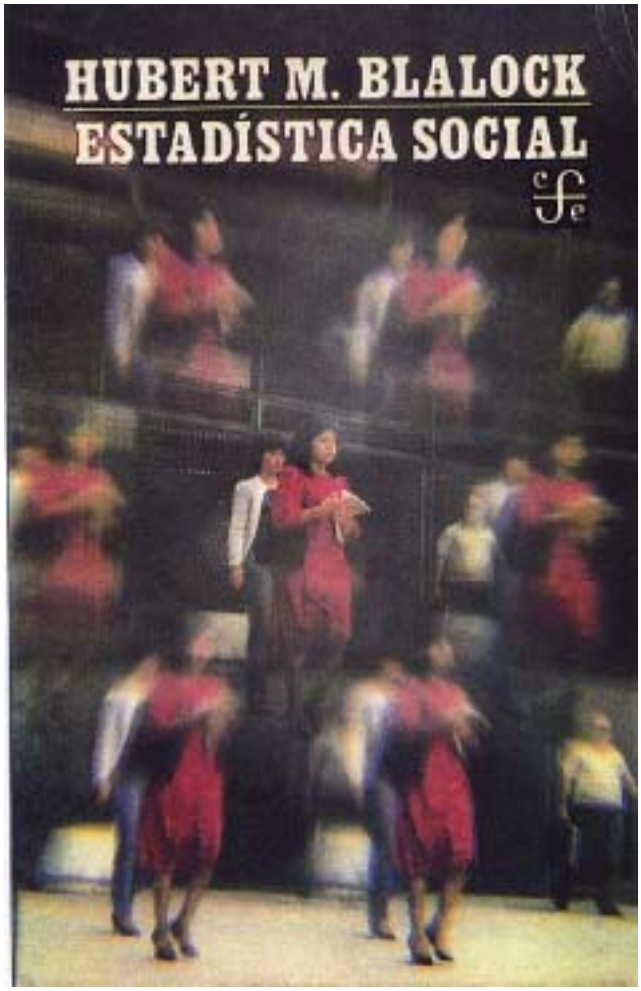
ESTADÍSTICA SOCIAL

HUBERT M. BLALOCK



FONDO DE CULTURA ECONÓMICA

México, 1986

Este material se utiliza con fines
exclusivamente didácticos

ÍNDICE

Prefacio 9

Primera Parte
INTRODUCCIÓN

1. Introducción: objetivos y límites de la estadística 15
1.1. Funciones de la estadística 16
1.2. El lugar de la estadística en el proceso de la investigación 19
1.3. Advertencia 20
Bibliografía 21

II. Teoría, medición y matemáticas 22
II.l. Teoría e hipótesis: definiciones operativas 22
II.2. El nivel de medición: escalas nominales, ordinales y de intervalo 26
II.3. Medición y estadística 32
II.4. Organización del libro 37
Bibliografía 40

Segunda Parte
ESTADÍSTICA DESCRIPTIVA UNIVARIADA

III. Escalas nominales: proporciones, porcentajes y razones 43
III.1. Proporciones 43
III.2. Porcentajes 45
111.3. Razones 49

Bibliografía 52

IV. Escalas de intervalo: distribuciones de frecuencia y representación gráfica 53
IV.1. Distribuciones de frecuencia: agrupamiento de los datos 53
IV.2. Distribuciones de frecuencia cumulativa 60
IV.3. Presentación gráfica: histogramas, polígonos de frecuencia y ojivas 61
Bibliografía 66

V. Escalas de intervalo: medidas de tendencia central 67
V.1. La media aritmética 67
V.2. La mediana 71
V.3. Cálculo de la media y la mediana de datos agrupados 73
V4. Comparación de la media y la mediana 81
V.5. Otras medidas de tendencia central 85
V.6. Deciles, cuartiles y percentiles 86
Bibliografía 88

VI. Escalas de intervalo: medidas de dispersión 90
VI.l. El recorrido 90
VI.2. La desviación cuartil 92
VI.3. La desviación media 92
VI.4. La desviación estándar 93
VI.5. El coeficiente de variabilidad 101
VI.6. Otras medidas resumidas 102
Bibliografía 103

VII. La distribución normal 104
VII.l. Distribuciones de frecuencias finitas versus infinitas 104
VII.2. Forma general de la curva normal 107
VII.3. Áreas bajo la curva normal 109

VII.4. Ilustraciones suplementarias del empleo de la tabla normal 113
Bibliografía 116

Tercera Parte
ESTADÍSTICA INDUCTIVA

VIII. Introducción a la estadística inductiva 119
VIII.l. Estadística y parámetros 119
VIII.2. Pasos en la verificación de una hipótesis 120
VIII.3. La falacia de afirmar el consecuente 123
VIII.4. La forma de las hipótesis estadísticas 124
Bibliografía 127

IX. Probabilidad 128
IX.1. Probabilidad a priori 129
IX.2. Propiedades matemáticas de las probabilidades 133
IX.3. Permutas 145
IX4. Valores esperados 151
IX.5. Independencia y muestreo aleatorio 153
Bibliografía 159

X. Pruebas de hipótesis: la distribución binomial 160
X1. La distribución de muestreo binomial 160
X.2. Pasos en las pruebas estadísticas 164
X.3. Aplicaciones de la binomial 177
X.4. Extensiones del binomio 181
X.5. Sumario 183
Bibliografía 186

XI. Pruebas de muestras simples que implican medias y proporciones 187
XI.l. Distribución en muestreo de las medias 187
XI.2. Prueba para la media de la población, conociendo. 194
XI.3. La distribución t de Student 199
XI.4. Pruebas que comportan proporciones 204
Bibliografía 210

XII. Estimación de punto e intervalo 211
XII.l. Estimación del punto 212
XII.2. Estimación del intervalo 215
XII.3. Intervalos de confianza para otros tipos de problemas 221
XII.4. Determinación del tamaño de la muestra 224
Bibliografía 227

Cuarta Parte
ESTADISTICAS BIVARIADAS Y MULTIVARIADAS

XIII. Pruebas de dos muestras: diferencia de las medias y las proporciones 231
XIII.l. Prueba de la diferencia de las medias 231
XIII.2. Diferencia de proporciones 240
XIII.3. Intervalos de confianza 245
XIII.4. Muestras dependientes: pares asociados 246
XIII.5. Comentarios a propósito de los esquemas experimentales y pruebas de significación 248
Bibliografía 255

XIV. Escalas ordinales: pruebas no paramétricas de dos muestras. 256
XIV1. Fuerza y eficiencia de la fuerza 257
XIV.2. La prueba de las secuencias (runs) de Wald-Wolfowitz 263

XIV.3. La prueba de Mann-Whitney o de Wilcoxon 269
XIVA. La prueba de Kolmogorov-Smirnov 277
XIV.5. La prueba de Wilcoxon de pares asociados y órdenes provistos de signo 280
XIV.6. Resumen 284
Bibliografía 288

XV. Escalas nominales: problemas de contingencia 289
XV.1. La prueba de la X-cuadrada 289
XV.2. La prueba exacta de Fisher 301
XV.3. Medidas de la fuerza de la relación 306
XV.4. Control de otras variables 319
Bibliografía 330

XVI. Análisis de la variancia 332
XVI.l. Análisis simple de la variancia 332
XVI.2. Comparación de medias específicas 343
XVI.3. Análisis bimodal de la variancia 349
XVIA. Alternativas no paramétricas del análisis de variancia 365
XVI.5. Medidas de asociación: correlación intraclase 370
Bibliografía 376

