese grupo, significa que se ha descubierto y probado un error sistemático que afecta a su técnica usual, evaluado en $ES = 3,05 - 2,9467 = 0,1033 \cdot 10^6$ gl/ml. Luego le quedan dos caminos: El más rápido y sencillo es corregir los resultados que logre, restando el valor del ES a cada medición efectuada con esa técnica. El segundo camino, más científico, es ponerse a investigar las causas que originan ese tipo de error, y así poder corregir su técnica.

Grupo 3: Otras técnicas.

Para los laboratorios que no emplean alguna de las dos técnicas anteriores resulta:

$$t = \frac{(\bar{x} - \mu)}{DS/\sqrt{n}} = \frac{(3,22 - 2,9467)}{0,08367/\sqrt{5}} = 7,303^{**} \text{ versus } t_{\alpha;v} = t_{0,99;4} = 4,604 \text{ y } t_{0,999;4} = 8,610$$

Hay evidencia como para rechazar la Ho al (99% de confianza). Como el resultado fue muy significativo, se rechaza la hipótesis nula de que el grupo 3 está calibrado. Se concluye que los 5 laboratorios que usan otras técnicas, como el recuento en cámara o el Macrohematocrito, deben ser informados de la gran diferencia que tienen con el laboratorio de referencia, se ha encontrado evidencia muy fuerte de un ES = 0,2733 gl/ml. Y para responsable de cada uno de estos 5 laboratorios del Grupo 3, significa que la técnica usual, no solo debe ser desechada como confiable, sino que es la peor de todas. Como era de esperar, de acuerdo a la bibliografía y por ello se desaconseja su uso en el laboratorio moderno.

Método 2) Calibración con intervalos de confianza (95%)

Ho : El método está calibrado: $\mu \in (\bar{X} \pm \Delta X)$ Con $\Delta X = t_{0,95;4} \cdot DS/\sqrt{n}$
 H1 : El método no está calibrado: $\mu \notin (\bar{X} \pm \Delta X)$ y $\mu = 2,9467$

Grupo	\bar{X}	ΔX	Límite Superior	Límite inferior	μ cae
1	2,95	0,012	2,962	2,938	Adentro
2	3,05	0,041	3,091	3,009	Fuera
3	3,22	0,104	3,324	3,116	Fuera

Se rechaza la Ho para los dos últimos casos.

17.3.2 Control de precisión

Para controlar la precisión de una técnica clínica o industrial, se necesita de un valor referencial de la *desviación máxima admisible* ($\sigma_{m\acute{a}x}$). En Bioquímica, tales valores se toman de algunos criterios existentes, tales como: usar el error relativo máximo, el criterio de Thonks, el de Aspen, etc. (que se verán más adelante). O bien, en los libros de texto de Análisis Clínicos. En Farmacia, se usan los mismos en los Laboratorios de Control de Calidad industriales. Pero, cuando se trata de sistemas de producción, los límites vienen especificados en el protocolo de fabricación del medicamento. En resumen se necesita definir primero un $\sigma_{m\acute{a}x}$ y luego efectuar un test de hipótesis con la distribución de Chi Cuadrado. El planteo es:

Ho : El sistema tiene una precisión aceptable: $DS \leq \sigma_{m\acute{a}x}$
 H1 : El sistema no tiene una precisión aceptable: $DS > \sigma_{m\acute{a}x}$

Lo que se testea con: $\chi^2 = (n - 1) DS^2 / \sigma_{m\acute{a}x}^2$ versus $\chi^2_{\alpha;v}$

Para ilustrar estas ideas se desarrolla el siguiente ejemplo, donde se usan los datos del ejemplo de calibración anterior:

Ejemplo) Usar los datos del Cuadro 17.4, con un CV% $_{m\acute{a}x}$ del 5%, para decidir si los Grupos tienen una precisión aceptable.

Grupo	DS	$\sigma_{m\acute{a}x} = 0,05 \cdot \mu$	χ^2	$\chi^2_{0,95;4}$	Resultado
1	0,01	0,1473	0,018	9,488	No se rechaza Ho
2	0,03317		0,203		No se rechaza Ho
3	0,08367		1,299		No se rechaza Ho

Conclusión: Todos los grupos tienen una precisión aceptable.

17.3.3 Control del factor humano

Para controlar la influencia del factor humano en las mediciones clínicas se debe hacer medir lo mismo a varios operadores diferentes. Así, se pueden comparar los valores medidos por cada uno, entre sí, con un modelo de Anova. En el ejemplo siguiente se usa un modelo de un factor para ilustrar el método, sin embargo, cuanto más factores se tomen en cuenta, mejor será la sensibilidad del modelo estadístico para detectar las diferencias.

