PARTE III
TEMAS RELACIONADOS CON LA INFERENCIA

La inferencia estadística ofrece más métodos que los que cualquier persona pueda conocer bien. Para demostrarlo, basta con dar un vistazo a cualquier programa estadístico importante. En una introducción a la estadística hay que ser selectivo. En la primera y segunda parte de este libro establecimos las bases que permiten comprender la estadística:

- La naturaleza y los objetivos del análisis de datos.
- Las ideas clave de los procedimientos para obtener datos.
- Los razonamientos que hay detrás de los intervalos de confianza y de las pruebas de significación.
- La experiencia de aplicar estas ideas a situaciones simples.

Cada uno de los tres capítulos de la tercera parte ofrece una breve introducción a un tema más avanzado de inferencia estadística. No importa el orden en que los leas.

¿Qué es lo que hace que un método estadístico sea "más avanzado"? La complejidad de los datos es un aspecto a considerar. En nuestra introducción a la inferencia, en la segunda parte, sólo consideramos métodos de inferencia aplicados a un solo parámetro poblacional y para comparar dos parámetros. El siguiente paso es la comparación de más de dos parámetros. El capítulo 8 enseña cómo comparar más de dos proporciones y el capítulo 9 discute la comparación de más de dos medias. En estos capítulos tenemos datos de tres o más muestras o tratamientos experimentales, no sólo de una o dos muestras. Otra situación más compleja hace referencia a la asociación entre dos variables. Describimos la asociación mediante una tabla de contingencia si las variables son categóricas, o mediante la correlación y la regresión si son cuantitativas. El capítulo 8 presenta la inferencia para tablas de contingencia. La inferencia para la regresión es el tema del capítulo 10. Los capítulos 8, 9 y 10, conjuntamente, llevan nuestro conocimiento sobre la inferencia al mismo punto que alcanzamos en nuestro estudio sobre el análisis de datos de los capítulos 1 y 2.

Una mayor complejidad exige una mayor dependencia de los programas estadísticos. En estos tres capítulos finales interpretarás con más frecuencia resultados de programas estadísticos o, si es posible, utilizarás tú mismo dichos programas. Los cálculos que se necesitan en el capítulo 8 puedes hacerlos sin necesidad de programas estadísticos o calculadoras con funciones estadísticas. En los capítulos 9 y 10, el esfuerzo de cálculo es demasiado grande y su contribución al conocimiento es demasiado pequeña. Afortunadamente, puedes comprender las ideas sin necesidad de hacer los cálculos paso a paso.

Otro aspecto de los métodos "más avanzados" es la aparición de nuevos conceptos e ideas. Llegados a este punto, hemos tenido que decidir qué temas estadísticos podemos enseñar en un primer curso de estadística. Los capítulos 8, 9 y 10 presentan unos métodos más complejos basados en los conocimientos de estadística adquiridos en los capítulos precedentes, sin utilizar conceptos fundamentales nuevos. Leyendo las secciones sobre "el problema de las comparaciones múltiples" de los capítulos 8 y 9 puedes darte cuenta de que la estadística aplicada no necesita de grandes ideas adicionales. Las ideas que ya has adquirido se encuentran entre las más importantes del mundo de la estadística.

8. INFERENCIA PARA TABLAS DE CONTINGENCIA

KARL PEARSON

Karl Pearson (1857-1936), catedrático en la University College de Londres, ya había publicado nueve libros antes de dirigir su gran energía hacia la estadística en 1893. Por supuesto, Pearson no se dedicó realmente a la estadística, ya que todavía no era un campo de estudio independiente. Pearson estudió problemas sobre la herencia y la evolución que le condujeron a la estadística.

Pearson desarrolló una familia de curvas –las llamaríamos curvas de densidad– para describir datos biológicos que no tenían una distribución normal. Posteriormente se preguntó cómo podría verificar si realmente una de estas curvas se ajustaba bien a un conjunto de datos. En 1900, inventó un método, la prueba ji cuadrado. La prueba ji cuadrado de Pearson tiene el honor de ser el procedimiento de inferencia estadística más antiguo que todavía se utiliza. En la actualidad se utiliza principalmente para problemas algo distintos de los que motivaron a Pearson, tal como veremos en este capítulo.

Después de Pearson, la estadística se convirtió en un campo de estudio. Fisher y Neyman, en los años veinte y treinta, dieron a la estadística buena parte de su forma actual. Pero he aquí lo que dice uno de los principales historiadores de la estadística sobre los orígenes:

Antes de 1900, veíamos a muchos científicos de diferentes campos que desarrollaban y utilizaban técnicas que ahora reconocemos como pertenecientes a la estadística moderna. Después de 1900, empezamos a ver verdaderos estadísticos que desarrollaban estas técnicas dentro de la lógica unificada de una ciencia empírica que va más allá de las partes que la componen. No hubo un momento bien marcado; pero con Pearson, Yule y el creciente número de estudiantes del laboratorio de Pearson, podríamos decir que la naciente disciplina se había consolidado.[1]

comprar

[1] Stephen M. Stigler, *The History of Statistics: The Measurement of Uncertainty before 1900*, Belknap Press, Cambridge, Mass., 1986. La cita es de la pág. 361.

8.1 Introducción

Los procedimientos z de dos muestras del capítulo 7 nos permiten comparar las proporciones de éxitos de dos grupos, ya sean dos poblaciones o dos tratamientos en un experimento. ¿Qué ocurre si queremos comparar más de dos grupos? Necesitamos una nueva prueba estadística. La nueva prueba empieza presentando los datos de una manera diferente, en forma de tabla de contingencia. Las tablas de contingencia tienen un uso más amplio que comparar la proporción de éxitos en varios grupos. Tal como vimos en la sección 6 del capítulo 2, las tablas de contingencia describen las relaciones entre dos variables categóricas cualesquiera. La misma prueba estadística que se utiliza para comparar varias proporciones, contrasta si la variable fila y la variable columna, cuyos valores se muestran en una tabla de contingencia, están relacionadas. Empezaremos con el problema de la comparación de varias proporciones.