XVII. Correlación y regresión 377
XVII.l. Regresión lineal y mínimos cuadrados 377
XVII.2. Correlación 393
Bibliografía 413

XVIII. Correlación y regresión [conclusión] 414
XVIII.l. Prueba de significación e intervalos de confianza 414
XVIII.2. Correlación no lineal y regresión 425
XVIII.3. Efectos de los errores de medición 431
XVIII.4. Escalas ordinales: correlación de rangos 433
Bibliografía 445

XIX. Correlación múltiple y parcial 447
XIX.1. Regresión múltiple y mínimos cuadrados- 447
XIX.2. Correlación parcial 451
XIX.3. Correlación parcial e interpretaciones causales 461
XIX.4. Mínimos cuadrados múltiples y los coeficientes beta 469
XIX.5. ¡Correlación múltiple 473
XIX.6. Regresión múltiple y no linealidad 479
XIX.7. Pruebas de significación e intervalos de confianza 484
Bibliografía 489

XX. Análisis de covariancia y variables simuladas 491
XX.1. Relación de dos escalas de intervalo, control de la escala nominal 492
XX.2. Relación de una escala de intervalo y una escala nominal, control de la escala de intervalo ... 510
XX.3. Extensiones del análisis de covariancia 516
XX.4. Análisis de la variable simulada 517
XX.5. Observaciones finales 521
Bibliografía 526

Quinta Parte
MUESTREO

XXI. Muestreo 531
XXI.1. Muestreo aleatorio sencillo 532
XXI.2. Muestreo sistemático 537

XXI.3. Muestreo estratificado 539
XXIA. Muestreo por conglomerados 546
XXI.S. Muestreo sin probabilidad 552
XXI.6. Errores no de muestreo y tamaño de la muestra 553
Bibliografía 554

APÉNDICES

I. Resumen de operaciones algebraicas 559
Cuadros 565
Índice de figuras 599
Índice de cuadros 603

II. TEORÍA, MEDICIÓN Y MATEMÁTICAS

Este capítulo tiene por objeto esbozar en líneas generales las relaciones existentes entre las proposiciones teóricas, las hipótesis empíricas, la medida y los modelos matemáticos. Muchos de los problemas tratados en este capítulo no suelen examinarse en conexión con los cursos de estadística, lo que se debe en parte a la tendencia poco afortunada consistente en dividir la materia en cursos con las apelaciones de "teoría", "métodos de investigación" y "estadística". Esto ocasiona que las relaciones internas entre dichas materias resulten a veces oscurecidas. Con objeto de situar a la estadística en la debida perspectiva, conviene prestar atención a las relaciones entre las proposiciones teóricas y las hipótesis de investigación por una parte, y entre estas últimas y los modelos matemáticos por la otra.

Se suele decir con frecuencia que el objeto de la investigación está en verificar hipótesis desarrolladas teóricamente y que los métodos estadísticos capacitan para efectuar dichas, pruebas. Hay que tener presente, sin embargo, que los procesos implicados en pasar de la teoría a las hipótesis reales de investigación y de éstas a los enunciados de probabilidad del tipo empleado en la inferencia estadística no son en modo alguno directos. En efecto, en ambos casos hay que tomar decisiones, las cuales pueden dar lugar a un grado considerable de controversia. Examinemos primero el carácter de las decisiones que se requieren para desarrollar, a partir de proposiciones teóricas, hipótesis verificables.

*II .l. Teoría e hipótesis: definiciones operativas

En el instante en que empezamos a diseñar un proyecto de investigación enderezado a verificar una proposición que puede aparecer en un trabajo teórico, resulta evidente que hay que hacer varias cosas antes de poder proceder a la prueba. Tomemos a título de ejemplo concreto la siguiente proposición: "Cuanto más elevada es la condición social de una persona, tanto menores son sus prejuicios en relación con los negros." Supongamos que la "condición social" se haya definido como la posición que la persona ocupa en relación con otras en la jerarquía social, y los "prejuicios" como tendencia latente a la discriminación de una minoría o como actitud negativa basada en juicios preformados. Aun si se prefiere sustituir por otras las definiciones de esos dos conceptos, se descubrirá sin duda alguna que, cualesquiera que sean las definiciones escogidas, resulta imposible servirse directamente de ellas para decidir cuál sea exactamente la condición de Jones, pongamos por caso, o el grado de sus prejuicios.

La razón de ello reside en que la mayoría de las proposiciones son más bien teóricas que operativas. En la definición teórica, en efecto, un concepto se define en términos de otros conceptos que se dan por comprendidos. En el modelo ideal del sistema completamente deductivo, se tomarían ciertos conceptos sin definir (primarios), y todos los demás se definirían en términos de aquéllos. En la geometría euclidiana, por ejemplo, los conceptos de *punto y recta* pueden tomarse sin definir, pudiendo luego definirse las nociones de *ángulo, triángulo o rectángulo* en función de aquéllos términos. Pese a que la elección de conceptos no definidos es hasta cierto punto arbitraria, el hecho de que tengan que existir siempre algunos conceptos primeros o primarios resulta de la necesidad inherente de definir los conceptos teóricos en términos unos de otros.

Por otra parte, las definiciones operativas son definiciones que enuncian efectivamente los procedimientos empleados en la medición ([8], pp. 58 a 65). La definición operativa de "longitud", por ejemplo, indicará exactamente cómo deba medirse el largo de un cuerpo. El ejemplo de una definición operativa del prejuicio implicará una prueba como la de la escala de la distancia social de Bogardus o, tal vez, una relación de conceptos antinegros en una lista de 24 puntos, juntamente con instrucciones detalladas para recoger los datos, valorar los puntos, etcétera. Como quiera que toda medición implica como requisito mínimo alguna clasificación, la definición operativa puede considerarse como un conjunto detallado de instrucciones que permiten clasificar a los individuos en forma inequívoca. De este modo, la noción de grado de confianza o garantía de seguridad queda integrada en dicho concepto de la definición operativa. La definición ha de ser lo bastante precisa para que todas las personas que se sirvan del procedimiento lleguen a los mismos resultados. Lo que las definiciones teóricas del prejuicio y de la condición social consignadas más arriba no lo permitirán, por supuesto, directamente.

---

* El asterisco que precede una sección, párrafo o ejercicio indica que la materia que contiene o es de comprensión difícil o trata de conceptos con los que probablemente no están familiarizados los estudiantes que sólo disponen de una preparación limitada , en materia de metodología de investigación. El estudiante principiante puede perfectamente omitir dichos pasajes o leerlos superficialmente. Por su parte, el asterisco que precede al título de una sección indica que la sección entera puede omitirse si se quiere.

Sostenemos, pues, que en toda ciencia se utilizan dos tipos distintos de definiciones. Diversas maneras alternativas de enfocar la relación entre la teoría y la investigación conducen esencialmente a la misma conclusión. Northrop designa lo que hemos llamado definiciones teóricas como "conceptos por postulación", y las definiciones operativas como "conceptos por intuición", [9].