Cuadro 17.5.1 : Comparaciones de 3 operadores distintos

1	27	27	20	18	Σ
	32	30	22	20	
	30	28	22	21	
	29	30	20	18	
	30	32	22	20	
	35	30	23	20	
					606
2	30	24	22	22	
	30	30	22	24	
	30	30	23	23	
	30	30	23	20	
	30	30	24	24	
	30	30	25	22	
					628
3	30	26	19	22	
	31	32	22	21	
	32	30	22	21	
	36	36	22	19	
	34	31	25	24	
	30	29	24	23	
$\Sigma \Sigma$					641
					1875

Fuente: Dras. Cech, N. ; Coschiza, M. y Lodeiro, N.

Ejemplo 1) Se desea saber si hay diferencias entre los tres operadores de un laboratorio en la medición de Antibiogramas. Lo que se mide es el diámetro del halo que deja cada antibiótico en la caja Petri donde se efectuó el cultivo. Los datos que se muestran en el Cuadro 17.5.1 corresponden a 6 pacientes diferentes, empleando 4 antibióticos para cada uno de ellos. Acá se resuelve el problema usando un modelo de 1 factor, pues no se toman en cuenta los diferentes antibióticos. Es como si se tuviesen 26 diámetros a medir, por los 3 operadores, con lo que se obtienen un total de 72 datos.

Paso 1) Se calcula la suma total de datos $T = \sum_1^a T_c = 1.875$

Paso 2) Se calcula la suma del cuadrado de todos los datos: $T_x^2 = \sum_1^a \sum_1^b X_{ijk}^2 = 50.459$

Paso 3) Se calcula el término de corrección $T^2 / N = (1.875)^2 / 72 = 48.828$

Paso 4) Se calcula la suma de los cuadrados entre los operadores:

$$T_E^2 = \sum_1^a \sum_1^n T_{ik}^2 / a.n = 48.854$$

Paso 6) Se calcula la suma de los cuadrados respectivos:

$$SS_{total} = T^2 - T^2 / N = Paso 2 - Paso 3 = 1.631$$

$$SS_{entre} = T_E^2 - T^2 / N = Paso 4 - Paso 3 = 26 \text{ (operadores)}$$

$$SS_{dentro} = SS_{total} - SS_{entre} = 1.631 - 26 = 1.605$$

Paso 7) Se calcula el cuadro de ANOVA como sigue:

Tabla 20.2: Cuadro de ANOVA para los datos de Tabla 20.1

Fuente de Variación	Suma de cuadrados	Grados de Libertad	Cuadrados Medios	F
Entre Filas (operadores)	26	2	13	0,56 (ns)
Dentro grupos (error)	1.605	69	23,26	
Total	1.631	71		

Los resultados no fueron significativos, por lo que se concluye que no hay diferencias entre la forma de medir de los tres operadores. Es decir, el factor humano no influye en los datos.

17.4 Modelos I y II del ANOVA

De acuerdo a la forma en que fue diseñado el experimento, es diferente la manera de proseguir usando el modelo de ANOVA. Hay dos modelos posibles, los cuales fueron descriptos por primera vez por Eisenhart en 1947.

Modelo I : se supone que el efecto de factor analizado es *constante* en cada grupo.

Modelo II : se supone que el efecto del factor analizado es *aleatorio* en cada grupo.

Hasta acá, tanto el Modelo I como el Modelo II, usan los mismos cálculos y se llega al Cuadro de ANOVA. Si no se encuentran diferencias significativas entre las muestras, no tiene sentido seguir adelante. Pero, si se rechaza la H_0 , entonces el problema siguiente es descubrir donde las diferencias se hacen significativas.

Cuando el investigador selecciona los diferentes grupos a investigar siguiendo un criterio determinado, quiere decir que cada grupo es elegido a propósito. En cuyo caso, se supone que el efecto del factor en cada grupo es una constante, que lo afecta por igual en todas sus mediciones. Aunque, esa constante puede variar de grupo a grupo, y ese efecto justamente es lo que se está tratando de descubrir. Es una especie de error sistemático actuando sobre el grupo y modificando su media de manera tal que la diferencia con el promedio general se hace significativo. Por esto, es el investigador quien opta por el Modelo I, al elegir cada grupo a su modo.