EJEMPLO 8.1

Los adictos a la cocaína necesitan esta droga para sentir placer. Quizás dándoles una medicación que combatiese la depresión se les ayudaría a dejar la cocaína. Un estudio de tres años comparó un antidepresivo denominado desipramina con el litio (un tratamiento habitual para combatir la adicción a la cocaína) y con un placebo. Los sujetos experimentales eran 72 cocainómanos que querían acabar con su drogodependencia. Se asignaron al azar veinticuatro sujetos a cada tratamiento. He aquí los recuentos y las proporciones de sujetos que no recayeron en el consumo de cocaína durante el estudio.[2]

Grupo	Tratamiento	Sujetos	No recayeron	Proporción
1	Desipramina	24	14	0,583
2	Litio	24	6	0,250
3	Placebo	24	4	0,167

Las proporciones muestrales de sujetos que se mantuvieron apartados de la cocaína son bastante distintas. El diagrama de barras de la figura 8.1 muestra las diferencias entre grupos. Estos datos, ¿constituyen una buena evidencia a favor de que las proporciones de éxitos en los tres tratamientos son distintas en la población de todos los consumidores de cocaína? ■

[2] D. M. Barnes, 1988, "Breaking the cycle of addiction", *Science*, 241, págs. 1.029-1.030.

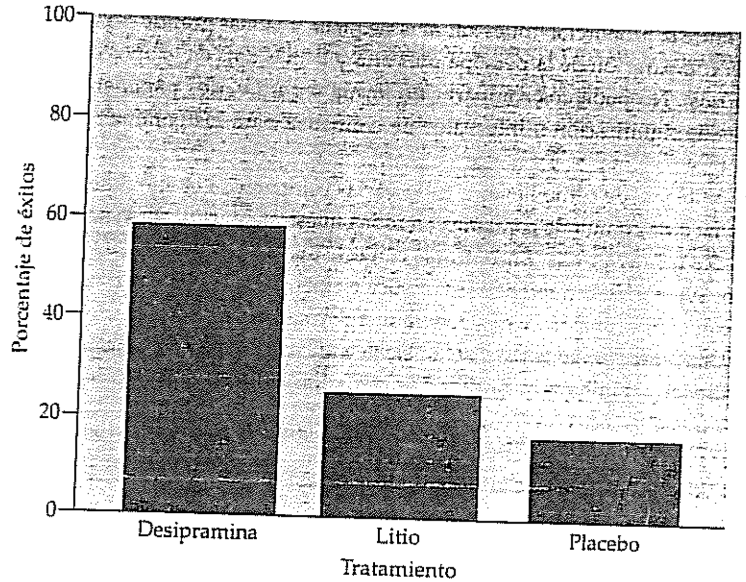

Figura 8.1. Diagrama de barras que compara las proporciones de éxitos de los tres tratamientos del ejemplo 8.1 para combatir la adicción a la cocaína.

8.1.1 El problema de las comparaciones múltiples

Llama a las proporciones poblacionales de éxitos en los tres grupos p_1, p_2 y p_3. De nuevo, utilizamos los subíndices para recordar qué grupo describe un determinado parámetro o estadístico. Para comparar estas tres proporciones poblacionales, podríamos utilizar los procedimientos z de dos muestras varias veces:

- Contrasta $H_0: p_1 = p_2$, para ver si la proporción de éxitos de la desipramina es distinta de la proporción de éxitos del litio.
- Contrasta $H_0: p_1 = p_3$, para ver si la desipramina difiere del placebo.
- Contrasta $H_0: p_2 = p_3$, para ver si el litio difiere del placebo.

El inconveniente de hacer estos contrastes es que obtenemos tres valores P, uno para cada uno de los contrastes. Esto no nos dice cuál es la probabilidad de que *tres* proporciones muestrales se encuentren tan separadas como lo están éstas. Puede ser que 0,167 y 0,583 sean significativamente distintas si sólo miramos dos grupos, pero que no lo sean si sabemos que son la proporción más pequeña y la más grande de los tres grupos. Conforme observamos más grupos, esperamos que la separación entre la proporción muestral más pequeña y la más grande sea mayor (piensa en la comparación de la persona más alta y la más baja en grupos cada vez más numerosos). No podemos comparar de forma segura varios parámetros haciendo pruebas de significación o intervalos de confianza para los parámetros de dos en dos.

El problema de cómo llevar a cabo varias comparaciones simultáneamente con alguna medida común de confianza en todas nuestras conclusiones es habitual en la estadística más avanzada. Encontrarse con este tipo de problema es un signo de que estamos superando la inferencia estadística elemental. Los métodos estadísticos para tratar con varias comparaciones generalmente constan de dos partes:

1. Una *prueba conjunta* para ver si existe suficiente evidencia de *alguna* desigualdad entre los parámetros que queremos comparar.

2. Un *análisis complementario* detallado para decidir cuáles son los parámetros distintos y para estimar el valor de las diferencias.

La prueba conjunta es con frecuencia razonablemente sencilla, aunque más compleja que las pruebas que hemos visto hasta ahora. El análisis complementario puede ser bastante complicado. En nuestra introducción básica a la estadística aplicada sólo veremos algunas pruebas conjuntas. En este capítulo presentamos una prueba para comparar varias proporciones poblacionales. El próximo capítulo muestra cómo comparar medias poblacionales.

