

AGENCY FOR INTERNATIONAL DEVELOPMENT
 WASHINGTON, D. C. 20523
BIBLIOGRAPHIC INPUT SHEET

FOR AID USE ONLY

Batch 57

1. SUBJECT CLASSIFICATION	A. PRIMARY	TEMPORARY
	B. SECONDARY	

2. TITLE AND SUBTITLE
 Guia para la medicion de resultados educacionales y sus usos; Sem.no.5: Aplicacion de las mediciones de resultados

3. AUTHOR(S)
 Glaser, Ezra

4. DOCUMENT DATE 1975	5. NUMBER OF PAGES 61p.	6. ARC NUMBER ARC
--------------------------	----------------------------	----------------------

7. REFERENCE ORGANIZATION NAME AND ADDRESS
 Georgetown

8. SUPPLEMENTARY NOTES (*Sponsoring Organization, Publishers, Availability*)
 (In English and Spanish; English, 50p.: PN-AAD-475)

9. ABSTRACT
 (EDUCATION R&D)

Volume 5 of a six-volume set of seminar guidebooks, each of which focuses on an aspect of educational outcome measurement and uses of such measurements in order to assist educational policymakers in developing countries to assess their needs and programs, is concerned with the topic of applying outcome measures. It is designed to convey to participants an understanding of the potential usefulness of statistical science to the program administrator. The basic theme of the seminar is that there are formidable difficulties in collecting, analyzing, and interpreting data. Thus, formal statistical methodology is needed. Moreover, it is in program accountability and output programming that data interpretations are needed. A correlary theme of the seminar is that administrators can draw on other professions and combine input from two sources: education and statistical science. The guidebook presents and discusses seven generic problems of acquiring, analyzing, and interpreting data for management uses. These concern characterizing and comparing a group of people, objects, or events; measuring the variability assignable to program influences; determining what principles govern collection of data for all types of analyses; and arranging for the economical collection of precise data. It also covers the application of statistical principles to output-oriented management: to management by objectives, output budgeting, cost-effectiveness studies, and benefit-cost studies of educational programs.

10. CONTROL NUMBER PN-AAD-480	11. PRICE OF DOCUMENT
	12. DESCRIPTORS
	13. PROJECT NUMBER
	14. CONTRACT NUMBER AID/CM/ta-C-73-8 GTS
	15. TYPE OF DOCUMENT

SEMINARIO N° V

APLICACION DE LAS MEDICIONES
DE RESULTADOS

Guía para la Medición de Resultados Educativos y Sus Usos



COSTA RICA - 1976

**GUIA PARA LA MEDICION
DE RESULTADOS EDUCACIONALES Y SUS USOS**

SEMINARIO N° I
TIPOS DE MEDICIONES DE RESULTADOS

SEMINARIO N° II
UTILIZACION DE LAS MEDICIONES DE RESULTADOS

SEMINARIO N° III
DISEÑO DE METODOS PARA MEDIR RESULTADOS
(RENDIMIENTO Y ACTITUDES)

SEMINARIO N° IV
DISEÑO DE METODOS PARA MEDIR RESULTADOS
(ECONOMICOS Y SOCIALES)

SEMINARIO N° V
APLICACION DE LAS MEDICIONES DE RESULTADOS

SEMINARIO N° VI
CONSECUENCIAS DE LA INFORMACION OBTENIDA Y
PASOS PARA SU APLICACION PRACTICA

Selma J. Mushkin, *Directora*
Bradley B. Billings, *Investigador Asociado*
PUBLIC SERVICES LABORATORY
Georgetown University

SEMINARIO N°

V

**APLICACION DE LAS MEDICIONES
DE RESULTADOS**

EZRA GLASER

PUBLIC SERVICES LABORATORY
Georgetown University
Noviembre de 1975

**Edición en español: Centro Multinacional de Investigación
Educativa (CEMIE)
San José, Costa Rica
1976**

Esta publicación, en su versión en inglés, fue producida por el PUBLIC SERVICES LABORATORY de la Universidad de Georgetown con sede en Washington, D. C., como parte de un proyecto patrocinado y financiado por la Agencia para el Desarrollo Internacional. Sin embargo, los puntos de vista que aquí se expresan no son necesariamente los de la mencionada Agencia.

La edición en español la hizo el Centro Multinacional de Investigación Educativa (CEMIE), proyecto de la OEA y del Ministerio de Educación Pública, con sede en San José, Costa Rica. Este Centro funciona bajo los auspicios del Programa Regional de Desarrollo Educativo de la Organización de los Estados Americanos.

PREFACIO

Este manual constituye básicamente una guía para quienes tienen la misión de informar a las autoridades de los ministerios de educación, administradores escolares, directores de escuelas y profesores sobre la medición de resultados de programas educativos.

Durante los últimos años las exigencias financieras de las instituciones educativas sobre los sistemas económicos nacionales se han incrementado. Dado que las necesidades de todos los servicios públicos en general, incluyendo educación, son cada vez mayores y sobrepasan el valor de los recursos disponibles, se hace urgente que los gobiernos inicien un proceso de revisión del empleo de los recursos actuales y futuros, preguntándose a la vez si se puede servir mejor a la población —y tal vez a menor costo— y en qué forma se podría lograr tal objetivo.

Los funcionarios encargados de la planificación y financiamiento de programas educativos han expresado su interés en medir los resultados de este sector, como un primer paso para resolver los interrogantes anteriores.

En este manual se presentan, en forma gráfica y esquemática, los siguientes temas para discusión en seminarios:

- I. Tipos de mediciones de resultados
- II. Utilización de las mediciones de resultados
- III. Diseño de métodos para medir resultados (rendimiento y actitudes)
- IV. Diseño de métodos para medir resultados (económicos y sociales)
- V. Aplicación de las mediciones de resultados

VI. Consecuencias de la información obtenida y pasos para su aplicación práctica.

Los esquemas de cada seminario, presentados a continuación, constituyen guías o planes de "lecciones preliminares". El moderador de cada grupo elegirá el material apropiado de trabajo. La información complementaria y el material ilustrativo que tengan especial importancia con respecto al sistema educativo de cada país (o región) podrían enriquecer la discusión.

Los gráficos presentados antes de cada tema de discusión tienen por objeto servir como una ayuda visual. El moderador de la discusión podría distribuir copias a los miembros de su grupo y tal vez completar estos datos con otros cuadros, o mostrar gráficos preparados con el mismo propósito, utilizando un proyector. Los gráficos aparecen de nuevo en el Apéndice B con el fin de que el lector tenga fácil acceso a ellos.

Los temas pueden seleccionarse de modo que proporcionen una orientación general acerca de los resultados educacionales en forma resumida, en un solo seminario, o desarrollarse con mayor profundidad en un ciclo de por lo menos seis seminarios. Se requiere de uno o más años de adiestramiento a fin de desarrollar las habilidades necesarias para diseñar instrumentos de medición o la capacidad analítica necesaria para aplicar tales medidas. Sin embargo, para quienes ya han adquirido esta capacidad analítica y cuantitativa, la participación en los seminarios aquí descritos les ayudará a comprender las implicaciones de este análisis de resultados en las políticas de educación. Los seminarios no constituyen un curso de capacitación técnica. Su propósito es más bien:

- Facilitar la comprensión de los diferentes tipos de resultados educacionales.
- Identificar las diversas aplicaciones de la medición de resultados.
- Ayudar a difundir tales conocimientos entre los funcionarios encargados de la política educativa, directores y profesores de centros docentes.
- Suministrar algunos conocimientos elementales sobre los instrumentos de análisis de las correlaciones de los resultados educacionales.

- **Destacar, lo más claramente posible, la importancia de estos resultados en la planificación de la educación y su financiamiento, así como en el trabajo docente.**

Los siguientes pasos son necesarios para lograr el desarrollo eficaz de cada seminario:

PASO 1

Elección de un moderador.

PASO 2

El moderador, junto con los funcionarios del sector educativo que patrocinen los seminarios, deberá definir los propósitos de las reuniones y resolver interrogantes tales como: ¿A quiénes están dirigidos los seminarios? ¿Cuál es el propósito del seminario?

PASO 3

Los seminarios deberán ser presentados indicando con claridad: a) su justificación al máximo y b) el uso práctico de los conocimientos que se adquieran, en el trabajo del funcionario, administrador, o profesor que asista al seminario.