Por nuestra parte nos hemos servido de una terminología que parece implicar que hay dos maneras distintas de definir un "mismo" concepto, en tanto que Northrop prefiere referirse a dos tipos distintos de conceptos. Otros, todavía prefieren pensar en términos de índices, más que de definiciones operativas. El concepto de *índice* implica por lo regular que el procedimiento empleado da sólo un indicador imperfecto de alguna variable señalada y un indicador de esta variable. Pero, independientemente del punto de vista que cada cual prefiera, es indispensable comprender el carácter del nexo entre las dos clases de definiciones, de conceptos o de variables. Podemos preguntar si existe o no, un método puramente lógico de juntar las dos clases de definiciones. Otra forma de plantear la cuestión consistiría en preguntar si existe o no modo lógico alguno de decidir si una definición operativa determinada (o un índice) mide "realmente" el concepto o la variable teóricamente definidos. La respuesta a ambas cuestiones parece ser negativa.

Northrop sostiene esencialmente que no hay manera alguna de asociar las dos clases de conceptos o definiciones, excepto por vía de convención o de común acuerdo. La gente en general está simplemente de acuerdo en que debería emplearse una determinada definición operativa como medida de un determinado concepto, si las operaciones parecen razonables sobre la base de la definición teórica. Puede presumirse que, si varias definiciones operativas son posibles, se escogerán aquellas que parecen más apropiadas y al mismo tiempo más seguras. El carácter de "apropiado" ha de juzgarse inevitablemente sobre la base de la comprensión que uno tiene de la definición de teórica. Se emplea a veces el término de *validez aparente* para designar el carácter apropiado de un índice o de una definición operacional (11), p. 165). Idealmente, según lo señala Bridgman, las operaciones y las definiciones teóricas habrían de asociarse sobre la base de uno a uno ([2], pp. 23ss). O en otros términos, si cambiamos la operación, deberíamos servirnos de otro concepto. Sin embargo, semejante ideal es tal vez irreal en el estado actual de desarrollo de las ciencias sociales. Su aplicación conduciría sin duda alguna ya sea a una rigidez capaz de ahogar todo nuevo progreso metodológico o a una proliferación de conceptos teóricos [1].

¿Qué puede hacerse, pues? Podemos admitir la posibilidad de tener asociado un número de diversas operaciones o de índices a cada concepto teórico. Pero en tal caso podemos encontrarnos con una dificultad común: dichos procedimientos pueden dar resultados distintos. Uno de los procedimientos empleados para medir el prejuicio puede llevar a resultados que indiquen que nuestra "hipótesis" ha sido confirmada. En tanto que, en otro caso, otro procedimiento puede conducir a la conclusión opuesta. En cierto sentido, así es como se opera el progreso, a condición de que no conduzca a una disputa interminable acerca de cuál procedimiento mide "realmente" el prejuicio (cuya esencia se supone comprendida). Con objeto de prevenir confusiones importa darse cuenta de que la prueba efectiva se hace en términos de los conceptos tal como se los ha definido operacionalmente. *Por lo tanto, las proposiciones que comportan conceptos definidos teóricamente no son verificables directamente*. Así, pues, si se dan dos definiciones operativas distintas del prejuicio se verificarán dos hipótesis distintas.

Se ha admitido que puede resultar deseable tener asociada más de una operación con cualquier concepto teórico dado, y se ha señalado que semejantes operaciones pueden conducir a resultados diversos. Estamos ahora en condiciones de proporcionar un criterio eficaz, pragmático, para una definición teórica empíricamente satisfactoria del concepto. Supongamos que tenemos un concepto definido teóricamente y varias definiciones operativas susceptibles de asociarse con dicha definición teórica. Sobre la base de esta última definición, la mayoría de los científicos estarán probablemente de acuerdo en que algunas de las operaciones deberían eliminarse por cuanto no se aplican a lo que está contenido en la definición teórica. Pueden decidir, por ejemplo, que las preguntas relativas a las tendencias delictivas o los gustos musicales no deberían emplearse para medir el prejuicio. Pero puede haber varias operaciones que ocupen más o menos el mismo lugar en la opinión de dichos jueces. En otros términos: sobre la base de la definición teórica, los expertos pueden no estar en condiciones de ponerse de acuerdo acerca de que un determinado procedimiento operativo debiera escogerse con preferencia a otros. Podemos decir en tal caso que, *en la medida en que dichos diversos procedimientos dan resultados diferentes* (en igualdad de circunstancias), *la definición teórica es deficiente*, en el sentido de que necesita probablemente revisión o aclaración. Por ejemplo: el concepto *prejuicio* puede acaso haberse definido de tal modo que resulte demasiado vago. Tal vez se considerará necesario distinguir entre varias clases o dimensiones del prejuicio, asociando operaciones distintas a cada una de ellas. En una forma por el estilo de ésta –que se la reconozca explícitamente o no–, el proceso de investigación puede utilizarse para ayudar a aclarar los conceptos teóricos.

Parece, pues, haber en esta forma, dos lenguajes distintos, relacionados por una especie de diccionario al que se ha llegado por consenso, que permite asociar los conceptos de uno de ellos con los del otro. Los científicos piensan en el lenguaje teórico y realizan sus experimentos en el lenguaje operativo. No es necesario asociar operaciones con todos los conceptos del lenguaje teórico. Sin embargo, importa percatarse de que los conceptos que no han sido definidos operativamente no deberían por lo regular aparecer en enunciados que pretenden construir hipótesis comprobadas En efecto, si esto ocurre, las cuestiones planteadas por las "hipótesis" carecerán por lo regular de sentido desde el punto de vista operativo y conducirán probablemente a un debate interminable.

II.2. El nivel de medición: escalas nominales, ordinales y de intervalo

Acabamos de ver que el proceso consistente en pasar de los conceptos definidos teóricamente a los definidos operativamente no es en modo alguno directo. En efecto, al asociar un tipo de concepto con el otro, han de tomarse ciertas decisiones. Y en forma análoga, el proceso enderezado a seleccionar el modelo matemático o estadístico apropiado para emplearlo en una técnica de investigación determinada o en un procedimiento operativo comporta asimismo cierto número de decisiones importantes. Podría acaso pensarse que, una vez que un fenómeno ha sido medido, la elección de un sistema matemático sería cosa de simple rutina. Esto depende de lo que se entiende por *medida*. Si empleamos el término para referirnos únicamente a aquellos tipos de medición usualmente empleados en una ciencia como la física (*v.gr*. la medición de la longitud, del tiempo o la masa), entonces la elección de un sistema matemático no constituye prácticamente problema. Pero si ampliamos el concepto de la medición para incluir en él ciertos procedimientos menos precisos de empleo corriente en las ciencias sociales como se hará en este texto, entonces el problema se hace más complejo. Podemos, pues, distinguir entre distintos niveles de medición, y habremos de encontrar diversos modelos estadísticos apropiados a cada uno de ellos.[1]

Escalas nominales

La operación básica y a la vez más sencilla en toda ciencia es la de la clasificación. Al clasificar tratamos de separar elementos desde el punto de vista de determinadas características, decidiendo acerca de cuáles son más semejantes y cuáles más distintos. Nuestro propósito consiste en agruparlos por categorías que sea lo más homogéneas posible en comparación con las diferencias entre as categorías. Si la clasificación es útil, se verá que las categorías son homogéneas con respecto a *otras* variables [10]. Así por ejemplo, agrupamos unas personas de acuerdo con sus respectivas religiones (metodistas, presbiterianos, católicos, etc.) y vemos si la religión guarda alguna relación con el prejuicio o el conservadurismo político. Podríamos acaso hallar que los presbiterianos tienden a ser más conservadores que los católicos, siendo las puntuaciones de aquéllos relativamente altas en comparación con éstos. Si se hubieran seleccionado los individuos según el color del pelo, criterio de clasificación perfectamente adecuado, probablemente no se habrían encontrado diferencias significativas entre las clases en relación con otras variables estudiadas. En otros términos: las diferencias *entre* las clases de color del pelo habrían sido ligeras en comparación con las diferencias *dentro* de cada categoría.