8.2 Tablas de contingencia

El primer paso hacia la prueba conjunta para comparar varias proporciones poblacionales consiste en disponer los datos en una ***tabla de contingencia*** que da los recuentos de éxitos y fracasos. He aquí la tabla de contingencia de los datos de la adicción a la cocaína:

	Recaídas	
	No	Sí
Desipramina	14	10
Litio	6	18
Placebo	4	20

Tabla f × c. Las celdas

Ésta es una tabla de contingencia de 3 × 2 ya que tiene 3 filas y 2 columnas. Una tabla con f filas y c columnas es una ***tabla*** $f \times c$. La tabla muestra la relación entre dos variables categóricas. La variable explicativa es el tratamiento (la desipramina,

el litio y el placebo). La variable respuesta es el éxito (no hay recaída) o el fracaso (hay recaída). La tabla de contingencia presenta los recuentos de las 6 combinaciones de los valores de estas variables. Cada uno de estos 6 recuentos ocupa una ***celda*** de la tabla.

8.2.1 Recuentos esperados

Queremos contrastar la hipótesis nula de que *no hay diferencias* entre las proporciones de éxitos de los adictos a la cocaína que han sido sometidos a cada uno de los tres tratamientos:

$$H_0 : p_1 = p_2 = p_3$$

La hipótesis alternativa es la de que existe alguna diferencia, es decir, que las tres proporciones no son iguales:

$$H_a : \text{ no se cumple que } p_1 = p_2 = p_3$$

La hipótesis alternativa ya no es de una o de dos colas. Es de "muchas colas", ya que permite cualquier relación distinta a "las tres proporciones son iguales". Por ejemplo, H_a incluye la situación en la que $p_2 = p_3$, pero que p_1 tenga un valor distinto.

Para contrastar H_0, comparamos los recuentos observados en la tabla de contingencia con los *recuentos esperados*, los recuentos que esperaríamos –excepto por el efecto del azar– si H_0 fuera cierta. Si las diferencias entre los recuentos observados y los esperados son grandes, estas diferencias constituyen una evidencia en contra de H_0. Es fácil hallar los recuentos esperados.

> **RECUENTOS ESPERADOS**
>
> El **recuento esperado** para cualquier celda de una tabla de contingencia cuando H_0 es cierta es
>
> $$\text{recuento esperado} = \frac{\text{total fila} \times \text{total columna}}{\text{total tabla}}$$

Para comprender por qué funciona esta fórmula, piensa primero sólo en una proporción.

EJEMPLO 8.2

María es una jugadora de baloncesto que encesta el 70% de los tiros libres. Si en un partido María lanza 10 tiros libres, esperamos que encestará el 70% de ellos, o 7 de los 10 tiros libres. Por supuesto, no siempre que María lanza 10 tiros libres encesta exactamente 7. En cada partido existe una variación debida al azar. Pero, a largo plazo, 7 de 10 es lo que esperamos. De hecho, 7 es el número *medio* de lanzamientos que acierta María cuando lanza 10 veces. ■

En un lenguaje más formal, si tenemos n observaciones independientes y la probabilidad de éxito de cada observación es p, esperamos np éxitos. Si obtenemos una muestra aleatoria simple de n individuos de una población en la cual la proporción de éxitos es p, esperamos np éxitos en la muestra. Ésta es la consideración que hay detrás de la fórmula de los recuentos esperados de una tabla de contingencia.

Vamos a aplicar esta consideración al estudio sobre la cocaína. La tabla de contingencia con los totales de filas y columnas es

	Recaída		
	No	Sí	Total
Desipramina	14	10	24
Litio	6	18	24
Placebo	4	20	24
Total	24	48	72

Hallaremos el valor esperado de la celda de la fila 1 (desipramina) y de la columna 2 (recaída). La proporción de recaídas entre todos los 72 sujetos es:

$$\frac{\text{recuento de recaídas}}{\text{total de la tabla}} = \frac{\text{total de la columna 2}}{\text{total de la tabla}} = \frac{48}{72} = \frac{2}{3}$$

Piensa en esta proporción como p, la proporción conjunta de recaídas. Si H_0 es cierta, esperamos hallar (excepto por el efecto del azar) esta misma proporción de recaídas en los tres grupos. En consecuencia, el recuento de recaídas esperado entre los 24 sujetos que tomaron desipramina es

$$np = (24)\ (2/3) = 16.$$

El cálculo de este recuento esperado sigue la forma indicada en el recuadro:

$$\frac{\text{total de la fila 1} \times \text{total de la columna 2}}{\text{total de la tabla}} = \frac{(24)(48)}{72}$$

EJEMPLO 8.3

He aquí los recuentos observados y los recuentos esperados:

	Observados		Esperados		
	No	Sí	No	Sí	
Desipramina	14	10	8	16	= 2/3 de 24
Litio	6	18	8	16	= 2/3 de 24
Placebo	4	20	8	16	= 2/3 de 24
				48	

Debido a que 2/3 del total de sujetos recayeron, esperamos que 2/3 de los 24 sujetos de cada grupo recaigan si no hay diferencias entre los tratamientos. De hecho, la desipramina tuvo menos recaídas (10) y más éxitos (14) de los esperados. El placebo tiene menos éxitos (4) y más recaídas (20). Ésta es otra manera de decir lo que las proporciones muestrales del ejemplo 8.1 dicen más directamente: la desipramina es mucho más efectiva que el placebo, estando el litio entre ambos. ■

EJERCICIOS

8.1. La North Carolina State University estudió los resultados obtenidos por los estudiantes de un curso obligatorio para la especialidad de ingeniería química. Uno de los temas de interés del estudio es la relación entre el tiempo invertido en actividades extracurriculares y el éxito de los estudiantes en este curso. He aquí datos sobre 119 estudiantes que respondieron a la pregunta sobre sus actividades extracurriculares.[3]

	Actividades extracurriculares (horas por semana)		
	< 2	2 a 12	>12
Aprobado o nota superior	11	68	3
Suspenso	9	23	5

(a) Tenemos una tabla $f \times c$. ¿Cuáles son los valores de f y c?