PASO 4

El moderador podría seleccionar, del material presentado en este manual, lo que considere particularmente útil para las discusiones de cada grupo.

Como ayuda en este proceso de selección, en el Apéndice A se sugieren algunos pasajes del texto que podrían resultar de particular importancia para los siguientes grupos:

- funcionarios de alta jerarquía, encargados de las políticas generales de planificación educacional
- funcionarios de jerarquía intermedia de los ministerios de Educación y de Hacienda
- administradores de escuelas
- directores de centros educativos
- profesores
- instructores de instituciones de formación del magisterio.

PASO 5

Se seleccionará material complementario de lectura. Este material se encuentra en publicaciones del Public Services Laboratory de la Universidad de Georgetown, tales como **Medición de los resultados de la educación en los países en vías de desarrollo: Bibliografía anotada y Medición de los resultados de la educación en los países en vías de desarrollo.**

PASO 6

Al finalizar cada seminario deberá llevarse a cabo una evaluación preliminar, con el propósito de determinar si se ha cumplido con los objetivos previstos. Esta evaluación puede realizarse por medio de pruebas, cuestionarios o comentarios de los participantes.

Este tipo de evaluación no indica realmente el efecto de los seminarios. Solamente la práctica indicará si en realidad contribuyen, en forma eficaz, a cambiar los sistemas tradicionales de educación.

TEMA: LA APLICACION DE LA MEDICION DE RESULTADOS

Objetivos del Seminario V

Al finalizar el seminario, los participantes deberán:

- 1) Entender y valorar la utilidad que puede tener la estadística para el administrador de programas educativos, cuando se trata de analizar los resultados de esos programas.
- 2) Conocer el tipo de problemas que pueden resolver los estadísticos y las técnicas que pueden proporcionar.
- 3) Entender cuál sería la labor del administrador de programas, para fomentar el uso de estadísticos profesionales.
- 4) Estar capacitados para consultar técnicos estadísticos en el desarrollo de sus propios programas y objetivos.
- 5) Estar mejor preparados para estudiar con sentido crítico los informes basados en el análisis y la interpretación de datos estadísticos.
- 6) Comprender y valorar las ventajas que puede tener el incrementar sus propias habilidades técnicas en el campo de la metodología estadística.
- 7) Ser conscientes de las dificultades inherentes a la compilación, al análisis y a la interpretación de datos cuantitativos y comprender los fundamentos de las técnicas que se apliquen.

Preguntas introductorias

Se sugieren varias preguntas a fin de dar flexibilidad de acción al moderador de la discusión. El tiempo dedicado a esta discusión preliminar dependerá, en parte, del tiempo disponible para todo el seminario.

- (1) ¿Cómo se podría caracterizar, con unos pocos elementos de medición, la población a la cual se dirigen los programas de los participantes en el seminario (gente, objetos, acontecimientos)?; ¿los que efectivamente reciben servicios en virtud de los programas?; ¿los elementos del programa propiamente dichos?
- (2) ¿Cómo se podrían resumir las relaciones entre la actividad del programa (es decir, la prestación de servicios en sí) y los cambios que se producen en la práctica (o diferencias que se crean y mantienen)?
- (3) ¿Cómo se podrían aislar los resultados de un programa educativo con respecto a otros factores e influencias que podrían intervenir simultáneamente?
- (4) ¿Cómo se podría atribuir a un cambio observado (por ejemplo, la realización de un objetivo del programa) alguna de las siguientes causas: a) operaciones específicas de los programas, b) factores externos, c) variaciones o cambios a los cuales no se puede atribuir una causa o influencia específica?
- (5) ¿Cuáles son los problemas más agudos que se plantean al tratar de medir los resultados de los programas en que hayan intervenido los participantes, y cómo relacionar estos resultados con las actividades y características de los programas propiamente dichos?

NOTAS PARA EL MODERADOR DE LA DISCUSIÓN

El tema básico del seminario es el estudio de las enormes dificultades que se plantean en la labor de compilación, análisis e interpretación de datos. Estas dificultades inherentes e inevitables, a dicha labor son las que crean la necesidad de una metodología estadística formal. Además, justamente, la misión de los administradores de programas y la programación de la

producción en sus diversas formas (PPBS*, Administración por objetivos, etc.), requieren esa interpretación de datos.

Un tema conexo del seminario sería la posibilidad que tienen los administradores de recurrir a varias profesiones y combinar la ayuda de dos fuentes profesionales. La primera corresponde a la competencia profesional con respecto a lo esencial de un programa específico: por ejemplo, la educación. No es posible llevar a cabo un análisis de programas sin basarse en los fundamentos teóricos y prácticos del tema esencial. La segunda es la técnica estadística: la metodología tradicional de base matemática que se aplica en los problemas de evaluación, tales como encuestas sociales, pruebas y mediciones.

De estos dos temas se podría derivar un tercero y último: no es una tarea simple para el administrador de un programa público obtener estos recursos técnicos y aprovecharlos. *Esta necesidad es justamente la idea central de este seminario: preparar mejor al administrador para dirigir sus labores, contando con la bien orientada asistencia de profesionales.*

A continuación se exponen siete puntos, cada uno de los cuales abarca el problema genérico que corrientemente afronta el administrador de programas y presenta el origen de la metodología estadística a la cual se puede recurrir al tratar de alcanzar una solución. El seminario puede complementarse con ejemplos proporcionados por la experiencia de los participantes. La aplicación específica a los programas en que han trabajado los participantes del seminario representa el principal elemento de aprendizaje y permitiría subrayar los conceptos básicos.

* N. del T. Planning-Programming-Budgeting System: Sistema de Planificación, Programación y Presupuestos.

DESARROLLO Y APLICACION DE LA MEDICION DE RESULTADOS EN LOS PROGRAMAS PUBLICOS

<i>Problema Genérico</i>	<i>Métodos de Análisis</i>	<i>Aplicaciones Típicas</i>
Caracterización y comparación de grupos de personas, objetos, acontecimientos.	Totales, subtotales, medias, distribuciones, pruebas de diferencias entre medias, etc.	Determinación de variaciones en el carácter o los resultados de los programas; diferencias en el tiempo, entre áreas geográficas, entre tecnologías de programación.
Medición de relaciones.	Tablas de contingencia, diagramas de dispersión, análisis de correlación y regresión.	Determinación de factores que influyen en los resultados de programas de manera favorable o desfavorable; elección de elementos programáticos para lograr mejores resultados.
Aislamiento de influencias; descomposición de tendencias o medición de resultados.	Análisis de tendencias; análisis de regresión; "control" mediante diseño experimental de interferencias que se desea evitar.	Atribución de efectos y consecuencias de un programa en presencia de factores que intervienen o interfieren.
Medición de la variación y descripción de sus componentes.	Análisis de la varianza; medición de componentes atribuibles a factores de un programa, a factores ajenos al programa (ambientales), a interacciones entre factores, y a "errores experimentales".	Medición de la importancia relativa de los factores del programa; factores ambientales y sus interacciones — todos estos componentes interpretados según medidas explícitas de la sensibilidad del sistema de medición.
Compilación de datos para determinar la especificación precisa de un modelo analítico y un plan de análisis.	Métodos de encuesta, selección de instrumentos adaptación de procedimientos administrativos y registros, etc.	Acoplo de datos sobre las poblaciones seleccionadas, manejo de los programas, costos, resultados de proyectos (inmediatos) y consecuencias de los programas (posiblemente a largo plazo).
Obtención de la información requerida de manera económica, manteniendo un alto nivel de calidad y aplicando medidas de precisión.	Muestreo y diseño experimental mediante control de calidad de las operaciones en el terreno.	Reducción de costos; reducción de las reacciones negativas originadas en las respuestas y de la interferencia con las operaciones; mantenimiento de un alto nivel de calidad.
Medición de los resultados y consecuencias del programa, asignándoles valores, eligiendo elementos y tecnología para el programa, con el fin de lograr el mayor nivel de producción por unidad de insumos de recursos.	Todas las técnicas anteriores, más esquemas de evaluación que permitan convertir los resultados y consecuencias en beneficios sociales, incluyendo experimentos de variación planificada.	Niveles de medida de resultados; variaciones en el tiempo; relación entre resultados y factores del programa (costo-eficacia) y comparación de diferentes programas según los resultados (análisis de beneficio y costo-beneficio.) Numerosas aplicaciones en planificación, análisis de política presupuestaria, etc.