Así, pues, la clasificación es fundamental para toda ciencia. Todos los demás niveles de medición, cualquiera que sea su precisión, comprenden básicamente la clasificación como operación mínima. Podemos, pues, considerar la clasificación como el nivel más bajo de medición, en el sentido mas amplio del término. Damos a las categorías nombres arbitrarios, a manera de etiquetas convenientes, sin formular supuesto alguno acerca de las relaciones entre aquéllas. Así por ejemplo, colocamos a los presbiterianos y a los católicos en categorías distintas pero no presuponemos que los unas sean "mayores que" o "mejores" que los otros. A condición de que las categorías sean exhaustivas (que comprendan todos los casos) y no se superpongan o se excluyan mutuamente (que ningún caso figure en más de una categoría), tenemos las condiciones mínimas necesarias para la aplicación de los métodos estadísticos. Se ha utilizado la expresión de *escala nominal* con referencia a ese nivel, el más simple de todos, de medición. Desde el punto de vista formal, las escalas nominales poseen las propiedades de simetría y transitividad. Por simetría entendemos que una relación que sea verdad entre A y B lo es también entre B y A. En tanto que por transitividad entendemos que si A = B y B = C, entonces A = C. Resumido, esto significa simplemente que si A está en la misma clase que B, o en una clase distinta, B está en la misma clase que A, o en una clase distinta,

[1] Para exámenes más detallados de estos distintos niveles de medición véanse [5], [7], [12] y [13].

respectivamente, y que si A y B están en una misma clase y B y C también en una misma clase entonces, A y C han de estar también en la misma clase.

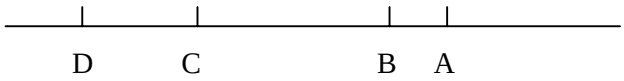
Habría que señalar que los números pueden asociarse arbitrariamente con cada categoría, pero esto no autoriza en ningún modo el empleo de las operaciones aritméticas usuales con dichos números. La función de los números, en este caso, es exactamente la misma que la de nombres, esto es la de designar las categorías. Es obvio que no tendría sentido alguno adicionar cifras de seguridad social y números de cuartos en un hotel. Pese a que nunca caeremos en la tentación de efectuar una operación tan ridícula como ésta se dan casos, sin embargo, en la investigación científica social, en que el absurdo no resulta en modo alguno tan obvio. Así pues, pese a que los valores numéricos puedan atribuirse arbitrariamente a diversas categorías el empleo de ciertas operaciones matemáticas de las más corrientes (suma, resta, multiplicación y división) requiere, con todo, la ejecución de ciertas operaciones *metodológicas* en el procedimiento de clasificación. Tendremos en breve ocasión de ver cuál deba ser el carácter de dichas operaciones.

Escalas ordinales

Resulta a menudo posible ordenar las categorías según el grado en que poseen una característica determinada, sin que por ello estemos en condiciones de decir cuántas poseen. Así pues, nos representamos un simple continuo a cuyo largo pueden ordenarse los individuos. Tal vez podamos colocar a los individuos de tal modo que nunca estén dos en el mismo lugar del continuo. Sin embargo, por lo regular existirá cierto número de conexiones. En tal caso no estamos en condiciones de distinguir entre determinados individuos, y los hemos agrupado juntos en una misma categoría. Pero estamos en condiciones, con todo, es decir que todos esos individuos tienen puntuaciones superiores a las de otros individuos determinados. Podemos, por ejemplo, clasificar familias conforme a su respectivo estado socioeconómico, en “superior”, “media superior” “media inferior” e “inferior”. Podríamos incluso limitarnos a dos categorías, la “superior” y la “inferior”

El tipo de medición que estamos examinando se sitúa manifiestamente a un nivel algo superior al que empleamos para obtener una escala nominal, ya que con ella podemos no sólo agrupar a los individuos en categorías separadas, sino ordenar además estas categorías unas con respecto a las otras. Designamos este nivel de medición como “escala ordinal”. Además de poseer las propiedades simétricas de la escala nominal, la escala ordinal es asimétrica en el sentido de que algunas relaciones especiales pueden ser verdad entre A y B y *no* serlo, en cambio, entre B y A. Así por ejemplo, la relación “mayor que” (>) es asimétrica, por cuanto si A > B, no puede ser cierto que B > A. La transitividad, en cambio, sigue subsistiendo, ya que si A > B y B > C, entonces A > C. Son estas propiedades, por supuesto, las que nos permiten colocar A, B, C... a lo largo de un mismo continuo.

Importa advertir que el nivel ordinal de medición no nos proporciona información alguna acerca de la *magnitud* de las diferencias entre los elementos. En efecto sabemos solamente que A es mayor que B, pero no sabemos cuánto mayor sea. Como tampoco podemos decir que la diferencia entre A y B sea menor que la que hay entre C y D.[2] Por consiguiente, no podemos adicionar o restar distancias sino en un sentido muy restringido. Así por ejemplo, si tuviéramos las siguientes relaciones



podemos decir que la distancia

$$\overline{AD} = \overline{AB} + \overline{BC} + \overline{CD}$$

pero no podemos tratar de comparar entre sí las distancias $\overline{AB}$ y $\overline{CD}$. En otros términos, cuando transformamos relaciones de orden en operaciones matemáticas, no podemos por lo regular servirnos de las operaciones usuales de suma, resta, multiplicación y división. En cambio sí podemos emplear las operaciones “mayor que” y “menor que”, siempre que esto presente alguna utilidad.

Escalas de intervalo de proporción

[2] Se ha empleado el término *métricamente ordenado* para designar escalas en las que es posible ordenar la magnitud de las diferencias entre elementos Véase [7].

En el sentido restringido de la palabra, el término *medida* puede emplearse para designar casos en los que no sólo estamos en condiciones de ordenar objetos según el grado en que poseen una característica determinada, sino que podemos indicar asimismo la distancia exacta entre ellos. Si esto es posible, podemos obtener lo que se designa como una *escala de intervalo*. No resulta difícil comprender que el nivel de medición de escala de intervalo requiere el establecimiento de algún tipo de unidad física de medición, que pueda considerarse por todos como una norma común y sea repetible, esto es, que pueda aplicarse indefinidamente con los mismos resultados. El largo se mide en términos de pie o metros, el tiempo en segundos, la temperatura en grados Fahrenheit o centígrados, el peso en libras o gramos, y el ingreso en dólares, etcétera. Por otra parte, no existen unidades semejantes de inteligencia, autoridad o prestigio, en las que todos los sociólogos puedan ponerse de acuerdo y de las que se pueda suponer que permanecen constantes de una situación a otra. Dada una unidad de medida, resulta posible decir que la diferencia entre dos marcas es de veinte unidades, o que una diferencia es dos veces mayor que otra. Esto significa que es posible adicionar o restar marcas en forma análoga a como añadimos pesas a una balanza o quitamos 6 pulgadas de una tabla con una sierra ([3], pp. 296 a 298). Y en forma semejante podemos adicionar los ingresos de marido y mujer, en tanto que carece de sentido adicionar sus cuotas de inteligencia (IO).