(b) Halla la proporción de estudiantes que tuvieron éxito (aprobado o nota superior) en cada uno de los tres grupos de actividades extracurriculares. ¿Qué tipo de relación entre las actividades extracurriculares y el éxito obtenido en el curso parecen indicar estas proporciones?

[3] Richard M. Felder, *et al.*, "Who gets it and who doesn't: a study of student performance in an introductory chemical engineering course", *1992 ASEE Annual Conference Proceedings*, American Society for Engineering Education, Washington, D.C., 1992, págs. 1.516-1.519.

(c) Dibuja un diagrama de barras para comparar las tres proporciones de éxitos.

(d) La hipótesis nula establece que las proporciones de éxitos son las mismas en los tres grupos si tomamos la población de todos los estudiantes. Halla los recuentos esperados si esta hipótesis fuera cierta y muéstralos en una tabla de contingencia.

(e) Compara los recuentos observados con los esperados. ¿Existen desviaciones grandes entre ellos? Estas desviaciones son otra manera de describir la relación que describiste en (b).

8.2. ¿Qué relación existe entre los hábitos de fumar de los estudiantes y los de sus padres? He aquí datos de una encuesta realizada a alumnos de ocho escuelas de secundaria de Arizona.[4]

	Estudiantes fumadores	Estudiantes no fumadores
Los dos padres fuman	400	1.380
Sólo uno de los padres fuma	416	1.823
Ninguno de los dos padres fuma	188	1.168

(a) Esto es una tabla $f \times c$. ¿Cuáles son los valores de f y de c?

(b) Calcula la proporción de estudiantes que fuma en cada uno de los tres grupos de padres. Luego describe con palabras la asociación entre los hábitos de fumar de padres e hijos.

(c) Dibuja un gráfico para mostrar dicha asociación.

(d) Explica con palabras lo que plantea la hipótesis nula $H_0: p_1 = p_2 = p_3$ sobre los hábitos fumadores de los estudiantes.

(e) Halla los recuentos esperados si H_0 fuera cierta y muéstralos en una tabla de contingencia similar a la tabla de los recuentos observados.

(f) Compara las tablas de los recuentos observados y esperados. Explica de qué manera la comparación expresa la misma asociación que viste en (b) y (c).

8.3 Prueba Ji cuadrado

La comparación de las proporciones muestrales de éxitos describe las diferencias entre los tres tratamientos para combatir la adicción a la cocaína. De todas formas, la prueba estadística que nos indica si esas diferencias son estadísticamente significativas no utiliza directamente las proporciones muestrales. La prueba compara los recuentos observados con los recuentos esperados. El estadístico de contraste que hace la comparación es el *estadístico Ji cuadrado.*

[4] S. V. Zagona (ed.), *Studies and Issues in Smoking Behavior*, University of Arizona Press, Tucson, 1967, págs. 157-180.

ESTADÍSTICO JI CUADRADO

El estadístico Ji cuadrado es una medida de la diferencia entre los recuentos observados y los recuentos esperados en una tabla de contingencia. La fórmula del estadístico es

$$\chi^2 = \sum \frac{(\text{recuento observado} - \text{recuento esperado})^2}{\text{recuento esperado}}$$

La suma es de todas las $f \times c$ celdas de la tabla.

El estadístico Ji cuadrado es una suma de términos, uno por cada celda de la tabla. En el ejemplo de la cocaína, 14 sujetos del grupo de la desipramina consiguieron evitar una recaída. El recuento esperado de esta celda es 8. En consecuencia, el componente estadístico Ji cuadrado correspondiente a esta celda es

$$\frac{(\text{recuento observado} - \text{recuento esperado})^2}{\text{recuento esperado}} = \\ = \frac{(14-8)^2}{8} = \frac{36}{8} = 4.5$$

Interpreta el estadístico Ji cuadrado, χ^2, como una medida de la distancia entre los recuentos observados y los recuentos esperados. Como cualquier distancia, su valor siempre es cero o positivo. Es cero sólo cuando los recuentos observados son exactamente iguales a los recuentos esperados. Los valores de χ^2 grandes constituyen una evidencia en contra de H_0, ya que indican que los recuentos observados están lejos de lo que esperaríamos si H_0 fuera cierta. Aunque la hipótesis alternativa H_a es de muchas colas, la prueba Ji cuadrado es de una cola, ya que cualquier violación de H_0 tiende a producir un valor de χ^2 grande. Los valores pequeños de χ^2 no constituyen ninguna evidencia en contra de H_0.

El cálculo manual de los recuentos esperados y el del estadístico Ji cuadrado lleva bastante tiempo. Como de costumbre, los programas estadísticos ahorran tiempo y siempre hacen los cálculos correctamente. Algunas calculadoras también calculan el estadístico Ji cuadrado.

EJEMPLO 8.4

Entra los datos de la tabla de contingencia (los 6 recuentos) del estudio sobre la cocaína en el programa estadístico Minitab y ejecuta la prueba Ji cuadrado. Los resulmas estadísticos presentan

entos observados y sitúa los servados. Minitab numera las a la fila y la columna de tota- adrado, χ^2. Para estos datos, or cada celda de la tabla. El co indica los términos indivi- valor que obtuvimos a mano. r al menos tan grande como estadísticos dan el valor P. babilidad de un valor de χ^2 arece al final de los resulta- or P pequeño nos da argu- de los tres tratamientos. ■

bserved counts

3948

s a la tabla de con- dan los recuentos) del estadístico Ji lidad de un valor consecuencia, el

La prueba Ji cuadrado es una prueba conjunta para comparar cualquier número de proporciones poblacionales. Si el contraste nos permite rechazar la hipótesis nula de que todas las proporciones son iguales, entonces procedemos a hacer el análisis complementario que examina las diferencias con detalle. No vamos a describir cómo hacer el análisis complementario, pero debes mirar qué efectos concretos sugieren los datos.