Ideas para la presentación

Esta sección se organiza en torno a siete problemas genéricos de compilación, análisis e interpretación de datos para aplicarlos en la administración de programas. El orden de la secuencia es intencional. En el proceso de presentación de la información y aplicación de datos estructurados, el procedimiento es el siguiente: 1) obtener información y reunir datos (formularios de inscripción, encuestas, archivos administrativos, pruebas —escritas u orales— o diversos formularios, registros de información, etc.); 2) preparar los datos según el tipo de análisis que deba efectuarse; 3) llevar a cabo el análisis propiamente dicho; 4) interpretar los resultados del análisis —descritos más adelante— y aplicar las conclusiones a los problemas administrativos, a los planes y a la acción del organismo y programa en cuestión.

Es importante observar que el diseño de proyectos sigue el camino inverso. En primer lugar, se determinan necesidades específicas con base en un orden de prioridades establecido por el administrador y sus superiores. Estas necesidades determinan, a su vez, el contenido y la forma del análisis estadístico y cómo deben interpretarse los resultados a fin de poder aplicarlos. Las mediciones, las relaciones y la sensibilidad indispensable para estos análisis determinan *en detalle* la forma, los alcances y el contenido del procedimiento de compilación de datos, es decir, qué datos deben reunirse, de qué manera, si mediante muestras (y qué tamaño y contenido deben tener esas muestras), etc.

También es posible proceder de otro modo. La información se obtiene sin cumplir primero con las etapas que podrían haberse programado en detalle por anticipado. A menudo se crea de este modo un "banco de datos"; no se llevan adelante, en forma explícita, los planes para analizar, interpretar y aplicar la información hasta haber establecido el "banco de datos". Decimos "en forma explícita" porque surgen muchas ideas sobre aplicación y análisis de datos al determinar la información que se va a compilar y al decidir en qué forma (incluyendo sistemas de clasificación de subtotales, etc.) y de qué fuentes se va a compilar. Sin embargo, la experiencia ha demostrado que las

necesidades del administrador de programas son demasiado precisas y compulsivas y no permiten comenzar con la información a base de conjeturas, con la esperanza de poder establecer las relaciones necesarias derivándolas de algún modo del "banco de datos". Tal vez la pobreza del análisis y evaluación de los programas de educación pública se deba, con frecuencia, a esta falla en la estrategia más que a cualquier otro factor ajeno a las decisiones de política.

Problema N° 1: Datos que caracterizan un grupo de personas, objeto o acontecimientos y que permitan efectuar comparaciones válidas entre dichos grupos.

A menudo se requiere que los datos proporcionen una descripción resumida de un grupo; por ejemplo, niños en edad escolar de una zona específica. Con frecuencia esta descripción consiste simplemente en el número de individuos registrados o estimado en el grupo: el total. Es probable que en algunos casos esta descripción resumida sea suficiente para ciertos propósitos, y que no se necesite mayor información. Sin embargo, más frecuentemente surgen otras interrogantes de importancia: ¿Cuántos niños de edad escolar asisten a la escuela? ¿Qué nivel de enseñanza han alcanzado aquellos que no asisten? Las respuestas tendrán sentido solamente si se relacionan con las edades específicas de los niños, y, en general, es conveniente contar con respuestas separadas para niños y niñas.

La posibilidad de dividir un grupo en subtotaes implica un sistema de *clasificación* bien definido, con *definiciones* precisas de las diferentes clases.

Más adelante veremos también que muchas de estas respuestas sólo pueden interpretarse provechosamente en la medida que sea posible *relacionar* las características de los subgrupos entre sí; los subtotaes por edad y nivel escolar alcanzado dejan muchas interrogantes de importancia sin respuesta. Es necesario establecer una clasificación cruzada por edad y por clase de modo que se pueda *distribuir* cada nivel escolar alcanzado por edad (y probablemente por sexo), y cada edad por asistencia escolar (grado actual) y por nivel escolar alcanzado (en el caso de los que no asisten a la escuela).

En resumen, podemos caracterizar un grupo presentando el gran total de individuos y clasificándolos para obtener subtotaes, lo cual significa que podemos *distribuirlos* en subgrupos mediante un sistema de clasificación. También podemos relacionar cualquier subtotal con otro para una variedad de distribuciones según los atributos de los miembros del grupo.

También es posible distribuir un grupo según una *variable* o escala de valores, tales como la estatura de los niños o los puntajes de las pruebas en un examen escolar de lectura. Por razones tanto prácticas como teóricas, un buen resumen de un grupo de puntajes de pruebas o estaturas es el *promedio* o media aritmética. Por lo tanto, es posible recurrir a esta sola estadística resumida para representar algún aspecto importante del grupo. Sin embargo, al igual que cuando se usa un total, el solo dato numérico no incluye una buena parte de la información.

Un primer paso para lograr mayor precisión consiste en presentar los promedios para cada uno de los subgrupos que se consideren útiles en el análisis de componentes de subtotales: puntajes de pruebas (o estatura) separados, para niños y niñas, y para cada edad, nivel escolar actual, en el caso de los que asisten a la escuela, y para cada nivel escolar alcanzado en el caso de los que no asisten a la escuela.

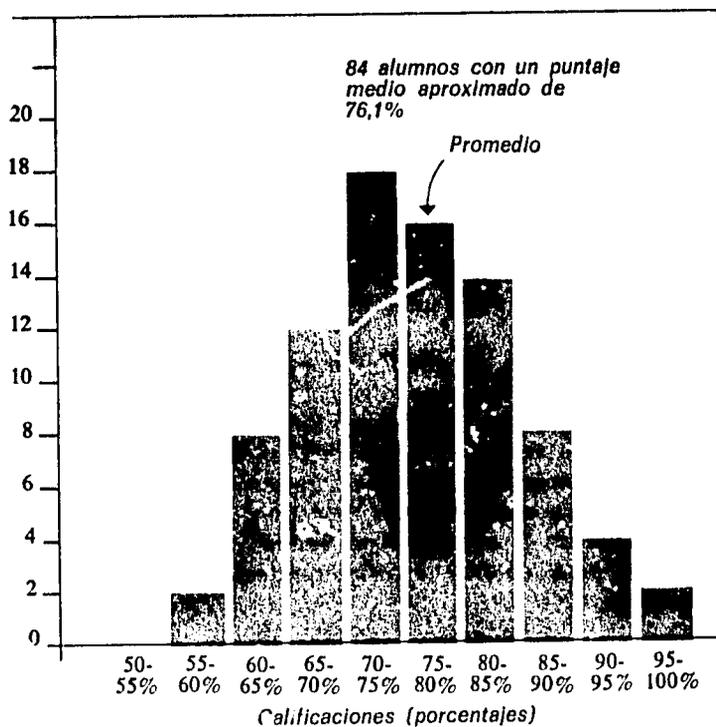
También es posible desarrollar un análisis de distribución para un solo subgrupo de individuos. Por ejemplo, las niñas de 10 años, en cuarto grado escolar, tienen un puntaje promedio en una prueba específica, pero no todas ellas obtuvieron ese puntaje (en realidad, es posible que ninguna de ellas haya obtenido ese promedio). Podemos distribuir los puntajes de cada niña en este subgrupo simplemente contando el número de niñas que obtuvo cada calificación y mostrando el resultado gráficamente (Gráfico V-1). También es posible indicar la proporción de puntajes que se sitúa entre ciertas distancias con respecto al promedio, por ejemplo, dentro de un límite de cinco puntos del promedio, diez puntos, etc. Esto proporciona un nuevo orden de medición o estadísticas que se agregan al promedio como descripción estadística del grupo. La estadística técnica depende mucho de esas medidas de dispersión con respecto a la media de la clase, ya que la teoría matemática de probabilidad proporciona mecanismos sumamente útiles para interpretarlas. (El estadístico utilizará la medida de la varianza, basada en las diferencias al cuadrado de cada puntaje con respecto al promedio.)

Ahora contamos con algunos mecanismos esenciales para la comparación analítica de dos o más grupos. Vol-

GRAFICO V-1

PUNTAJES DE PRUEBAS APLICADAS A NIÑAS DE DIEZ AÑOS DE EDAD EN CUARTO GRADO

Número de alumnas



viendo al caso de las niñas de 10 años de edad, ahora es posible comparar el grupo de niñas que tienen 10 años actualmente con un grupo de niñas que tenían 10 años en alguna fecha específica del pasado, con el propósito de determinar si ha habido progreso en la asistencia a clases y en los niveles de educación escolar alcanzados. También es posible comparar una comunidad con otra. Si se aplican pruebas tipificadas de rendimiento escolar, es posible efectuar las mismas comparaciones para grupos de 10 años de edad en diferentes años o comunidades.