Si además es posible situar en la escala un punto cero absoluto o no arbitrario, entonces tenemos un nivel de medición algo mayor, que suele designarse como *escala de razón* (*ratio scale*). En tal caso estamos en condiciones de comparar marcas sirviéndonos de su proporciones. Podemos, por ejemplo, decir que una marca es dos veces más alta que otra. Si el punto cero fuera arbitrario, como es el caso en las escalas de grados centígrados y de Fahrenheit, aquello no sería legítimo. Así, por ejemplo, no decimos que 70º Fahrenheit son el doble de 35º centígrados si bien podemos decir que la *diferencia* entre dichas temperaturas es la misma que entre 105º C y 70º Fahrenheit. Sin embargo, prácticamente en todos los casos que conoce el autor, esta distinción entre escala de intervalo y escala de proporción es puramente académica, ya que es extremadamente difícil encontrar una escala legítima de intervalo que no sea al propio tiempo una escala de proporción. Esto se debe al hecho de que, una vez establecida la magnitud de la unidad, es casi siempre posible concebir cero unidades, pese a que nunca podamos hallar un cuerpo que no posea largo o masa, u obtener una temperatura de cero absoluto. Así pues, prácticamente en todos los casos en que se dispone de una unidad de la aritmética, incluidas las raíces cuadradas, las potencias y los logaritmos.

Suscítanse algunas importantes cuestiones acerca de la legitimidad de servirse de escalas de intervalo en el caso de cierto número de variables sociológicas y sociopsicológicas. Infortunadamente, no resulta posible discutir estas cuestiones en detalle en un texto general como el presente, pese a lo cual vamos a mencionar brevemente alguna de ellas. Se sostiene en ocasiones que una variable como, por ejemplo, el ingreso no constituye en realidad, si se calcula en dólares, una escala de intervalo ya que una diferencia de $ 1000 posee un significado psicológico, distinto según que se dé entre ingresos de $ 2000 y $ 3000 o entre ingresos de $ 30000 y $ 31000. Al parecer, este argumento confunde la cuestión. Porque lo que aquí se dice efectivamente es que el ingreso calculado en dólares y el “ingreso psicológico” (a suponer que se lo pueda medir en términos de alguna unidad) no se relacionan directamente o en forma lineal. Y esto es una cuestión de hecho que carece de importancia en relación con la cuestión de saber si existe o no una unidad legítima de medida.

- Al llevar a cabo la enumeración de actos de conducta, de personas, de ocupaciones, o de grupos de diversa índole, se obtendrán muchas escalas de razón. “ proporción de actos criminales, por ejemplo, se obtiene contando, el número registrado de tales actos y comparándolos con la base de población. La mayor parte de nuestros datos censales de ciudades, estados o regiones, se obtienen contando varias clases de gente y dividiendo las cifras así obtenidas por la base de la población: por ciento urbano, porcentaje de la fuerza de trabajo en situación de desempleo, tamaño promedio de las familias, porcentaje de no blancos, etcétera. La complejidad de la división del trabajo puede ser medida en función del número de ocupaciones diferentes, o bien puede obtenerse un índice de la complejidad organizativa contando el número de oficinas sucursales. En ocasiones surgen discusiones acerca de si tales mediciones constituyen realmente “escalas de razón” (ver Coleman [4] para un excelente planteamiento de este problema). Si se toma el punto de vista estrictamente operativo, según el cual la medida utilizada constituye la definición de la variable de interés, cabrá poca duda de que se habrá así obtenido una legítima escala de razón, ya que han sido contadas unidades precisas, y tales unidades han sido tomadas como equivalentes (y por tanto intercambiables). De esta manera, si añadimos a una determinada población 1000 negros y le restamos 1000 blancos, haremos la suposición fundamental de que, en orden a la medida usada, no hay diferencia, bien sean unos u otros los negros usada, no hay diferencia, bien sean uno u otros los negros o los blancos implicados. Por otra parte el punto cero

está bien definido. La afirmación de que el porcentaje de no blancos en una ciudad es igual a cero no ofrece ambigüedad.

- En cuantas ocasiones surge una discrepancia acerca de la adecuación de las medidas enumeradas, o de si éstas legitiman la adopción de mediciones relativas al nivel proporcional, experimento una vehemente sospecha de que el problema básico tiene un carácter totalmente distinto, a saber: el de la relación entre la medida utilizada y la construcción teórica que intenta medirse. Por ejemplo: la proporción de desempleados puede ser utilizada como indicación del mal funcionamiento de la economía; un porcentaje minoritario, como indicador de una amenaza planteada por la minoría, o un porcentaje urbano como indicador de la influencia ejercida por los valores urbanos. En tales casos nunca podrá la estadística resolver *per se* una controversia, resultando por ello necesario soslayar los problemas básicos, suponiendo, por el contrario, que sólo nos interesa la variable que intentamos medir.
- Puede suscitarse otra cuestión a propósito de si es posible o no conseguir una escala de intervalo en materia de medición de la actitud. Se han efectuado varios intentos enderezados a conseguir dicho fin. En el método Thurstone de los intervalos de apariencia igual, se pide a los jueces que agrupen objetos en montones situados a distancias iguales a lo largo del continuo de la actitud ([ll], pp. 359 a 365). Se discurre esencialmente diciendo que, si se da un alto grado de consenso entre los jueces, puede emplearse legítimamente una escala de intervalo. Este procedimiento, así se sostiene, es esencialmente el mismo que se emplea para obtener escalas de intervalo en otras disciplinas. Este argumento parece legítimo, a condición que se dé efectivamente un alto grado de consenso entre los jueces y a condición que éstos dispongan de un gran número de montones en los que se puedan clasificar los objetos. Así, por ejemplo, si se vieran obligados a clasificar los objetos en uno de tres o cuatro montones, podríamos contar con un grado elevado de consenso, debido simplemente a la tosquedad del instrumento de medición. Habría, en efecto, tal margen de variabilidad dentro de cada montón, que difícilmente podría sostenerse que los objetos de los diversos montones estaban a igual distancia unos de otros. Pero aun admitiendo una concordancia perfecta y la máxima libertad en el agrupamiento de los objetos en montones, aun así sigue el método de Thurstone presentando dificultades por lo que se refiere al concepto de la unidad de referencia. Se hace necesario postular que es la existencia de dicha unidad la que hace posible el acuerdo entre los jueces. Puede afirmarse razonablemente que, en este punto del desarrollo de la medición de la actitud, la mayoría de las técnicas dan unas aproximaciones muy mediocres de las escalas de intervalo. Probablemente de muchas de ellas no debiera siquiera considerarse que proporcionan escalas ordinales legítimas. Las consecuencias de ello por lo que se refiere al análisis estadístico se irán haciendo más claras a medida que vayamos avanzando.

II.3. Medición y estadística

Hemos visto que existen diversos niveles de medición, con sus propiedades peculiares cada uno. Debe observarse que estos distintos niveles forman una escala cumulativa ellos mismos. En efecto, la escala ordinal posee todas las propiedades de la escala nominal además de la ordinal. A su vez, la escala de intervalo posee todas las propiedades de las escalas nominal y ordinal y, además, una unidad de medida, en tanto que la escala de proporción presenta el nivel más elevado, ya que posee no sólo una unidad de medida, sino, además, un cero absoluto. El carácter cumulativo de estas escalas significa que, al analizar nuestros datos estamos siempre autorizados a descender uno o más grados en el nivel de medición. En efecto, si tenemos una escala de intervalo, tenemos al propio tiempo una escala ordinal, y podemos servirnos de esta circunstancia en nuestros análisis estadísticos. Esto resultará a veces necesario, cuando no dispongamos de técnicas estadísticas o éstas sean en algún modo deficientes en cuanto a manipular la variable como escala de intervalo. Sin embargo, al proceder así perdemos información. Así, por ejemplo, si sabemos que Jones tiene un ingreso de $ 11000 y Smith uno de $ 6 000 y sólo nos servimos del hecho de que Jones cuenta con el mayor de los dos ingresos en cuestión, entonces desperdiciamos la información relativa a que la diferencia de los ingresos es de $5000. Por lo tanto, en la mayoría de los casos resultará ventajoso servirnos del nivel de medición más alto que podamos legítimamente adoptar.