EJEMPLO 8.5

El estudio sobre la cocaína encontró diferencias significativas entre las proporciones de éxitos de los tres tratamientos para combatir la adicción a la cocaína. Podemos ver diferencias concretas de tres maneras.

Primero mira las *proporciones muestrales*:

$$\hat{p}_1 = 0{,}583 \qquad \hat{p}_2 = 0{,}250 \qquad \hat{p}_3 = 0{,}167$$

Las proporciones muestrales indican que la mayor diferencia entre las proporciones corresponde a la desipramina, que tiene una proporción de éxitos mucho mayor que el litio o el placebo. Éste es el efecto que esperaba encontrar el estudio.

A continuación, *compara los recuentos observados con los recuentos esperados* que aparecen de la figura 8.2. El Tratamiento 1 (la desipramina) tiene más éxitos y menos fracasos de lo que esperaríamos si los tres tratamientos tuvieran la misma proporción de éxitos en la población. Los otros dos tratamientos tienen menos éxitos y más fracasos de los esperados.

Finalmente, Minitab proporciona debajo de la tabla las 6 "distancias" individuales entre los recuentos observados y los esperados. Estas distancias se suman para obtener χ^2. La disposición de los ***componentes de*** χ^2 es la misma que la disposición de los recuentos en la tabla 3 × 2. Los componentes mayores muestran qué celdas contribuyen más a la distancia conjunta χ^2. El mayor componente corresponde, de lejos, a la celda situada en la parte izquierda superior de la tabla: la desipramina tiene más éxitos de los que en principio cabría esperar si H_0 fuera cierta.

Las tres maneras de examinar los datos apuntan hacia la misma conclusión: la desipramina funciona mejor que los otros tratamientos. Ésta es una conclusión informal. Existen métodos más avanzados que proporcionan pruebas e intervalos de confianza que hacen que el examen complementario sea más formal. ■

cia del estudio sobre la cocaína, de tres tratamientos y dos resul- 2 columnas. Esto es, $f = 3$ y $c = 2$. El estadístico Ji cuadrado, en s grados de libertad:

$$(f-1)(c-1) = (3-1)(2-1) = (2)(1) = 2$$

e ordenador de la figura 8.2 dan 2 grados de libertad para la Ji

do del estadístico Ji cuadrado es $\chi^2 = 10{,}500$. Fíjate en la fila = 2 de la tabla E. El valor de $\chi^2 = 10{,}500$ se sitúa entre los valo- 005 de la distribución Ji cuadrado con 2 grados de libertad. ba Ji cuadrado siempre es de una cola. Por tanto, el valor P de tra entre 0,01 y 0,005. Las diferencias entre las tres proporciones ativas al nivel $\alpha = 0{,}01$. ■

ordenador de la figura 8.3 dan dos grados de libertad para la

e esto sea correcto.

da el valor del estadístico Ji cuadrado, $\chi^2 = 6{,}926$. ¿Entre qué la E se encuentra este valor? ¿Qué te dice la tabla sobre el

ordenador de la figura 8.4 dan dos grados de libertad para la

e esto sea correcto.

dá el valor del estadístico Ji cuadrado, $\chi^2 = 37{,}568$. ¿Dónde se n la tabla E? ¿Qué te indica la tabla sobre el valor P?

s de la prueba Ji cuadrado

ncia pueden surgir de diversas situaciones. El estudio sobre la ento que asigna 24 adictos a tres grupos. Cada grupo es una ación distinta correspondiente a un tratamiento diferente. El el tamaño de cada muestra por adelantado. Para cada sujeto cuál de los dos resultados posibles se observa. La hipótesis " entre los tratamientos toma la forma de "igual proporción de

DISTRIBUCIONES JI CUADRADO

Las **distribuciones Ji cuadrado** son una familia de distribuciones que sólo toman valores positivos y que son asimétricas hacia la derecha. Una distribución Ji cuadrado concreta viene determinada por un parámetro, llamado **grados de libertad**.

La prueba Ji cuadrado para una tabla de contingencia de f filas y c columnas utiliza los valores críticos de la distribución Ji cuadrado con $(f-1)(c-1)$ grados de libertad. El valor P es el área situada a la derecha de χ^2 por debajo de la curva de densidad de la Ji cuadrado.

La figura 8.5 representa las curvas de densidad de tres miembros de la familia de distribuciones Ji cuadrado. A medida que aumentan los grados de libertad, las curvas de densidad son menos asimétricas y los valores mayores son más probables. La tabla E, situada al final del libro, proporciona los valores críticos de las distribuciones Ji cuadrado. Puedes utilizar la tabla E si los programas estadísticos no te dan los valores P de la prueba Ji cuadrado.

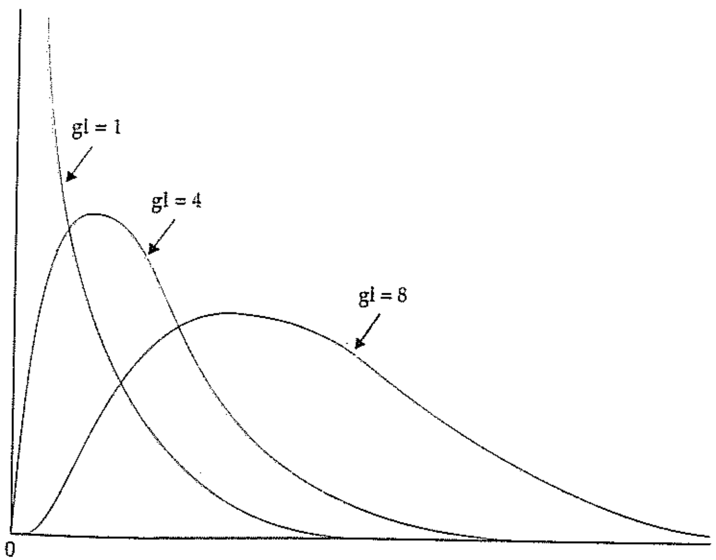

Figura 8.5. Curvas de densidad para las distribuciones Ji cuadrado con 1, 4 y 8 grados de libertad. Las distribuciones Ji cuadrado sólo toman valores positivos.