El uso de totales, subtotales (basados en un sistema de clasificación bien definido), distribuciones de atributos y distribuciones de variables puede a menudo hacer posibles todas las comparaciones útiles que son esenciales.

Con frecuencia el administrador de un programa educativo deberá afrontar un problema más: el análisis de las diferencias que desearía atribuir a su propio programa. Por ejemplo, la utilización de un nuevo texto de estudios debería mejorar los puntajes en las pruebas de los niños. Es necesario, pues, comparar los beneficios con ese nuevo texto en comunidades y clases anteriores que no contaron con él. Por razones de sensibilidad, a menudo es necesario distribuir los niños en subgrupos de diferentes maneras (edad, sexo, etc.), a fin de captar mejor el impacto del posible progreso.

Más adelante volveremos a este problema. En todo caso, los mecanismos estadísticos relativamente simples que hemos esbozado constituyen un buen punto de partida. Veremos que la teoría de la distribución desempeña un papel importante y sutil en la interpretación de datos experimentales, de encuesta y de muestreo. En general, habrá que confiar en la capacidad profesional de un técnico matemático para obtener una interpretación válida de los análisis.

Problema N° 2: Medir las relaciones entre dos factores o características, en particular los resultados posibles de un programa y compararlos con alguna medida de un producto del programa.

Comenzaremos con un ejemplo elemental (que requerirá mayor precisión en la mayoría de los casos reales, como se verá al tratar el problema N° 3). Supongamos que se establece, en forma experimental, un nuevo programa en virtud del cual se distribuyen textos gratuitamente a los alumnos de tercer grado, sólo en determinadas escuelas. Se desea evaluar la influencia de estos textos en la promoción de alumnos al grado siguiente. Consideraremos todos los alumnos del tercer grado a quienes se han entregado textos gratuitamente y una cantidad semejante de alumnos de tercer grado de otras escuelas. Nuestra medida de eficacia del programa será el porcentaje de alumnos de este año que se matriculan en el cuarto grado el año próximo.

Los resultados pueden resumirse (datos hipotéticos) en una simple tabla de contingencia:

Escuelas

	<i>Con textos gratuitos</i>	<i>Sin textos gratuitos</i>
Porcentaje de alumnos de tercer grado matriculados en el cuarto grado del año-calendario siguiente	80	61
Porcentaje de alumnos no matriculados en el cuarto grado	20	39

Esta tabla parece demostrar que un 19 por ciento más de alumnos de tercer grado continuaron sus estudios en el grado siguiente como consecuencia de la distribución gratuita de textos. (Podrían surgir interrogantes sobre la existencia o inexistencia de otras diferencias entre las escuelas en que se aplicó el programa y las otras escuelas

cuyos datos se utilizaron en la comparación. Esta posibilidad será analizada en el problema N° 3. Podríamos preguntarnos también cómo se eligen las escuelas experimentales: ¿Por qué esas escuelas en particular? Esta es una cuestión técnica que se tratará en el problema N° 6.)

Tomemos un ejemplo de la medicina: se nos dice que todas las ovejas que recibieron la vacuna de Pasteur contra la bacera sobrevivieron con buena salud al mezclarse con ovejas enfermas, mientras que las que no fueron vacunadas murieron.

El resultado:

	<i>Vacunadas</i>	<i>No vacunadas</i>
Ovejas que sobrevivieron	100%	0%
Ovejas que murieron	0%	100%

Nadie podría contradecir este resultado. Sin embargo, la medicina moderna (y la educación) a menudo debe tratar de resolver problemas complejos cuya solución no es tan fácil. Supongamos que Pasteur hubiera atacado las enfermedades más difíciles, con los siguientes resultados:

	<i>Con tratamiento</i>	<i>Sin tratamiento</i>
Ovejas que sobrevivieron	41%	37%
Ovejas que murieron	59%	63%

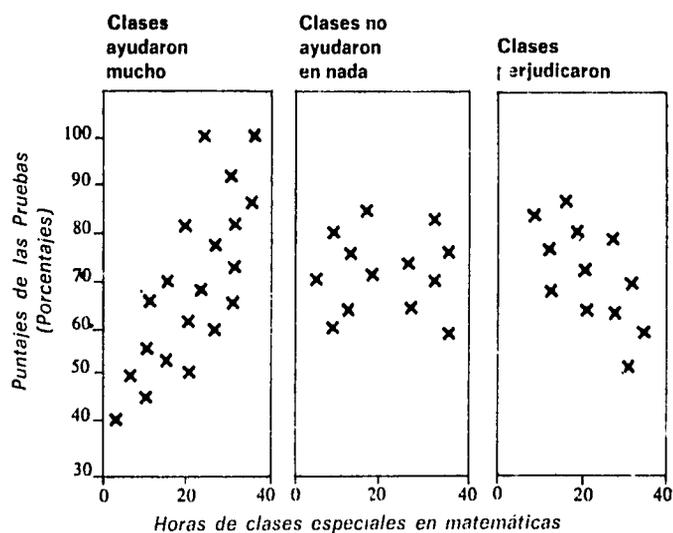
¿Qué posibilidad hay de que surja esta diferencia del cuatro por ciento a favor del tratamiento en casos en que el mismo no ha tenido realmente ningún efecto? Esta pregunta puede contestarla el técnico estadístico, utilizando la teoría matemática de la probabilidad. Probablemente resultará que la diferencia es una diferencia real, aunque el tratamiento no es muy eficaz, sólo si se toman muestras bastante grandes —y aleatorias— de casos con la enfermedad, con tratamiento y sin tratamiento. Obsérvese la misma lógica básica del problema N° 1, en el cual se requeriría medir la diferencia entre dos grupos de individuos, utilizando medias y distribuciones.

Si pasamos de los atributos (que puedan estar presentes o ausentes, como las vacunas y supervivencias) a las

GRAFICO V-2

RELACION ENTRE PUNTAJES EN MATEMATICAS Y HORAS DE CLASES ESPECIALES

(Ejemplos de diagramas de dispersión)



(Datos hipotéticos)

variables (que pueden incluir una serie de valores), es posible recurrir a otras técnicas más eficaces.

Supongamos que tenemos un programa especial de enseñanza de aritmética, en un grado escolar determinado, con un grupo de alumnos que reciben clases especiales de tutores individuales durante cantidades variables de horas de clase, y otro grupo de alumnos que no recibe clases especiales (cero horas de clase). Supongamos que luego se aplica una prueba para registrar los puntajes de modo que cada uno de éstos se puedan asociar al número de horas de clases especiales recibidas. Esto permite establecer

un *diagrama de dispersión*, en el cual cada punto corresponde a un alumno cuyo puntaje en la prueba se mide en la escala de la izquierda (distancia vertical), mientras que el número de horas de clases especiales recibidas por el alumno se mide en la escala inferior (distancia horizontal). En el Gráfico V-2, el diagrama de la izquierda correspondería a la situación en que la enseñanza especial ha ayudado mucho a los alumnos; en el diagrama del centro dicha ayuda ha sido escasa o nula, y el diagrama de la derecha muestra la situación en que la enseñanza especial, en realidad, fue un estorbo a la capacidad del alumno para obtener un buen puntaje en la prueba. (Todos los datos son hipotéticos.)