¿Y qué puede decirse del proceso inverso consistente en subir la escala de medición, pasando, por ejemplo, de la ordinal a la de intervalo? Estamos a veces tentados de hacerlo, ya que estaríamos en condiciones de servirnos de técnicas estadísticas más potentes. Es incluso posible que lo hagamos sin darnos cuenta en absoluto de lo que ha sucedido exactamente. Importa percatarse de que no hay nada en los procedimientos estadísticos o matemáticos de los que nos servimos en última instancia que nos permita

verificar la legitimidad de nuestros métodos de investigación. *El empleo de un determinado modelo matemático supone que se ha alcanzado cierto nivel de medición*. La responsabilidad en cuanto a decidir si sus procedimientos operativos permiten o no el empleo de determinadas operaciones matemáticas recae exclusivamente sobre el investigador. Éste ha de decidirse en primer lugar por el nivel de medición adecuado, y esto decidirá a su vez acerca del sistema matemático apropiado. En otros términos: un determinado modelo matemático puede asociarse a cierto nivel de medición conforme a las consideraciones examinadas en la sección precedente. Así, por ejemplo, las operaciones aritméticas corrientes sólo pueden emplearse por lo regular con las escalas de intervalo y de razón.

Nos enfrentamos aquí una vez más con el problema de tener que traducir de un lenguaje a otro. El lenguaje operativo comporta determinadas operaciones físicas , tales como el empleo de una unidad de medida. El lenguaje matemático, a su vez, implica un sistema totalmente abstracto de símbolos y operaciones matemáticas, y es útil no sólo porque es preciso y está altamente desarrollado, sino debido también a que su carácter abstracto permite la aplicación a una gran variedad de problemas empíricos. Las matemáticas se sirven del razonamiento deductivo por el que se pasa de un conjunto de definiciones, supuestos y reglas de operación a un conjunto de conclusiones mediante un razonamiento puramente lógico. En sí mismas, las matemáticas nada no dicen acerca de la realidad, ya que todas las conclusiones están contenidas en las definiciones, los supuestos y las reglas originales, no habiéndose determinado empíricamente. Así, pues, si han de ser de alguna utilidad para el científico las conclusiones matemáticas han de traducirse inversamente a los lenguajes operativos y teóricos[5].

Sostenemos, pues, que no es legítimo servirse de un sistema matemático que comporta las operaciones de sumar o restar, si esto no está legitimado por el método de medición. Aunque el sentido de este hecho sólo nos resultará plenamente claro cuando empecemos a servirnos de las diversas escalas de medida, estamos diciendo en realidad que no podemos remontar legítimamente en la jerarquía de medición, a menos que el proceso mismo de la medición haya sido mejorado. Lo que ninguna manipulación matemática puede hacer. ¿Cómo decidimos, pues, cuál nivel de medición es el legítimo? Infortunadamente, el problema no es tan sencillo como podría suponerse. Unos pocos ejemplos bastarán para dar una idea de la complejidad del mismo.

*Para ilustrar uno de estos problemas es necesario distinguir las escalas ordinales y de intervalo de la escala *parcialmente ordenada* que resulta de la combinación de dos o más escalas ordinales (o de intervalo) en un solo índice. Ocurre con frecuencia, en sociología y en las otras ciencias sociales, que aquello que por lo pronto parece ser una simple escala ordinal (o de intervalo) es en realidad una combinación de varias escalas ordinales (o de intervalo), con el resultado de que no puede hacerse una clasificación inequívoca de individuos sin adoptar previamente ciertas otras decisiones. Tomemos, por ejemplo, el caso de la condición socioeconómica. Por lo regular, solemos determinar la condición de una persona examinando cierto número de criterios distintos, tales como su ingreso, ocupación, educación, antecedentes familiares o la zona de residencia. Si A se clasifica mejor que B según todos y cada uno de dichos criterios, entonces A puede obviamente clasificarse como más alto que B por lo que se refiere a la condición general. Pero, ¿qué ocurre si A tiene un ingreso superior al de B, y éste, en cambio, un nombre de familia más prominente? ¿Cuál de los dos ocupa en este caso el mejor rango social? Tenemos aquí varias alternativas. La primera consiste en dejar de lado la noción de condición general y pensar en términos de dimensiones separadas de la misma, cada una de las cuales pueda acaso admitir un nivel ordinal de medición. Terminamos así no con una, sino con varias escalas ordinales, y la cuestión empírica está en saber hasta qué punto las distintas dimensiones puedan relacionarse entre sí . Por supuesto, si existe una relación perfecta entre todas las dimensiones, la cuestión se convierte en puramente académica, ya que A, si es superior a B en cada una de las dimensiones, lo será también en todas ellas. En la práctica, sin duda, esto no ocurre nunca.

*Nuestra segunda alternativa consiste en tratar de “forzar” la aplicación de una escala ordinal a los datos, adoptando algunas decisiones acerca del peso relativo de cada dimensión y de las equivalencias que ello implica. Así, por ejemplo, si podemos admitir que un año suplementario de instrucción equivale a $ 1338.49 de ingreso suplementario, podemos traducir las unidades educativas en unidades de ingreso, llegando así a una escala unidimensional. Obviamente, el problema de traducir los antecedentes familiares o el área de residencia es más complicado todavía. El método de medición que aquí estamos examinando comporta un tipo de construcción de índice. Baste decir que semejante construcción de índice comporta usualmente algunas decisiones arbitrarias a propósito de los pesos relativos que haya que atribuir. Si el sistema de ponderación se deja justificar, entonces puede emplearse una escala ordinal; en caso contrario, subsiste la duda acerca de si los individuos pueden o no clasificarse legítimamente en relación con el rango.

*Uno de los métodos comúnmente empleados para obtener una escala ordinal consiste en servirse de uno o más jueces para clasificar a los individuos conforme a un criterio como por ejemplo, el del poder o del

prestigio. Supongamos, para simplificar, que no hay más que un juez y que se le ha impuesto la tarea de clasificar a los individuos según su "posición social" en la localidad. Suponiendo que la persona coopere, el método empleado nos garantiza la obtención de una escala ordinal *independientemente* de cómo los individuos se comparen realmente a los ojos del juez. Es posible que, si se hubiera empleado otro método, no se habría obtenido escala ordinal alguna. Si se hubiera utilizado una técnica de comparaciones apareadas, en la que se pronunciaran juicios entre cada combinación por pares, el juez podría haber tasado a Smith más alto que a Brown, a éste más alto que a Jones, pero a este último más alto que al primero, Smith; violando en esta forma la propiedad de transitividad de las escalas ordinales. El investigador ha de proceder ahora a una elección. Puede llegar a la conclusión de que existe una escala parcialmente ordenada de una clase u otra. O puede considerar que el juez es inconsecuente o comete "error". Como lo señala Coombs, este problema relativo a lo que haya que designar como error de medición es un dilema básico con el que se encuentra el sociólogo ([7], pp. 485 a 488). En términos generales, éste puede adoptar un alto nivel de medición y considerar las desviaciones del tipo que se acaba de señalar como errores de medición, o puede descender a un nivel más bajo de ésta.