EJEMPLO 8.6

La tabla de contingencia del estudio sobre la cocaína, de tres tratamientos y dos resultados, tiene 3 filas y 2 columnas. Esto es, $f = 3$ y $c = 2$. El estadístico Ji cuadrado, en consecuencia, tiene los grados de libertad:

$$(f-1)(c-1) = (3-1)(2-1) = (2)(1) = 2$$

Los resultados de ordenador de la figura 8.2 dan 2 grados de libertad para la Ji cuadrado.

El valor observado del estadístico Ji cuadrado es $\chi^2 = 10.500$. Fíjate en la fila correspondiente a gl = 2 de la tabla E. El valor de $\chi^2 = 10.500$ se sitúa entre los valores críticos 0,01 y 0,005 de la distribución Ji cuadrado con 2 grados de libertad. Recuerda que la prueba Ji cuadrado siempre es de una cola. Por tanto, el valor P de $\chi^2 = 10.500$ se encuentra entre 0.01 y 0.005. Las diferencias entre las tres proporciones de éxitos son significativas al nivel $\alpha = 0.01$. ■

EJERCICIOS

8.5. Los resultados de ordenador de la figura 8.3 dan dos grados de libertad para la tabla del ejercicio 8.1.

(a) Comprueba que esto sea correcto.

(b) El ordenador da el valor del estadístico Ji cuadrado, $\chi^2 = 6.926$. ¿Entre qué dos valores de la tabla E se encuentra este valor? ¿Qué te dice la tabla sobre el valor P?

8.6. Los resultados de ordenador de la figura 8.4 dan dos grados de libertad para la tabla del ejercicio 8.2.

(a) Comprueba que esto sea correcto.

(b) El ordenador dá el valor del estadístico Ji cuadrado, $\chi^2 = 37.568$. ¿Dónde se encuentra este valor en la tabla E? ¿Qué te indica la tabla sobre el valor P?

8.3.2 Más aplicaciones de la prueba Ji cuadrado

Las tablas de contingencia pueden surgir de diversas situaciones. El estudio sobre la cocaína es un experimento que asigna 24 adictos a tres grupos. Cada grupo es una muestra de una población distinta correspondiente a un tratamiento diferente. El diseño del estudio fija el tamaño de cada muestra por adelantado. Para cada sujeto del estudio se registra cuál de los dos resultados posibles se observa. La hipótesis nula de "no diferencia" entre los tratamientos toma la forma de "igual proporción de

```
Expected counts are printed below observed counts

         C1       C2       C3      C4    Total
  1      58      874       15       8      955
      39.08   896.44    14.61    4.87

  2     222     3927       70      20     4239
     173.47  3979.05    64.86   21.62

  3      50     2396       34      10     2490
     101.90  2337.30    38.10   12.70

  4       7      533        7       4      551
      22.55   517.21     8.43    2.81

Total   337     7330      126      42     8235

ChiSq =  9.158 + 0.562 + 0.010 + 2.011 +
        13.575 + 0.681 + 0.407 + 0.121 +
        26.432 + 1.474 + 0.441 + 0.574 +
        10.722 + 0.482 + 0.243 + 0.504 = 67.397
df = 9
2 cells with expected counts less than 5.0

  Chisquare 9.
  67.3970    1.0000
```

Figura 8.6. Resultados obtenidos con Minitab a partir de la tabla 4 × 4 del ejemplo 8.8.

una evidencia contundente a favor de que la categoría laboral está relacionada con el estado civil.

La tabla E nos da un resultado similar. Para una tabla 4 × 4, los grados de libertad son $(f-1)(c-1) = 9$. Mira en la fila correspondiente a gl = 9 en la tabla E. El mayor valor crítico de esta fila es 29,67, que corresponde a un valor P igual a 0,0005. La $\chi^2 = 67{,}397$ observada está más allá de este valor; por tanto, $P < 0{,}0005$.

A continuación, haz el análisis complementario para describir la relación. Al igual que en la sección 6 del capítulo 2, describimos la relación entre dos variables categóricas comparando porcentajes. He aquí la tabla de los porcentajes de los hombres de cada estado civil cuyos puestos de trabajo pertenecen a cada una de las categorías laborales. Cada columna de esta tabla da la distribución condicional de los niveles laborales entre los hombres de un determinado estado civil. Cada

columna suma el 100%, ya que se considera a todos los hombres de un estado civil concreto.

		Estado Civil			
		Soltero	Casado	Divorciado	Viudo
	1	17,2%	11,3%	11,9%	19,1%
Nivel	2	65,9%	50,8%	55,5%	47,7%
laboral	3	14,9%	31,0%	26,9%	23,8%
	4	2,0%	6,9%	5,6%	9,6%
		100%	100%	100%	100%