En cambio, cuando el investigador selecciona al azar, entre todos los casos posibles, a los grupos de trabajo, entonces se trata de un Modelo II y se supone que el efecto del factor en cada grupo es de tipo aleatorio. Por ejemplo, si un farmacéutico desea investigar el efecto de los antibióticos sobre cierta enfermedad, y mediante un sorteo, selecciona a 5 de ellos, entre los vigentes en el mercado, entonces se trata de un Modelo II. Si en cambio, tal selección la hace buscando a los que contengan sulfato de Neomicina, de 5 marcas diferentes, entonces se trata de un Modelo I porque tal selección no la hizo al azar.

Cuando se realiza un diseño en Modelo II, generalmente se trata de estudios poblacionales, donde se trata de determinar si el factor de análisis, produce variabilidad en las muestras. El problema allí, es calcular el porcentaje de incremento en la varianza, que es la medida cuantitativa de la dispersión en general. La varianza total, se descompone en dos términos, uno debida al factor y otro remanente, resto o error. El porcentaje de varianza debida al factor se denomina: *Componente Añadida de Varianza* y en el próximo capítulo se verá como calcularla.

Cuando se efectúa un experimento diseñado como Modelo I, es porque se está interesado en determinar y cuantificar, el efecto constante del factor en la muestra. De lo que se trata en tal caso, es de poder comparar entre sí las medias muestrales para detectar si hay diferencia significativa entre ellas. En los ejemplos vistos en el punto de calibración anterior, la elección de los tres grupos no se hizo al azar, sino que se eligieron a propósito y por lo tanto se trataba de un Modelo I. Las comparaciones entre medias se hicieron con el modelo Student, el que solo puede comparar de a dos muestras cada vez. Entonces, el camino fue hacer todas las comparaciones posibles de a pares. Sin embargo, cuando la cantidad de muestras es grande, este proceder puede ser muy engorroso y difícil de entender. Para estos casos, hay una forma más elegante, basada en el ANOVA, denominada: *Comparaciones Múltiples*. Objeto central del próximo capítulo.

Para entender mejor la utilidad de estos conceptos, conviene hacerse una pregunta básica: ¿ Por que el segundo valor del primer grupo del ejemplo anterior fue 2,96 en lugar de cualquier otra cantidad ? Una primera aproximación a la respuesta es pensar que esto se debe a que fue extraído de una población con una media $\mu \in (2,9467 \pm 0,0004) \cdot 10^6$ gl/ml, pero si se desconoce tal valor, se puede usar su mejor estimación:

$\bar{X} = T / N = 3,07 \cdot 10^6$ gl/ml. Luego como primera aproximación se puede imaginar la relación:

$$X_{21} \cong \mu$$

Pero en este caso el valor de X_{21} es 2,96 y $\mu = 3,07$ (o con sangre calibrada $\mu = 2,9467$). Hay una diferencia de 0,11 unidades que sigue existiendo. Para el Modelo I, se puede efectuar una segunda aproximación y pensar que ese valor forma parte del grupo 1 cuya media es de 2,95. Luego el efecto de medir con un contador automático provoca en las mediciones una diferencia constante α_1 , para el grupo 1 de $(3,07 - 2,95)$. O sea, un valor negativo $\alpha_1 = -0,12$ a lo largo de todo el grupo. Entonces, como una segunda aproximación se puede expresar :

$$X_{21} \cong \mu + \alpha_1 = 3,07 - 0,12 = 2,95$$

Todavía hay una diferencia de 0,01. Pues $X_{21} = 2,96$ y con la segunda aproximación se lo estima en 2,95. La pregunta siguiente es ¿ Por qué este valor individual no puede ser totalmente explicado ? La respuesta es: porque hay una variabilidad inherente de tipo genético o ambiental que afecta el valor medido. Pero, si se trabaja con una misma sangre se ha acotado la variabilidad genética, que aparece en las mediciones biológicas entre individuos. Queda todavía la variabilidad ambiental, para acotarla se puede realizar todas las mediciones en un mismo ámbito, con temperaturas y humedad controladas. En el ejemplo visto: en caso de trabajar con varios laboratorios, ubicados en diferentes lugares físicos; no hay manera de reducir la variabilidad ambiental. De todas maneras, se podría haber establecido a los laboratorios intervinientes en el experimento, las condiciones ambientales, tales como humedad ambiente, temperatura, alimentación eléctrica regulada, etc. De manera tal, de dejar constantes esos valores y acotar la variabilidad ambiental. Pero, siempre hay algo incontrolable, por más que trabaje en las condiciones asépticas de un quirófano, todavía falta fijar variaciones propias de mediciones repetidas. Y en el hipotético caso, de fijar ese tipo de variabilidad, todavía queda el irremediable error de medición. Entonces, la tercera aproximación y final que se puede plantear es:

$$X_{21} = \mu + \alpha_1 + \sigma_{ij}$$

Donde se supone que el tercer término agregado, es un valor particular que adopta una función error de medición, de tipo gaussiana con parámetros $(0 ; \sigma_{ij})$. Entonces, se puede expresar el Modelo I en términos generales con:

Modelo I de ANOVA :

$$X_{ij} = \mu + \alpha_1 + \sigma_{ij}$$

Esto ya es una mejor explicación, para saber porque se produce una medida en particular y arroja cierta luz sobre la naturaleza de las mediciones. El ANOVA permite cuantificar estas ideas para poder trabajar en optimizaciones.

En un Modelo II, se supone que el efecto del factor en el grupo es de tipo aleatorio, en lugar de constante. Entonces, en lugar de poner un valor constante α_i dentro de un grupo cualquier

ra, se plantea un efecto de tipo aleatorio dado por A_i , que tiene una distribución gaussiana de parámetros $(0 ; \sigma_A)$

Modelo II de ANOVA :

$$X_{ij} = \mu + A_i + \sigma_{ij}$$

17.5 Supuestos básicos del ANOVA

De acuerdo a lo visto, para un modelo unifactorial de ANOVA se necesitan hacer los cuatro supuestos siguientes:

- 1) *Las muestras deben ser aleatorias.*
- 2) *Las muestras deben ser independientes.*
- 3) *La población de donde son extraídas debe ser gaussiana.*
- 4) *Sus varianzas deben ser iguales.*

Si se verifican estos supuestos, entonces con hipótesis nula de igualdad de medias muestrales se testea si todas las muestras fueron extraídas de la misma población, contra la hipótesis alternativa de que el efecto del factor analizado es significativo.

17.6 Modelo no paramétrico de Kruskal-Wallis

Cuando no se verifique alguno de los cuatro supuestos básicos del Anova (normalidad, aleatoriedad, independencia y homoscedasticidad), o bien, la magnitud estudiada no sea continua, se tiene el recurso de usar un modelo no paramétrico equivalente: *Modelo de Kruskal-Wallis*.

Este tipo de modelo se hizo muy popular debido a que los cálculos son muy sencillos y no hay que preocuparse por las hipótesis del Anova. Aunque claro está, el Anova tiene mayor efectividad para detectar las diferencias entre las medias muestrales. La manera de proceder con el modelo de Kuskal-Wallis es muy similar al de la U de Mann y Whitney, se trata de ordenar primero los datos en forma creciente, ignorando la división en grupos. Se le asigna un rango a cada dato y en el caso de empates o "ligas" se usa el rango promedio entre ellos. Luego se reemplaza cada dato de la tabla original, por su rango equivalente y con los nuevos datos se calcula el estadígrafo H de Kruskal-Wallis, cuya función distribución se muestra en la Tabla 17 del Anexo.

El estadígrafo H depende de dos cosas: el rango de cada uno de los valores obtenidos y del tamaño muestral respectivo. Su cálculo es algo engorroso pero muy sencillo de efectuar como se muestra en el ejemplo siguiente.

$$H = \left[\frac{12}{\sum_1^a n_i (\sum_1^a n_i + 1)} \sum_1^a \frac{(\sum_1^{n_i} R)_i^2}{n_i} \right] - 3 \left(\sum_1^a n_i + 1 \right)$$

Si se tiene en cuenta que para el caso de muestras iguales $N = \sum_1^a n_i = a \cdot n$ resulta:

$$H = \left[\frac{12}{N(N+1)} \sum_1^a \frac{(\sum_1^{n_i} R)_i^2}{n} \right] - 3(N+1)$$

Para ilustrar este modelo se toman los datos del Cuadro 17.6 de un estudio de diferentes drogas para el tratamiento de una enfermedad, tanto puras como mezcladas y un control o placebo. Se mide el tiempo de curación en días, desde que los pacientes comienzan el tratamiento.

Cuadro 17.6: Estudio comparativo de las drogas A, B y C.