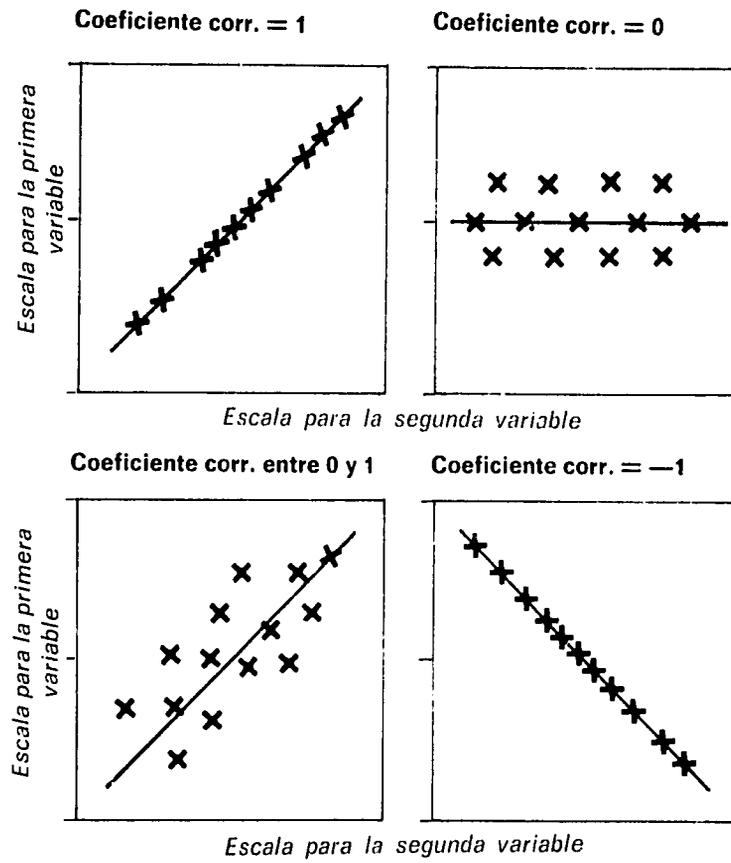
¿Cómo podrían resumirse estas relaciones mediante estadísticas adecuadas? Hay dos métodos estrechamente ligados entre sí para analizar relaciones de este tipo. El primer método es el llamado *análisis de correlación*, que mide la correlación entre las dos variables. En el Gráfico V-3 el primer diagrama muestra una correlación que podría calcularse como *coeficiente de correlación* igual a 1; en el segundo diagrama ese coeficiente sería igual a cero; en el tercero se sitúa entre cero y uno, y en el último sería igual a menos uno. (El coeficiente se calcula como la razón entre los cuadrados de las diferencias con respecto a la línea trazada en cada diagrama y los cuadrados de las diferencias con respecto a la media aritmética; la línea representa el ajuste de "mínimos cuadrados" porque cualquiera otra línea trazada en el mismo diagrama de dispersión tendría un total de cuadrados de diferencias mayor que la línea que se ha trazado.)

El segundo método de análisis es el llamado *análisis de regresión*, en el cual se utiliza la misma línea de "mínimos cuadrados" —o línea de "regresión"— para estimar el resultado que habrían arrojado las pruebas si el grupo hubiese recibido determinado número de horas de clases especiales. Obsérvese que un alumno obtendrá este rendimiento estimado sólo si existe un coeficiente de correlación exactamente igual a uno. En otros casos, es posible que no haya ningún alumno con el número especificado de horas de clases especiales que alcance el puntaje estimado como "el más probable" para ese número de horas de clase.

GRAFICO V-3

CORRELACION ENTRE DOS VARIABLES

(Muestran el coeficiente de correlación)



(Datos hipotéticos)

Es posible estudiar más de dos variables simultáneamente mediante estas técnicas. El "coeficiente de correlación múltiple" podría expresar la influencia del número de horas de clases especiales, el número de alumnos por clase regular y los años de experiencia del profesor de aritmética. El análisis de regresión permitiría una estimación del puntaje de pruebas para cualquier combinación de horas de clases especiales, número de alumnos por clase y experiencia del profesor.

La aplicación de este método a los análisis de programas es perfectamente directa. Hay que elegir una población estudiantil adecuada (problema Nº 6), establecer una escala que pueda aplicarse como medida de éxito (problema Nº 5) y organizar la variación de los factores que deben estudiarse (probablemente Nº 6 nuevamente) de modo que la influencia de cada uno pueda medirse con respecto a la medida de éxito seleccionada (resultado). Hay una situación especial en la cual dos factores son importantes, pero sólo si ambos están presentes. El análisis de regresión no puede medir esta condición necesaria pero insuficiente de una sola variable; este tipo de "interacción" requiere una técnica estadística diferente (problema Nº 4).

Problema N° 3: Separar o aislar la influencia de un programa con respecto a otras influencias que intervienen y que podrían producir resultados similares.

Esta es una situación frecuente en la experiencia real de la evaluación de programas mediante el análisis de resultados. En realidad, el mundo no se detiene cuando las autoridades deciden tratar de aminorar o remediar un problema social. El mundo cambiante continúa cambiando y las tendencias a menudo se mantienen, aunque sean modificadas por un nuevo programa.

Entonces, ¿cómo aislar el efecto del programa con respecto a otros cambios?

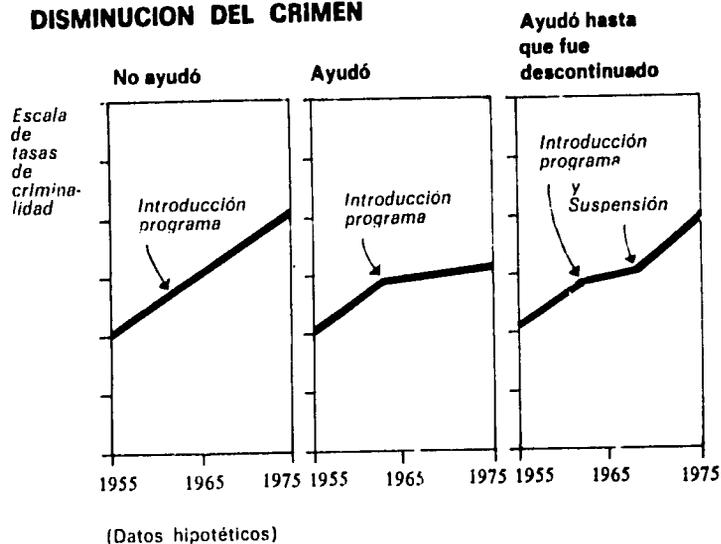
Un enfoque de este problema es el *análisis de tendencias*, un estudio de la forma de los cambios en el tiempo. Con frecuencia se leen en publicaciones no técnicas comentarios como: "habiendo invertido millones en la prevención del crimen, la tasa de criminalidad continúa aumentando; es pues, evidente que el programa es un fracaso y que el dinero de los contribuyentes ha sido malgastado". Este juicio se justificaría tal vez —en virtud de un análisis elemental e indudablemente incompleto— si se tratara de la situación indicada en el diagrama de la izquierda del Gráfico V-4. Sin embargo, en el segundo diagrama se podría discutir la afirmación de que el dinero ha sido "malgastado", a pesar de que la tasa de criminalidad ha seguido aumentando. Lo que el gráfico indica es que existen influencias tan fuertes que favorecen la tendencia al aumento del crimen (valdría la pena estudiar las características demográficas cambiantes, las condiciones económicas y las reacciones sociales, así como la seguridad pública, la corrección y rehabilitación de criminales y los programas conexos) y que dicho aumento continúa a pesar de una baja significativa de la *tasa de aumento* de la criminalidad, que coincide con la aplicación del nuevo programa. Tal vez, en un análisis más completo, el tercer diagrama sería aún más convincente. Lo que indica este gráfico es que el programa hizo bajar la tasa de aumento pero sin un efecto duradero;

GRAFICO V-4

EFFECTOS DE LOS PROGRAMAS DE PREVENCIÓN DEL CRIMEN

(Ilustración del análisis de tendencias)

INTRODUCCION DE UN PROGRAMA PARA DISMINUCION DEL CRIMEN



la supresión del programa coincidió con una reaparición de la tendencia anterior, junto con una mayor proporción de incremento en la tasa de criminalidad.

La educación también ha aportado conclusiones vagas y discutibles, disfrazadas como resultados de análisis serios: "Habiendo invertido todo el dinero en el Título I (de la ley de 1965 sobre educación primaria y secundaria en los Estados Unidos), los escolares aún no saben leer". En este caso es necesario: 1) definir y *cuantificar* con mayor

precisión el concepto "niños que no saben leer"; 2) medir las variaciones de esta medida de la capacidad para leer, y 3) aislar los efectos del programa del Título I (y sus numerosos y variados elementos) de otras influencias coincidentes.