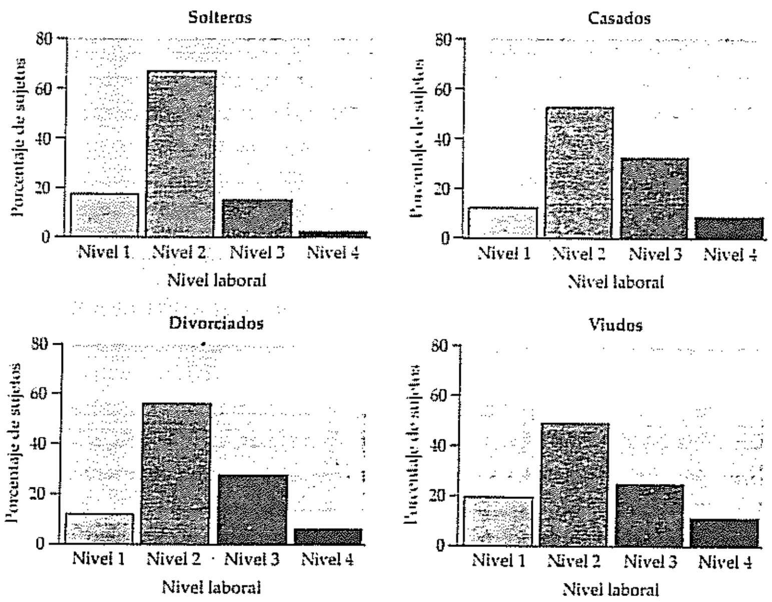

Figura 8.7. Diagramas de barras de los datos del ejemplo 8.8. Cada diagrama presenta el porcentaje de cada nivel laboral entre los hombres de un determinado estado civil.

El diagrama de barras de la figura 8.7 nos ayuda a comparar estas cuatro distribuciones condicionales. Vemos, enseguida, que entre los hombres solteros los menores porcentajes corresponden a los puestos de trabajo de niveles más altos, 3 y 4. No sólo los hombres casados, sino incluso los hombres que se casaron una vez y que ahora están divorciados o son viudos, tienen una mayor probabilidad de ocupar puestos de trabajo de mayor nivel. Fíjate en los 16 componentes de la suma Ji cuadrado de los resultados de ordenador para confirmarlo. Las cuatro celdas de los hombres solteros tienen los mayores componentes de la χ^2. La tabla de los valores obtenidos con Minitab muestra que los recuentos observados en los hombres solteros son superiores a los valores esperados en los niveles 1 y 2, e inferiores a los esperados en los niveles 3 y 4.

Por supuesto que esta asociación entre el estado civil y la categoría laboral no indica que estar soltero sea la *causa* de tener un puesto de trabajo de menor nivel. La explicación puede ser tan sencilla como el hecho de que los hombres solteros tienden a ser más jóvenes y que, por tanto, todavía no han alcanzado los niveles más altos. ■

8.3.3 Recuentos exigidos en las celdas para la prueba Ji cuadrado

Los resultados de ordenador que aparecen en la figura 8.6 tienen una característica más. Nos indican que los recuentos esperados de dos de las 16 celdas son menores que 5. La prueba Ji cuadrado, al igual que los procedimientos z de comparación de dos proporciones, es un método aproximado que se hace cada vez más exacto a medida que el número de observaciones de las celdas de la tabla se hace mayor. Afortunadamente, la aproximación es precisa para un número de observaciones modesto. He aquí una regla práctica.[6]

> **RECUENTOS EXIGIDOS EN LAS CELDAS PARA LA PRUEBA JI CUADRADO**
>
> Puedes utilizar de forma segura la prueba Ji cuadrado con los valores críticos de la distribución Ji cuadrado cuando no más de un 20% de los recuentos esperados son menores que 5, y todos los recuentos esperados son mayores o iguales que 1. En particular, en una tabla 2 × 2, los cuatro recuentos esperados deben ser mayores o iguales que 5.

El ejemplo 8.8 cumple fácilmente esta regla. Todos los recuentos esperados son mayores que 1, y sólo 2 de los 16 recuentos (el 12,5%) son menores que 5.

[6] Existen muchos estudios por ordenador sobre la precisión de los valores críticos de la Ji cuadrado para χ^2. Para una discusión breve y algunas referencias, consulta la sección 3.2.5 de David S. Moore, "Tests of chi-square type", en Ralph B. D'Agostino y Michael A. Stephens (eds.), *Goodness-of-Fit Techniques*, Marcel Dekker, Nueva York, 1986, págs. 63-95. Si los recuentos esperados en las celdas son aproximadamente iguales, la aproximación Ji cuadrado es adecuada cuando los promedios de los recuentos esperados son tan pequeños como 1 o 2. Si los recuentos esperados son desiguales, hay que tener en cuenta la regla orientativa dada en el texto. Para tener una guía sobre la inferencia con muestras pequeñas, consulta Alan Agresti, 1992, "A survey of exact inference for contingency tables", *Statistical Science*, 7, págs. 131-177.

8.3.4 Prueba Ji cuadrado y prueba z

Podemos utilizar la prueba Ji cuadrado para comparar cualquier número de proporciones. Si comparamos f proporciones y formamos dos columnas, una de éxitos y otra de fracasos, los recuentos resultantes forman una tabla $f \times 2$. Los valores P proceden de una distribución Ji cuadrado con $f - 1$ grados de libertad. Si $f = 2$, estamos comparando sólo dos proporciones. Tenemos dos maneras de hacerlo: la prueba z de la sección 7.3 y la prueba Ji cuadrado con un grado de libertad para una tabla 2 × 2. *Estas dos pruebas siempre concuerdan*. De hecho, el estadístico Ji cuadrado χ^2 es exactamente el cuadrado del estadístico z y el valor P de χ^2 es exactamente igual al valor P de dos colas de z. Te recomendamos la utilización del estadístico z para la comparación de dos proporciones, ya que te da la posibilidad de hacer pruebas de una cola y está relacionado con un intervalo de confianza de $p_1 - p_2$.

EJERCICIOS

8.7. El estudio de la North Carolina State University descrito en el ejercicio 8.1 también prestó atención a la relación entre las aspiraciones de los estudiantes y la nota real obtenida en el curso. El informe del estudio dice: "La probabilidad de pasar el curso CHE 205 era distinta para los estudiantes que hubieran estado satisfechos con un aprobado (el 36% de 14 estudiantes), con un notable (el 64% de 64), con un sobresaliente (el 90% de 30) o con una matrícula de honor (el 82% de 11)".[7]

(a) Utiliza esta información para construir una tabla 4 × 2 sobre las aspiraciones y los éxitos o fracasos de los 119 estudiantes.