"Puede ilustrarse el mismo dilema en el caso de la escala de :Guttman. En el tipo perfecto de ésta, las preguntas tienen una propiedad. cumulativa que justifica la adopción de una escala al [14]. Las preguntas pueden ordenarse de modo que se vaya pasando de un límite inferior a un límite superior, de tal modo que el tipo exacto de respuesta de un individuo pueda reproducirse a partir de su puntuación total. Así, por ejemplo, si se tienen cinco problemas aritméticos que vayan del más fácil al más difícil; la persona que resuelva el más difícil estará también lógicamente en condiciones de resolver los más fáciles. Si resuelve correctamente tres de los problemas, éstos serán los tres más fáciles, fallando en los otros dos. En una escala perfecta de distancia social, las preguntas relativas al prejuicio pueden disponerse conforme al grado de la intimidad de contacto con la minoría considerada. Una persona que esté dispuesta a casarse con un negro, estará dispuesta, por supuesto, a vivir en la misma calle de uno de ellos; si lo acepta como vecino, no tendrá inconvenientes en sentarse a su lado en el autobús. Así, pues, podemos ver en la escala perfecta de Guttman que la persona que conteste afirmativamente cuatro preguntas habrá contestado exactamente las mismas que una persona con tres afirmaciones, *más una*. Si la escala sólo estuviera parcialmente ordenada, podría decirse que en ciertos aspectos A tiene más prejuicios que B, y en otros, aspectos menos, ya que los dos individuos han aceptado combinaciones distintas de preguntas.

* Sin embargo, en la práctica raramente alcanzamos una escala perfecta de Guttman, si es que la alcanzamos alguna vez. En efecto, hay siempre algunas personas cuyo tipo de respuestas se desvía del tipo ideal. ¿Son éstas acaso inconsecuentes porque aceptan a un negro como vecino pero se niegan a sentarse a su lado en el autobús? Tal vez. Pero, por otra parte, tal vez no sea así. A menos que el investigador esté dispuesto a *suponer* que dispone de una escala ordinal legítima, no puede sostener que el individuo considerado cometa error. Y si el número de errores aumenta, empezamos a sospechar de nuestra escala. Por otra parte, siempre estamos dispuestos a tolerar cierto número relativamente pequeño de errores. Es este principio el que se halla a la base de la decisión relativa a aceptar la escala de Guttman como escala ordinal, si el número de errores, medido por el coeficiente de reproductibilidad, es muy pequeño. Sin embargo, conviene percatarse del hecho de que la decisión es hasta cierto punto arbitraria, y de que en última instancia nos enfrentaremos con el problema de decidir a qué debamos llamar error.

*Estos ejemplos deberían bastar para indicar que no siempre es cosa fácil decidir cuál tipo de escala pueda emplearse legítimamente. Desde un punto de vista ideal, habría que servirse de una técnica de reunión de datos que permita los niveles más bajos de medición, si éstos son los únicos que los datos admiten, antes que recurrir a técnicas que adapten violentamente la escala a los datos. Así, pues, el método de las comparaciones apareadas sólo dará una escala ordinal si el juez está efectivamente en condiciones de clasificar a los individuos. Por otra parte, si se le invita a colocarlos en un orden preciso de clasificación, habrá de hacerlo, así crea o no que esto puede lograrse legítimamente. Habiéndose servido de este último método de reunión de datos y no estando en condiciones de demostrar empíricamente que los individuos pueden ordenarse sin violentar los datos, habrá de *suponer* la existencia de un solo continuo.

Con objeto de insistir en el hecho de que toda técnica estadística considerada presupone siempre un nivel específico de medición, nos acostumbraremos a indicar siempre el nivel de medición requerido por cada procedimiento. Al elegir entre procedimientos alternativos una de las preguntas más importantes a formular es ésta: "¿Es legítimo aceptar el nivel de medición que una determinada técnica requiere?" Si no lo es, tal vez deba encontrarse un procedimiento alternativo. Si la única consideración fuese el nivel de medición, se simplificaría el problema de la elección entre procedimientos alternativos.

Encontramos con frecuencia, sin embargo, que ciertos procedimientos que no tienen grandes exigencias en cuanto a la medida, y que por tal razón parecen preferibles, resultan menos satisfactorios en

relación con otras características deseables. Así se ve uno enfrentado con decisiones difíciles, en las que está implícita la necesidad de sopesar la seriedad relativa de las diversas clases de presunciones violadas. En tales casos puede resultar deseable analizar nuestros datos mediante la aplicación de diferentes métodos, observando si las conclusiones así obtenidas difieren entre sí en forma considerable.

En este punto puede ocurrir que nuestro examen de estos, diferentes niveles de medición y de los problemas de elección entre pruebas y mediciones alternativas, no nos ilustre gran cosa. Uno de los peligros de la estadística "de recetario" consiste en una excesiva simplificación de los criterios y los problemas implícitos en la adopción de decisiones relativas al análisis de datos. Es imposible exagerar la importancia que tiene, al utilizar cualquier técnica estadística, el tener presente las presunciones implícitas que el procedimiento requiere. En el curso del presente examen, una de las primeras preguntas a formularse es la relativa al nivel de medición que puede legítimamente aceptarse.

II. 4. Organización del libro

La organización de los restantes capítulos viene determinada por ciertas consideraciones, la primera de las cuales consiste en presentar ante todo las ideas más simples, pasando gradualmente de mayor complejidad. Como cada una de las secciones pone el conocimiento de materiales que previamente han sido tratados, resulta conveniente seguir esta organización, pasando por alto solamente los párrafos o secciones precedidos de asterisco. El capítulo XIV, por excepción, puede ser saltado en su totalidad, o bien englobado con las prueba y procedimientos "no paramétricos" contenidos en los capítulos XVI y XVIII. El capítulo XXI, acerca del muestreo, puede ser leído en relación con el capítulo IX relativo a la probabilidad, aun cuando el capítulo sobre muestreo contiene varias secciones que sólo podrán entenderse cuando hayan sido leídos los capítulos XI, XIII y XVI. Lo fundamental del capítulo XVII podrá ser asimilado sin haber previamente abordado el capítulo XVI sobre análisis de diferencias. Se recomienda, en general, que se estudien los distintos temas en el orden que se presentan.

Los instrumentos estadísticos no son fácilmente agrupables bajo uno o dos apartados, y por tal razón los títulos que ostentan las principales divisiones del libro son sólo parcialmente adecuados, limitándose por el contrario a centrar la atención primaria. La Segunda Parte se limita al tema de la estadística descriptiva, en tanto que en las partes Tercera y Cuarta el principal, aunque no el único, foco de atención, se refiere a la inducción, a la prueba de hipótesis y a la estimación de parámetros de población basada en datos de muestreo. En las partes Segunda y Tercera nos limitaremos casi por entero a los procedimientos que traen implícita una sola variable por vez, en tanto que en la Cuarta Parte pasamos a problemas más difíciles tales como el manejo simultáneo de dos o más variables.