Nº	Placebo	Droga A	Droga B	Mezcla A+B	Droga C
1	125	107	108	108	112
2	117	108	111	109	116
3	120	110	106	108	115
4	125	109	108	111	113
5	115	112	107	107	114
6	121	110	106	106	112
7	117	110	111	108	115
8	117	107	110	107	115
9	126	109	107	107	112
10	118	111	108	109	117

Fuente: Ejemplo derivado del libro de Sokal y Rohlf

Los pasos a seguir para resolver este modelo son los siguientes:

Paso 1) Se calculan los rangos de cada valor obtenido:

valor	rango	R	valor	rango	R	valor	rango	R	valor	rango	R	valor	rango	R
106	1	2	108	11	14	109	21	19,5	112	31	31,5	117	41	42,5
106	2	2	108	12	14	110	22	23,5	112	32	31,5	117	42	42,5
106	3	2	108	13	14	110	23	23,5	112	33	31,5	117	43	42,5
107	4	7	108	14	14	110	24	23,5	113	34	34	117	44	42,5
107	5	7	108	15	14	110	25	23,5	114	35	35	118	45	45
107	6	7	108	16	14	111	26	27,5	115	36	37,5	120	46	46
107	7	7	108	17	14	111	27	27,5	115	37	37,5	121	47	47
107	8	7	109	18	19,5	111	28	27,5	115	38	37,5	125	48	48,5
107	9	7	109	19	19,5	111	29	27,5	115	39	37,5	125	49	48,5
107	10	7	109	20	19,5	112	30	31,5	116	40	40,0	126	50	50

El tiempo de cura medido en días se ordena de menor a mayor, sin tomar en cuenta a que grupo pertenece cada dato. El valor mínimo obtenido fue de 106 días y el mayor de 126. Estos valores se vuelcan a un cuadro como el de más abajo, en las primeras columnas de cada bloque de 10 datos. Hay 5 bloques de 3 columnas cada uno. En la segunda columna de cada bloque se coloca el rango correspondiente: 1 para el menor y así sucesivamente hasta llegar al mayor de rango 50. En caso de empates, se asigna a los datos empatados, el rango promedio. Por ejemplo, los tres primeros datos, tienen el mismo valor 106, le tocan los rangos 1, 2 y 3, luego su promedio 2 es asignado a cada uno de ellos.

Paso 2) Luego se reemplazan los datos del cuadro original por sus respectivos rangos:

Cuadro 17.7 : Reemplazo de los datos del Cuadro 17.6 por sus rangos.

Nº	Placebo	Droga A	Droga B	Mezcla A+B	Droga C	Totales
1	48,5	7	14	14	31,5	
2	42,5	14	27,5	19,5	40	
3	46	23,5	2	14	37,5	
4	48,5	19,5	14	27,5	34	
5	37,5	31,5	7	7	35	
6	47	23,5	2	2	31,5	
7	42,5	23,5	27,5	14	37,5	
8	42,5	7	23,5	7	37,5	
9	50	19,5	7	7	31,5	
10	45	27,5	14	19,5	42,5	
Total	450	196,5	138,5	131,5	358,5	1275

Paso 3) Luego se reemplazan los datos obtenidos en la fórmula de H :

$$H = \left[\frac{12}{50(50+1)} \cdot \frac{(450)^2 + (196,5)^2 + (138,5)^2 + (131,5)^2 + (358,5)^2}{10} \right] - 3(50+1) = 38,109$$

Paso 4) Dado que se encontraron empates, el valor de H debe ser corregido dividiendo por D :

$$H_{\text{corregido}} = H / D \quad \text{y para ello D se calcula con:}$$

$$D = 1 - \frac{\sum_1^m T_j}{(N-1)N(N+1)} \quad \text{donde } T_j = (E_j^3 - E_j) \text{ con } m = 10 \text{ empates } E_j$$

En el primer empate hay 3 valores empatados en 106 días, luego es:

$$T_1 = (E_1^3 - E_1) = (3^3 - 3) = 24$$

En el segundo y tercer empate hay 7 valores empatados en 107 días, luego es:

$$T_2 = (E_2^3 - E_2) = (7^3 - 7) = 336 = T_3$$

Luego hay cuatro empates seguidos de 4 valores de 109 a 112 y otros dos en 115 y 117

$$T_4 = (E_4^3 - E_4) = (4^3 - 4) = 60 = T_5 = T_6 = T_7 = T_8 = T_9$$

Finalmente hay un solo empate de 2 valores en 125

$$T_{10} = (E_{10}^3 - E_{10}) = (2^3 - 2) = 6$$

Luego el total de los puntajes por empate es:

$$\sum_1^m T_j = 24 + 336 + 336 + 6 \cdot (60) + 6 = 1062$$

Reemplazando estos valores, el factor de corrección es:

$$D = 1 - \frac{\sum_1^m T_j}{(N-1)N(N+1)} = 1 - \frac{1062}{49(50)51} = 1 - 0,0085 = 0,99151$$