(b) Halla los recuentos esperados. Los recuentos esperados no cumplen nuestra regla sobre el número de observaciones exigido. ¿Por qué? Sin embargo, los valores P de la distribución Ji cuadrado son razonablemente exactos.

(c) Una calculadora estadística da el valor del estadístico Ji cuadrado de esta tabla $\chi^2 = 14{,}986$. Da los grados de libertad. Utiliza la tabla E para aproximar el valor P de este estadístico.

(d) Describe brevemente la relación entre las aspiraciones de los estudiantes y las notas realmente obtenidas en el curso.

8.8. El ejercicio 7.20 compara los miembros de una mutua de asistencia sanitaria que se querellaron contra la mutua con una muestra aleatoria simple de miembros de la

[7] La cita es de la página 1.517 del artículo citado en la nota 3.

mutua que no presentaron demandas. El estudio dividió a los que se querellaron en dos grupos: los que se querellaron en relación con el tratamiento médico y los que se querellaron por otros motivos. He aquí los datos del número total de miembros de cada grupo y del número de asociados que voluntariamente se dieron de baja de la mutua.

	No se querellaron	Querellas médicas	Querellas no médicas
Total	743	199	440
Bajas	22	26	28

(a) Halla el porcentaje de los asociados que se dieron de baja en cada grupo.

(b) Construye una tabla de contingencia que tenga en cuenta las distintas opciones ante la idea de querellarse y la decisión de darse de baja o no.

(c) Halla los recuentos esperados y comprueba que puedas utilizar de forma segura la prueba Ji cuadrado.

(d) El estadístico Ji cuadrado de esta tabla es $\chi^2 = 31{,}765$. ¿Qué hipótesis nula y alternativa contrasta este estadístico? ¿Cuáles son sus grados de libertad? Utiliza la tabla E para aproximar el valor P.

(e) ¿Cuál es tu conclusión a partir de estos datos?

8.9. Un estudio a gran escala sobre el cuidado de los niños construyó una muestra a partir de datos de la *Current Population Survey* durante un periodo de varios años. El resultado puede considerarse, de forma aproximada, como una muestra aleatoria simple de los trabajadores que cuidan niños. La *Current Population Survey* distingue tres tipos de trabajadores que cuidan a niños: los que trabajan en casas privadas, los que no trabajan en casas privadas y los maestros de preescolar. He aquí cifras sobre el número de mujeres negras entre las mujeres que trabajan en cada una de las tres categorías.[8]

	Total	Negras
En casas privadas	2.455	172
Fuera de casa privadas	1.191	167
Maestras	659	86

(a) ¿Qué porcentaje de cada clase de trabajadoras que cuidan a niños son negras?

(b) Construye una tabla de contingencia que contenga el tipo de trabajadora y la raza (negras y otras).

[8] David M. Blau, 1992, "The child care labor market", *Journal of Human Resources*, 27, págs. 9-39.

(c) ¿Puedes utilizar de forma segura la prueba Ji cuadrado? ¿Qué hipótesis nula y alternativa contrasta χ^2?

(d) El estadístico Ji cuadrado para esta tabla es $\chi^2 = 53{,}194$. ¿Cuántos grados de libertad tiene el estadístico? Utiliza la tabla E para aproximar el valor P.

(e) ¿Qué conclusión sacas a partir de estos datos?

8.10. En el pasado, se recomendaba la congelación gástrica como tratamiento para las úlceras intestinales. Esta técnica dejó de utilizarse cuando se demostró mediante unos experimentos que no era efectiva. Un experimento comparativo aleatorizado halló que 28 de 82 pacientes sometidos a la congelación gástrica mejoraron, mientras que 30 de 78 pacientes del grupo placebo también lo hicieron.[9] Podemos contrastar la hipótesis de "no diferencias" entre los dos grupos de dos maneras: utilizando el estadístico z de dos muestras o el estadístico Ji cuadrado.

(a) Plantea la hipótesis nula con una alternativa de dos colas y lleva a cabo la prueba z. ¿Cuál es el valor P de la tabla A?

(b) Presenta los datos con una tabla 2 × 2. Utiliza la prueba Ji cuadrado para contrastar las hipótesis de (a). Comprueba que el estadístico χ^2 es el cuadrado del estadístico z. Utiliza la tabla E para comprobar que el valor P de la Ji cuadrado concuerda con el resultado de z hasta donde permite la precisión de la tabla.

(c) ¿Cuáles son tus conclusiones sobre la efectividad de la congelación gástrica como tratamiento para las úlceras?

RESUMEN

La prueba Ji cuadrado para tablas de contingencia contrasta la hipótesis nula de que no existe ninguna relación entre la variable fila y la variable columna.

Es frecuente utilizar la prueba Ji cuadrado para comparar varias proporciones poblacionales. La hipótesis nula plantea que todas las proporciones poblacionales son iguales. La hipótesis alternativa plantea que no todas ellas son iguales, de manera que permite cualquier otra relación entre las proporciones poblacionales.

El recuento esperado de cualquier celda de una tabla de contingencia cuando H_0 es cierta es

$$\text{recuento esperado} = \frac{\text{total fila} \times \text{total columna}}{\text{total tabla}}$$

[9] Lillian Lin Miao, "Gastric freezing: an example of the evaluation of medical therapy by randomized clinical trials", en John P. Bunker, Benjamin A. Barnes y Frederick Mosteller (eds.), *Costs, Risks, and Benefits of Surgery*, Oxford University Press, Nueva York, 1977, págs. 198-211.