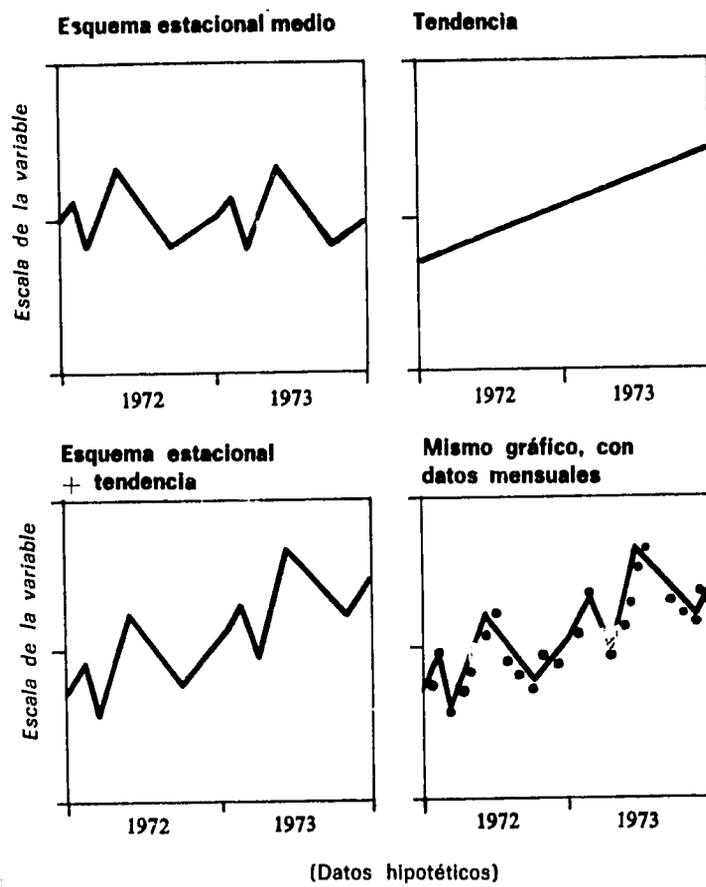
El análisis de tendencias depende, pues, de una coincidencia en el tiempo para poder atribuir "causas" a los factores de programa y otras influencias. Un enfoque más avanzado implica la separación de los componentes de una serie cronológica. Por ejemplo, el primer diagrama del Gráfico V-5 presenta un esquema normal de cambio a través del año promedio, un componente "estacional"; el segundo diagrama muestra una tendencia (posiblemente calculada para un período de varios años); el tercer diagrama muestra la combinación de dos componentes; el cuarto muestra una verdadera serie cronológica (datos hipotéticos), y el último indica la *variación residual* o variación inexplicada, habiendo tomado en cuenta una tendencia media y un esquema estacional medio. Esto debe atribuirse a lo siguiente: (1) al hecho de que la tendencia *media* y el esquema estacional *medio* no describen lo que sucede en un año determinado, es decir, la *distribución* de datos en torno a los promedios (véase problema N° 1), y (2) todas las demás causas e influencias fuera de la tendencia y los esquemas estacionales. El estudio de los "ciclos comerciales" en economía ha hecho uso depurado de estas técnicas, para demostrar finalmente que no se puede confiar en la repetición de hechos a través del tiempo y, por lo tanto, no es posible utilizar, en forma satisfactoria, esta estrategia en la economía y la previsión económica.

Se puede recurrir a otro tipo de estrategia para aislar los factores que deseamos no intervengan en el estudio de resultados de programas y para evaluar programas públicos. Los datos pueden compilarse de tal manera (véase problema N° 5) que es posible aislar un factor de otro. Esto podría lograrse con la preparación de una tabla de contingencia (problema N° 2), dejando que los requerimientos de la tabla determinen las especificaciones para la compilación de datos (véase nuevamente el problema N° 5). Supongamos que en una comunidad determinada algunas escuelas obtienen típicamente puntajes altos en las pruebas de ren-

GRAFICO V-5

VARIACIONES A TRAVES DEL TIEMPO

(Ilustración de los componentes de una serie cronológica)



dimiento escolar, mientras que en otras son bajos. Supongamos que en la escuela que obtiene puntajes altos la mayoría de los estudiantes provienen de familias de ingresos relativamente altos, cuyos padres han recibido una buena educación, mientras que los de las escuelas que obtienen puntajes bajos provienen, en su mayoría, de familias de bajos ingresos, cuyos padres han alcanzado un nivel inferior de educación.

¿Es el medio familiar o la escuela la causa de la diferencia en el rendimiento escolar de los niños? Si seleccionamos un grupo de alumnos de cada clase de las escuelas (aún si esta selección se efectúa de manera justa e imparcial, como se comentará en el problema Nº 6), los resultados, probablemente, se presentarían como el cuadro de contingencia del primer diagrama del Gráfico V-6. Toda la interpretación se basaría en unos pocos niños que provienen de familias de bajos ingresos pero que asisten a escuelas de alto puntaje y unos pocos que provienen de familias con mejor situación económica pero asisten a escuelas más pobres. El resto de los casos aportaría muy poca información útil, puesto que en ellos los factores que desean aislar han sido "confundidos" en el diseño de la compilación de datos. Existe una técnica que permite aumentar el tamaño de la muestra en las casillas en que se diferencian los dos tipos de influencia y reducir la muestra en las casillas en que esas influencias se confunden. El segundo diagrama muestra el resultado con respecto a la estructura de las muestras. El último diagrama presenta los puntajes promedio obtenidos en las pruebas, en cada casilla, siendo más precisas las dos casillas que diferencian los antecedentes familiares y la "calidad de las escuelas". En este ejemplo hipotético, el análisis indicaría que el medio familiar tuvo mayor importancia en la diferencia en el rendimiento de los niños que las características de las escuelas. Esto podría tener gran valor para la formulación de políticas y presupuestos y también para la administración racional de programas educativos.

CUADRO V-6

PUNTAJES DE PRUEBAS ESCOLARES EN DOS COMUNIDADES

(Muestra la combinación de influencias explicativas y la sensibilidad del análisis estadístico)

Escuelas de comunidades en que el nivel de ingresos familiares es:

Niños cuyos padres han alcanzado un nivel educacional:

	Alto	Bajo
Alto	214: 83%	2: *
Bajo	4: *	307: 72%
Alto	50: 83%	200: 81%
Bajo	200: 75%	50: 72%

Clave: Cada casilla indica el número de niños sometidos a prueba y el puntaje medio obtenido.

(Datos hipotéticos)

* *Muestra demasiado pequeña; promedio no es significativo.*

Hemos evitado dos problemas en el análisis anterior. En primer lugar: ¿cómo podemos conocer el medio familiar antes de recoger los datos? Este problema puede resolverse diseñando la compilación de datos en dos etapas. En la primera se determinan los antecedentes familiares de un número de niños suficientemente elevado para que la muestra final en la segunda etapa pueda contener, en la mayoría de los casos, a los niños que sólo representan una pequeña fracción del total de alumnos. En la segunda etapa combinamos dos variables en nuestra descripción del "medio familiar": ingresos y nivel de educación de los padres. Nuevamente, este tipo de confusión puede resolverse mediante una etapa preliminar de compilación de datos. Si bien la mayoría de los padres de las familias de altos ingresos seguramente ha alcanzado un nivel de educación más alto que el de los padres de las familias de bajos ingresos, es necesario que haya numerosas combinaciones de altos ingresos y bajo nivel de educación, y de bajos ingresos y alto nivel de educación para representar la muestra en tablas de contingencia que diferencien estas influencias.

Hay otros métodos para aislar las influencias que podrían confundir o obstaculizar la medición de resultados de programas. Estos otros métodos implican un "control" de factores que podrían interferir, mediante el diseño específico del análisis y compilación de datos, de modo que esto permita aislar factores por medio del *análisis estadístico*, aún si no fuere posible encontrar *individuos* que ejemplifiquen los casos diferenciados como en los cuadros de contingencia indicados anteriormente. Estudiaremos este asunto en el problema N° 4.

Cabe observar que este lógico enfoque del análisis estadístico se basa en una sólida teoría de estadística matemática, lo cual puede salvar al analista de la decepción que producen los datos en muchos sentidos. Lo que hemos presentado aquí constituye una lógica básica de los métodos. En la práctica, el análisis puede presentar numerosas trampas a quien carezca de preparación y prudencia (recuérdese el problema N° 1). El propósito del seminario es preparar a los participantes de modo que sean conscientes de las estrategias que el estadístico cali-

ficado puede utilizar en la solución de los problemas del administrador de programas y, más importante, que sepan cómo instruir al estadístico con respecto a las necesidades del análisis, a fin de cumplir el objetivo del análisis de programas, presupuestos y operaciones conexas.