Entonces $H_{\text{corregido}} = H / D = 38,109 / 0,99151 = 38,44$

Para muestras grandes el estadígrafo H se distribuye aproximadamente según una distribución Chi cuadrado, mientras que si los tamaños muestrales son pequeños ($N < 6$) se puede emplear la distribución exacta Tabla 17 del Anexo. En este caso, como $N=50$ se formula la hipótesis nula H_0 de que los 5 grupos testeados no difieren en “localización”, es decir no hay diferencia entre ellos, y se contrasta con $\chi^2_{\alpha(v)} = \chi^2_{0,01(5-1)} = 13,277$. Como este resulta menor que el valor del estadígrafo $H_{\text{corregido}} = 38,44^{**}$ se rechaza la hipótesis nula. Se puede afirmar que hay evidencia científica de que hay diferencias entre los grupos investigados.

17.7 Comparación de varias proporciones

En los capítulos anteriores se analizó la comparación de dos proporciones en dos casos básicos. La primera era usar los modelos de Gauss y Student para hacerlo usando el error de muestreo. La segunda era usar el modelo de la Chi-cuadrado para analizar la independencia en tablas de contingencia de 2×2 . En estos casos la magnitud era dicotómica y se analizaban dos factores (Tabla de la Verdad, Tabla de contingencia, Epidemiología, etc.). Cuando más de dos proporciones deben ser comparadas, el cálculo del error estándar entre un par de proporciones, se parece mucho a las comparaciones múltiples que se verán en el próximo capítulo. Sin embargo, el test general para determinar si hay diferencias significativas entre todas las proporciones que se quieren analizar es análogo al F-test del Anova. En realidad es una extensión del análisis de la Chi-cuadrado para varias proporciones.

Suponiendo que hay k grupos de observaciones, y que en el grupo número i de tamaño n_i hay una cantidad de individuos r_i que muestran cierta característica (presente / ausente, sano / enfermo, etc.); entonces la proporción de individuos con la característica será $p_i = r_i / n_i$. Los datos se pueden presentar como en el Cuadro 17.8

Cuadro 17.8 Comparación de k proporciones.

Grupos	1	2	3	...	i	...	k	Todos los grupos
(+)	r_1	r_2	r_3	...	r_i	...	r_k	R
(-)	$n_1 - r_1$	$n_2 - r_2$	$n_3 - r_3$...	$n_i - r_i$...	$n_k - r_k$	N - R
Total	n_1	n_2	n_3	...	n_i	...	n_k	N
Proporción	p_1	p_2	p_3	...	p_i	...	p_k	$P = R / N$

Las frecuencias de esta tabla de contingencia de 2 x k son dicotómicas (+) y (-) y se muestran en las dos filas de datos. En la tercer fila están los tamaños grupales y en la cuarta las proporciones de positivos respectivas. El test de la Chi-cuadrado sigue el mismo esquema presentado antes, donde para cada una de las frecuencias observadas, se puede calcular su frecuencia esperada con el producto de los respectivos totales marginales dividido el tamaño muestral, o sea:

$E_i = (\text{Total de la fila } i \times \text{Total de columna } i) / N$ y el valor del estadígrafo con:

$$\chi^2 = \sum_1^n \frac{(O_i - E_i)^2}{E_i}$$

La H_0 es que todas las k proporciones de positivos son iguales, porque provienen de la misma población de donde fueron extraídas aleatoriamente. Se hace el test comparando el valor muestral contra el de tablas de la Chi-cuadrado, con k – 1 grados de libertad. No se requiere corrección por continuidad a menos que las frecuencias observadas sean muy pequeñas, ya que se usan los mismos totales marginales para el número de tablas individuales de 2 x 2, y el número total es muy grande. Por tanto, estos se ajustan mucho mejor a la distribución Chi-cuadrado que en cada una de las tablas individuales de 2 x 2 que la componen. La proporción promedio se puede obtener usando la media ponderada con:

$$P = R / N = \frac{\sum_1^k n_i \cdot p_i}{\sum_1^k n_i} \quad \text{y} \quad Q = 1 - P$$

La contribución de cada grupo al total de χ^2 es, por ejemplo para el grupo i, igual a:

$(r_i - P \cdot n_i)^2 [(1 / P \cdot n_i) + (1 / Q \cdot n_i)]$ Donde reemplazando y manipulando la fórmula queda:

$$\chi^2 = \sum_1^k \frac{n_i \cdot (p_i - P)^2}{P \cdot Q}$$

Esta fórmula es equivalente a la anterior vista en el Capítulo 15 y 16, pero los supuestos básicos han sido reemplazados en dos aspectos: a) la variación de las proporciones siguen una distribución binomial en lugar de la normal y b) la varianza verdadera ha sido estimada con $(P \cdot Q) / n_i$. A Medida que los tamaños muestrales aumentan, estas aproximaciones pierden importancia. Para ilustrar este procedimiento se presenta el ejemplo siguiente, tomado del libro de Armitage y Berry (página 207): Se analizaron muestras de materia fecal de 288 individuos, agrupados por edades, para ver la presencia de huevos de la *Schistosoma mansoni*.

Frecuencias observadas

Edad (años)	0-9	10-19	20-29	30-39	40-49	Total
Positivo	14	16	14	7	6	57
Negativo	87	33	66	34	11	231
Total	101	49	80	41	17	288

Frecuencias esperadas

Edad (años)	0-9	10-19	20-29	30-39	40-49	Total
Positivo	20	9,7	15,8	8,1	3,4	57
Negativo	81	39,3	64,2	32,9	13,6	231
Total	101	49	80	41	17	288

Contribución de χ^2 para cada celda ($\chi^2 = 10,36$)

Edad (años)	0-9	10-19	20-29	30-39	40-49	Total
Positivo	1,79	4,09	0,21	0,15	2,07	8,31
Negativo	0,44	1,01	0,05	0,04	0,51	2,05
Total	2,23	5,10	0,26	0,19	2,58	10,36

Hay 4 grados de libertad y el valor crítico de tablas es $\chi^2_{0,95; 4} = 9,49$; por lo tanto la H_0 debe ser rechazada. Hay diferencias significativas entre las proporciones halladas. Por otra parte se pueden calcular:

$$P = R / N = \frac{\sum_1^k n_i \cdot p_i}{\sum_1^k n_i} = 0,1979 \quad \text{y} \quad Q = 1 - P = 0,8021 \quad \text{Entonces,}$$

$$\chi^2 = \sum_1^k \frac{n_i \cdot (p_i - P)^2}{P \cdot Q} = \{101 [(14/101) - 0,1979]^2 + \dots + 17 [(6/17) - 0,1979]^2\} / [(0,1979) (0,8021)] = 10,37^*$$

La pequeña discrepancia entre los valores de χ^2 calculados por ambos métodos se deben a errores de redondeo, pero en ambos casos la H_0 debe ser rechazada. Si en lugar de la proporción de positivos, se hubiese usado la de negativos el resultado final es el mismo.

17.8 Problemas propuestos

1) Marcar la respuesta correcta a cada una de las afirmaciones siguientes, o completar la frase:

- | | | |
|--|-------|-------|
| 1) Con los modelos de ANOVA se pueden comparar más de dos muestras a la vez. | V | F |
| 2) Y también más de un factor, con varias muestras a la vez. | V | F |
| 3) Para poder aplicar los modelos de ANOVA las muestras deben ser: | | |
| 4) En ANOVA se pueden usar muestras de diferente tamaño. | V | F |
| 5) La SS total se particiona en dos partes que son: | | |
| 4) Los grados de libertad se dividen en dos partes, lo mismo que la Suma de Cuadrados. | V | F |
| 5) Los tres MS se pueden calcular con:..... | | |
| 6) Los MS entre y dentro, sumados, son igual a los MS totales. | V | F |
| 7) Con el Cuadro de ANOVA se pueden presentar los resultados en forma resumida. | V | F |
| 8) El estadígrafo de comparación F se calcula como el cociente de los MS entre y dentro. | V | F |
| 9) La distribución teórica del F es una distribución Chi Cuadrado. | V | F |
| 10) Los diez pasos a seguir para calcular el Cuadro de Anova en forma rápida son: | | |
| 11) La ventaja del ANOVA respecto a Student es que puede comparar más de 2 muestras. | V | F |
| 12) En el Modelo I de ANOVA se supone que el factor actúa en forma constante. | V | F |
| 13) En el Modelo II de ANOVA se supone que el factor actúa en forma aleatoria. | V | F |
| 14) El factor actúa constante del Modelo I, actúa por igual en todas las muestras. | V | F |
| 15) El modelo de Kruskal-Wallis es el equivalente no paramétrico del ANOVA. | V | F |