Entrelazada en estas distinciones entre descripción e inducción, así como entre las estadísticas univariables y bivariables o multivariables, se observa un tercer principio organizativo, a saber: el relativo a los niveles de medición para cada una de las variables. Muchos de los títulos de los capítulos señalan este nivel de medición, pero tal vez el mejor método para lograr una perspectiva resumida del contenido consista en acudir al cuadro de pruebas y mediciones que aparece en las guardas. En su primera columna aparecen los procedimientos a usar con variables simples. Vemos allí que en el capítulo III nos ocuparemos de las mediciones muy simples (porcentajes, proporciones y razones), utilizados tanto con las dicotomías como con las escalas generales nominales con más de dos categorías. Las pruebas de hipótesis relativas a escalas nominales simples serán estudiadas en los capítulos X, XI y XII. Las medidas (mediana, desviación cuadril) adecuadas para ser utilizadas con una escala ordinal, sencillas, serán examinadas brevemente en los capítulos V y VI, en tanto que en el capítulo X se discutirá una muy sencilla prueba (la binomial), aplicable a datos ordinales. Dedicaremos algo más de nuestra atención a las escalas de intervalos y razones, las que examinaremos en los capítulos IV a VII, relativos a procedimientos descriptivos univariados, y de nuevo en los capítulos XI y XII de la Tercera Parte, relativos a la estadística inductiva.

Comenzando con el capítulo XIII volveremos nuestra atención a las relaciones entre dos o más variables, lo que desde luego supone que habremos de ocuparnos del nivel de medición de la variable segunda (y adicional), tanto como de la primera. Las columnas 2 a 5 del cuadro ofrecen varias combinaciones en relación con el nivel de medición de las dos variables. Por ejemplo: la casilla superior de la columna 2 se refiere a aquellas situaciones en que se dan dos dicotomías relacionadas entre sí (por ejemplo, sexo contra preferencias políticas). En la segunda casilla de la columna 2 se tiene en cuenta la posibilidad de que la primera escala nominal tenga más de dos categorías (por ejemplo, protestantes, católicos y judíos. En la tercera casilla una de las variables es una dicotomía (por ejemplo, sexo) en tanto que la segunda es una escala ordinal, y así sucesivamente. Hay una sola casilla sin ocupar, a saber: aquella en que una variable se mide en el nivel ordinal y la segunda en el nivel de intervalo o de razón. Aun cuando tales

situaciones pueden por supuesto ser resueltas, carecemos de instrumentos realmente satisfactorios que no requieran una pérdida de información al reducir el nivel de medición en cualquiera de las dos variables. No hay necesidad alguna de llenar las casillas situadas sobre las que ocupan la diagonal del cuadro, ya que las cubren aquellas situadas bajo dicha diagonal.

Resulta prematuro en estos momentos poner a discusión cada una de las posibilidades relacionadas en el cuadro. El punto más importante a señalar es el de que el nivel de medición afectado constituye una de las consideraciones más importantes al llevar a cabo una elección entre diversos procedimientos. La elección es relativa, aunque no enteramente, simple, en tanto uno se encuentre limitado al uso de sólo dos variables. Resulta mucho más difícil en el caso del análisis multivariado, en el que con frecuencia puede uno trabajar con cinco o incluso hasta quince o veinte variables al mismo tiempo, y donde resalta como sumamente improbable que todas ellas sean medidas al mismo nivel, y donde a menudo es poco deseable utilizar demasiados métodos de prueba y medición. En los capítulos XV, XVI, XIX y XX se examinan estos problemas de análisis multivariado. En ciertos lugares, particularmente al final de los capítulos XIV y XX, se encontrarán, en forma resumida, algunas de las consideraciones relativas a la selección entre procedimientos alternativos.

Como se observará, no todas las combinaciones posibles son manejadas en este texto con el mismo grado de minuciosidad. Ocurre así, no sólo por limitaciones de espacio y por la necesidad de detenerse en el examen de las ideas fundamentales, sino por la teoría estadística se encuentra mucho más avanzada en aspectos. Se ha trabajado mucho más, en particular, en el sector de la llamada “estadística paramétrica”, relativa a las escalas de intervalo y de razón, que en el de los procedimientos ordinales, por la cual nuestros instrumentos para el uso de las escalas de intervalo y de tiempo están mucho más desarrollados, especialmente en el caso del análisis multivariado. La diferencia entre las escalas de intervalo y razón no ha sido tampoco explotada en la teoría estadística, por lo menos hasta el nivel que a nosotros nos interesaría. La razón básica estriba en que los modelos estadísticos con los que generalmente trabajamos están basados en una ecuación lineal general que es aditiva, en lugar de abarcar razones o variables. Por tal motivo, y para todo objetivo práctico, no es necesario tener presente tal distinción según se avanza en la lectura. Será empero necesario acudir periódicamente al cuadro de las guardas.

El lector hará bien en acostumbrarse a explicar en sus propios términos el significado de los conceptos importantes. Los nuevos conceptos introducidos en este capítulo son:

la escala de intervalo,
la escala nominal,
la definición operativa,
la escala ordinal y
la escala de razón.

BIBLIOGRAFÍA

Blalock, H. M.: "The Measurement Problem: A Gap between the Languages of Theory and Research"; en H. M. Blalock y Ann B. Blalock (E.) *Methodology in Social Resear*ch, McGraw-Ñill Book Company, Nueva York, 1968, cap. 1.

Bridgman, P. W.: *The Logic of Modern Physics*, The Macmillan Co., Nueva York, 1938, pp. 1-39.

Cohen, M. R. y E. Nagel: *An Introduction to Logic and Scientifie Method*, Harcourt, Brace and Company, Inc., Nueva York, 1937, caps. 12 y 15.

Coleman, James S.: *Introduction to Mathematical Sociology*, The Free Press, Nueva York, 1964, cap. 2.

Coombs, C. H., H. Raiffa y R. M. Thrall: "Some Views on Mathematical Models and Measurement Theory", *Psychological Review*, vol. 61, pp. 132-144, marzo de 1954.

Coombs, C. H.: *A Theory of Data*, John Wiley & Sons, Inc., Nueva York, 1964.

Coombs, C. H.: 'Theory and Methods of Social Measurement, en L. Festinger y D. Katz (ed.), *Research Methods in the Behavioral Sciences*, The Dryden Press, Inc., Nueva York, 1953, pp. 471-535.

Lundberg, G. A.: *Foundations of Sociology*, The Macmillan Company, Nueva York, 1939, caps. 1-2.

Northrop, F. S. C., *The Logic of the Sciences and the Humanities*, The Macmillan Company, Nueva York, 1947, caps. 5-7.

Radcliffe-Brown, A. R.: *A Natural Science of Society*, The Free Press of Glencoe, Ill., Nueva York, 1957, pp. 28-42.

Selltiz, C., M. Jahoda, M. Deutsch y S. W. Cook: *Research Methods in Social Relations*, Henry Holt and Company, Inc., Nueva York, 1959, caps. 5 y 10.

Senders, V. L., *Measurement and Statistics*, Oxford University Press, Nueva York, 1958, cap. 2.

Stevens, S. S.: "Mathematics, Measurement, and Psychophysic", en S. S. Stevens (ed.), *Handbook of Experimental Psychology*, John Wiley & Sons, Inc., Nueva York, 1951, pp. 1-49.

Stouffer, S. A., et al.: *Measurement and Prediction*, Princeton University Press, Princeton, N. J., 1950, caps. 1 y 3.

Weiss, R. S.: *Statistics in Social Research*, John Wiley & Sons, Inc., Nueva York, 1968, cap. 2.