Problema N° 4: Medir la variación total en alguna escala de valores (por ejemplo, una medición de resultados) y asociar los componentes de esa variación con las influencias del programa y otros factores relacionados. La estadística es el estudio sistemático de la variabilidad. Si no hay variabilidad, no hay objeto de estudio.

Consideremos el problema del alumno que progresa poco en todas las situaciones escolares. Si se diera en la realidad esta increíble condición, el análisis mediante métodos cuantitativos sería imposible. Supongamos que todos los niños reciben las mismas calificaciones, los mismos puntajes en las pruebas y obtienen el mismo resultado en todas las demás evaluaciones, en todos los casos. Puesto que no hay variación, no es posible relacionar ninguna influencia con el éxito escolar.

En la realidad, un niño difiere de otro en muchos aspectos. Nuestro problema consiste en establecer una estructura metódica de estas variaciones, de tal modo que podamos comprender, tanto como sea posible, los factores que intervienen en las variaciones.

Supongamos que se ha aplicado una prueba de lectura a todos los alumnos de tercer grado en tres escuelas. Supongamos, además, que no contamos con ninguna información específica sobre las escuelas, aunque éstas podrían diferir en diversos aspectos, ya que cada una pertenece a una comunidad diferente. Los niños también difieren en muchos aspectos, pero todo lo que conocemos de ellos —que permitiría asociar estos factores a los puntajes de cada niño— es el *sexo, la edad y el nivel de educación alcanzado por el padre* (o la madre, si el padre no vive en el hogar).

Trazamos un marco para el análisis —un “*modelo*”— que determine las medidas y relaciones que permitirán la “*explicación*” —la *asociación*— de las diferencias de puntaje con respecto a los diversos factores antes mencionados:

- (variaciones asociadas con diferencias entre las escuelas (aunque no exista ningún registro sobre estas posibles diferencias);

- variaciones (asociadas con el sexo de los niños de tercer grado;
- variaciones asociadas con la edad de los alumnos de tercer grado;
- variaciones asociadas con las diferencias en el nivel de educación alcanzado por los padres.

Cabe observar que probablemente sólo el primer factor se consideraría un factor "de programa", por el hecho de que puede ser controlado mediante la modificación de presupuestos, la preparación o adiestramiento de profesores, la distribución de textos de estudio gratuitamente o la entrega de nuevos textos, etc. La evaluación de las diferencias entre escuelas constituye sólo el primer paso en el proceso de averiguar qué es lo que causa las diferencias y utilizar la información para mejorar *todas* las escuelas en la medida en que lo permitan la tecnología (libros, material didáctico, métodos de enseñanza, etc.) y los recursos.

La asociación de las diferencias entre puntajes de pruebas con las diferencias en el sexo y edad entre alumnos y nivel de educación alcanzado por los padres se miden, no porque el administrador de programas tenga la posibilidad de controlar estos elementos, sino porque es esencial evitar la confusión de las diversas influencias. También resulta útil medir la proporción de la variación total que está determinada por el sistema educativo y no por el medio en que el sistema funciona.

La lista anterior de fuentes de variación no está completa. Faltan medidas de *interacción*, y la *variación residual* (medida como "varianza" residual) que puede ser útil considerar como "error experimental", porque indica los límites de la posibilidad de explicar los hechos mediante el modelo.

La interacción difiere de otras medidas tales como la diferencia entre escuelas. En el caso de esta última, el análisis mide el *efecto aislado* de uno de los factores en el modelo (particularmente importante en este ejemplo). La interacción mide el efecto de dos o más factores que actúan juntos. En un modelo supersimplificado podríamos imaginar la existencia de aulas sin profesor o profesores sin salas de clase, pizarra o material didáctico. El estudio de los puntajes de las pruebas en estas condiciones segu-

ramente indicará que se obtendría escasa ventaja del hecho de contar con material didáctico pero no con profesor, y un éxito muy limitado en el caso de que se cuente con profesores pero éstos no dispongan de material didáctico. Sin embargo, la *combinación* —interacción de profesores, aulas, pizarras y textos— se asociaría, probablemente, con los puntajes más altos.

Otra ilustración del significado de la interacción puede obtenerse del siguiente ejemplo:

Promedio de puntajes de pruebas de aritmética

	<i>Grupo étnico A</i>	<i>Grupo étnico B</i>	<i>Ambos grupos</i>
Niños	40	60	50
Niñas	60	40	50
Ambos sexos	50	50	50

La media aritmética para uno u otro sexo y ambos grupos étnicos es 50 por ciento. Cada sexo y cada grupo étnico tiene a su vez un puntaje medio de 50 por ciento. Pero el esquema de puntajes indica que estas estadísticas no logran demostrar la presencia de otro fenómeno. Algo ocurre con el niño en el grupo A, o niña en el grupo B (en este ejemplo hipotético) que produce puntajes mucho más altos —como término medio— que las dos combinaciones inversas. Estas estadísticas sólo describen la *asociación* pero no las *causas* de las diferencias. Sin embargo, no hay razón para buscar causas mientras no esté claro que debe haber al menos una. (Una conjetura que podría someterse a prueba es que los diferentes grupos étnicos asisten a diferentes escuelas y que una escuela tiene un profesor que atrae especialmente a los niños y la otra tiene un profesor que atrae en particular a las niñas; los factores de actitud y motivación de esta situación se expresan en el patrón de puntajes de prueba.)

Nuestro método consistirá en medir la *variación total* y luego restar la *variación asociada* con cada uno de los factores aislados, las interacciones entre cada combinación de dos factores (interacciones "de primer orden") y las interacciones entre combinaciones de más de dos factores

simultáneos ("orden más elevado"). Después de sustraer estos componentes, siempre quedará algo de la variación original, no asociada con ninguno de los incrementos anteriores. Esto se llama "variación residual" y representa todas las influencias de factores no incluidos en el modelo, así como todas las interacciones no incluidas. Estas últimas deben incluir todas las diferencias de capacidad de los niños, sus habilidades para aprender, actitud, motivaciones, asistencia a clases, influencias del hogar no relacionadas con el nivel de educación de los padres, condiciones físicas del niño, etc. Se incluyen también todas las variaciones asociadas con la prueba misma y su aplicación.

La variación residual cumple dos finalidades fundamentales: (1) indica la importancia relativa de los factores cuya influencia ha sido medida, así como todas las demás influencias en los puntajes de pruebas, y (2) proporciona una base para medir la sensibilidad de cualquier diferencia que efectivamente se mida. Para esta última finalidad, podríamos dudar de la validez de un 2 por ciento de la variación asociada con las diferencias entre escuelas si el 65 por ciento de la variación total no pudiese asociarse con ningún factor o interacción en el experimento. En cambio, una diferencia del 15 por ciento entre escuelas, con una variación residual de sólo 10 por ciento, parecería importante. (El técnico estadístico volvería a la teoría matemática de probabilidad para lograr una interpretación más exacta de los datos de este tipo y los tamaños de las muestras influirían en la interpretación.)

¿Cómo debemos medir la variabilidad total y sus diversas partes indicadas anteriormente? Hay razones teóricas importantes para utilizar la varianza con respecto a la media aritmética como escala de medida, obtenida tomando cada puntaje individual (por ejemplo 73 por ciento) y la diferencia con respecto a la media (por ejemplo 3 puntos porcentuales si la medida es setenta), elevando al cuadrado esa diferencia (9 en este caso), y, finalmente, sumando todas estas diferencias al cuadrado. Mediante cálculos similares se podrán indicar las diferencias entre promedios por escuela, edades, etc. Siempre habrá una varianza residual, la parte que queda, y que el experimento no explica.

El "*análisis de la varianza*" tiene muchas aplicaciones potenciales: la técnica descrita antes con respecto a la administración de programas públicos de educación y otros servicios sociales, y también para la evaluación de resultados. Aún más importante, es posible asociar mediciones de resultados con características del programa y su aplicación, especialmente en algunos de los recursos o insumos del programa, tales como fondos, personal, locales, material, etc. (Las implicaciones de esta posibilidad serán analizadas más adelante, en el problema N° 7.)

Problema N° 5: Determinar los principios que deberán regir el acopio de datos para todos los tipos de análisis descritos anteriormente.

Ahora podemos discutir los problemas de obtención de datos, después de haber revisado varios tipos de análisis que resultarán útiles en la evaluación de resultados de los programas públicos, en general, y de los programas educativos, en particular.