- 16) En caso de empates se debe calcular el rango promedio entre los empates o ligas. **V F**
17) La idea de K-W es reemplazar los valores originales por sus rangos y calcular un H. **V F**
18) El factor de corrección se usa siempre y vale H/D. **V F**
19) El estadígrafo H se distribuye según una función F de Fisher. **V F**
20) Para tamaños muestrales muy pequeños $N < 40$ se puede usar la Tabla 17 del Anexo. **V F**
2) Se desea investigar el efecto de los alimentos balanceados en la cría de pollos para un productor de la zona. El experimento consiste en pesar los pollos antes de comenzar y al final de un mes de pruebas. Las diferencias en peso encontradas en cada uno se muestran en la tabla siguiente. Como control se alimenta a un grupo de la forma tradicional. Se escogen al azar 10 pollos por grupo.

Nº	Control	Marca 1	Marca 2	Marca 3
1	150	207	230	221
2	160	208	235	225
3	140	210	228	219
4	135	209	240	217
5	155	212	238	225
6	151	210	226	222
7	147	220	234	223
8	137	207	225	217
9	146	209	239	215
10	138	211	237	224

Averiguar si hay diferencias significativas entre las marcas testeadas usando:

- El modelo de ANOVA.
 - El modelo de Kruskal Wallis
 - Comparar las medias usando el Modelo de Student.
 - ¿Cuál es la diferencia si las marcas hubiesen sido escogidas al azar entre las del mercado?
- 3) Resolver el problema de ejemplo 1, del punto 17.3, usando el modelo no paramétrico de Kruskal Wallis, y comparar con los resultados allí obtenidos.
- 4) Un bioquímico desea averiguar si hay diferencia entre tres marcas comerciales de kits, para efectuar la técnica de glucosa. Para ello, prepara un “pool” de sueros, lo homogeneiza bien y los fracciona en 15 alícuotas. Asigna cinco viales para cada marca y mide la glucosa para el mismo suero en forma repetida. Si la homogeneización está bien realizada, se dan por cumplidos los supuestos de ANOVA. Los resultados se muestran en la tabla siguiente. Decidir si existen diferencias significativas entre las marcas probadas en el experimento.

Medición Nº	Marca 1	Marca 2	Marca 3
1	1,04	1,10	1,21
2	1,05	1,11	1,19
3	1,06	1,09	1,18
4	1,07	1,07	1,17
5	1,05	1,08	1,21

5) Un farmacéutico desea saber si una promoción produce algún efecto en sus ventas. Para ello toma como control el promedio histórico de ventas de los días miércoles, escogidos al azar, de

sus datos. Tal día se tomó como referencia luego de un sorteo. Prueba un mes con el primer tipo de promoción, deja pasar dos meses, y prueba un mes con un segundo tipo de promoción que incluye premios a los clientes. Sus resultados se muestran en la tabla siguiente, sobre la base de ellos decidir si las promociones dan resultado.

Medición N°	Control	Promo 1	Promo 2
1	245	230	276
2	176	252	289
3	198	261	293
4	217	245	282

6) En una industria farmacéutica hay cuatro líneas de producción para la fabricación de analgésicos, con distinta tecnología. Sus rendimientos horarios son similares, pero el encargado desea averiguar si la cantidad de productos rechazados es la misma para los cuatro. Para ello toma de los registros históricos de producción 7 días elegidos al azar, del último semestre. Sus resultados fueron :

Dato N°	1	2	3	4
1	452	322	298	340
2	379	345	312	358
3	412	367	280	345
4	320	341	310	362
5	350	372	235	370
6	390	317	304	326
7	378	324	320	368

Decidir si hay diferencias entre las cuatro tecnologías.

7) Se ha determinado una magnitud clínica mediante cuatro operadores diferentes usando una misma técnica y con la misma muestra sanguínea. El objeto es detectar si hay diferencias debido al factor humano en los resultados del laboratorio. Los datos obtenidos son:

Dato N°	A	B	C	D
1	45	44	40	44
2	43	44	44	45
3	41	45	45	46
4	42	43	46	45

a) Se pide determinar usando el modelo de ANOVA unifactorial si hay diferencia entre los cuatro operadores. En caso de encontrarla, investigar usando el modelo Student todos los pares posibles para determinar entre cuales de ellos hay diferencias.

b) Si se sabe que el valor real de la magnitud es 45 estudiar a cada operador en exactitud y precisión para determinar un orden de mérito entre los mismos.