En primer lugar, es necesario tener conocimiento de la materia que se ha de investigar. Un ingeniero químico, altamente calificado, probablemente no estaría al tanto de las características importantes del sistema de educación primaria ni de las necesidades para llevar a cabo una evaluación de programas, al menos no con el detalle requerido para un análisis estadístico riguroso. Aún la especificación de una "prueba de aritmética de tercer grado" o del "nivel de educación alcanzado por los padres" requiere una conceptualización cuidadosa y una definición precisa. Estos problemas no son tema de estudio de los estadísticos teóricos (si bien los que trabajan en estadísticas aplicadas conocen muy bien todas las trampas, si es que no conocen también las soluciones); en realidad, este campo corresponde a la preparación y experiencia de los educadores profesionales y administradores de programas educativos.

¿Cuál prueba de aritmética debe aplicarse? ¿Qué se sabe de la capacidad de la prueba para discriminar entre los alumnos situados cerca del centro de la distribución? ¿Es ésta una prueba utilizada en los sectores con los cuales sería útil establecer comparaciones? ¿Qué defectos de dicha prueba se conocen? ¿Con qué población ha sido "normalizada"?

Los niveles educacionales alcanzados, ¿deben describirse en tres clasificaciones (por ejemplo, ningún nivel de educación formal, tres años de escuela primaria completos, cuatro años completos) o diez? ¿La conceptualización abarca también la educación en el trabajo, la formación autodidacta y los métodos no tradicionales, tales como el empleo

de la televisión? ¿Deben emplearse sustitutos (por ejemplo un trabajo que requiera mucha lectura y habilidad de cálculo)?

Después de enumerar la información necesaria para satisfacer las necesidades del modelo analítico, definir y precisar cada punto, preparar formularios de encuesta, entrevista, etc., todavía falta mucho por hacer. Para las entrevistas habrá que contratar y preparar personal que trabaje en el campo. Para los cuestionarios por correo habrá que preparar cartas adecuadas con razones convincentes para lograr la cooperación de los encuestados, y un formulario de insistencia para los que inicialmente no respondan. Será necesario preparar a los empleados que han de obtener estos datos de los registros escolares u otros archivos administrativos. Todo este proceso debe ser cuidadosamente supervisado y deberán adoptarse medidas para mantener la calidad (tales como un sistema de "control de calidad") que garantice que todas las instrucciones se cumplen con exactitud.

No sería raro que el descuido en este aspecto del trabajo haga fracasar una investigación en otro sentido correcta. Los organismos especializados en encuestas no sólo dedican gran atención a todos los aspectos de compilación de datos, sino que se aseguran también de que el personal que labora en el campo cumpla las instrucciones, que los encuestados comprendan las preguntas, que la información que proporcionan tenga una base adecuada, etc. La verificación a posteriori de todos los factores que garantizan la calidad del estudio, por ejemplo, se ha hecho común en las encuestas correctamente realizadas. Un pequeño grupo de personal muy bien capacitado vuelve a entrevistar a una pequeña muestra aleatoria de los encuestados que respondieron, esta vez interrogando con mayor profundidad, a fin de establecer la validez y precisión de los datos.

Aún la información de los archivos administrativos requiere una evaluación de calidad, pues también puede haber errores allí, o personas que registran datos sin atenerse estrictamente a las instrucciones oficiales con respecto a dichos archivos.

Problema N° 6: Organizar el acopio de datos estadísticos (es decir, la información cuantitativa estructurada) de modo que pueda obtenerse en forma económica y se puedan establecer generalizaciones y evaluar la precisión de dicha información.

En algunos casos, existe un interés particular por un grupo de niños de una escuela. No hay alternativa en cuanto a la selección de casos si el grupo es tan pequeño que resulta posible incluir a todos los alumnos sin incurrir en un costo exagerado ni en la creación de problemas administrativos. La información resultante, el análisis y las interpretaciones pertenecerán exclusivamente a este grupo específico de niños y a nadie más.

En la administración de programas de servicio público, como la educación, es más frecuente el interés inherente en grandes grupos, demasiado grandes para observar a todos los individuos. A veces el interés se centra en grupos que no pueden ser observados, como por ejemplo los alumnos que pasarán a tercer grado. El problema consistirá, tal vez, en sacar conclusiones sobre los individuos de los cuales no se tienen datos.

(En otras actividades, fuera de la educación, existen claros ejemplos de "aplicación de pruebas destructivas", en las cuales los individuos observados son destruidos o "agotados". Esto ocurre en las pruebas educacionales en forma más sutil: los estudiantes sometidos a un examen particular no pueden representar a niños que no han tenido esta experiencia si se considera que el procedimiento de prueba implica aprendizaje o práctica. La siembra es una buena ilustración de este problema. Si se utilizaran todas las semillas en pruebas para obtener la proporción de germinación, no habría semillas para plantaciones comerciales. Por lo tanto, es necesario encontrar la manera de sembrar pocas semillas y sacar conclusiones válidas para el total, a partir de la reacción de un pequeño grupo.)

Muy a menudo, considerando problemas de costo, de tiempo, interrupciones en las aulas de la escuela y la nece-

alidad de aplicar muy cuidadosamente los métodos de compilación de datos, resulta necesario deducir las conclusiones relativas a un grupo grande, a partir de la observación de una pequeña parte de este grupo, es decir, de *una muestra*.

Sin embargo, los educadores ya saben que los niños acusan grandes variaciones con respecto a todo tipo de características, incluyendo algunas muy difíciles de conceptualizar y medir. Y ¿qué ocurre con las características sobre las cuales no sabemos nada y pueden afectar de manera importante la conducta del niño?

El principio fundamental consiste en establecer una muestra *aleatoria*, de modo que cada individuo del grupo tenga la misma posibilidad de formar parte de ella. La elección tiene que depender estrictamente del azar, sin que entre en juego ningún otro tipo de juicio. Es necesario evitar todo mecanismo de selección que pueda producir parcialidad en los resultados. Las personas que responden a cuestionarios por correo podrían ser significativamente diferentes de las que no responden y esta diferencia puede afectar la información que se desea obtener mediante la encuesta. Las visitas a los hogares podrían desvirtuarse, simplemente porque los encuestadores evitan las entradas llenas de barro de ciertas casas, o las casas en que hay perros que no estén atados, por ejemplo.

Si cada miembro del *grupo sobre el cual se desean realizar inferencias* tiene igual posibilidad de participar en la muestra, es posible determinar inferencias con respecto al grupo total a partir de la muestra aleatoria, mediante la cual se obtiene la información. Esto se aplica por igual a los registros administrativos, a los resultados de pruebas, a la información sobre los hogares y familias y a otros tipos de información y formatos de presentación de la información.

He aquí algunas características fundamentales de una muestra estrictamente aleatoria:

Comentario N° 1: La muestra no debe contener ningún elemento de parcialidad. Ningún juicio de analistas, expertos o autoridades políticas debe influir en la elección de los individuos que proporcionan información (o de aquellos sobre quienes se compile información). Supongamos que

existe una característica personal que no ha sido identificada como importante en los resultados de las pruebas a un grupo de estudiantes; podemos designarla simplemente como "característica A". Supongamos, además, que el 38 por ciento de los estudiantes, que participan en la muestra, tienen esta característica, mientras que el resto no la tiene. La muestra aleatoria también tendrá un 38 por ciento de los individuos con la característica A (considerando sólo la posibilidad de varianza de muestreo, analizada a continuación) y, por lo tanto, *no contendrá juicios preestablecidos con respecto a este factor aunque no se conozca su importancia o siquiera su existencia.*

Comentario N° 2: La teoría matemática de probabilidad permite que el estadístico calcule la *precisión* de una muestra a partir del tamaño; por ejemplo, el número de individuos que integran la muestra, *siempre que haya sido determinada en forma aleatoria.* El número total de individuos del grupo que se estudia (tanto los que forman parte de la muestra como los demás) no afecta la precisión de la muestra ni su estadística, si bien esto molesta intuitivamente a muchos observadores no técnicos. (El caso de un grupo pequeño y una muestra que incluya la mayoría de ellos constituye una rara excepción.) La precisión depende solamente del tamaño de la muestra aleatoria y es posible obtener cualquier grado de precisión determinando una muestra de tamaño suficiente (la "ley de los grandes números", para muestras aleatorias solamente, y no depende de que la distribución sea "normal" o de cualquier otra forma en particular). Por consiguiente, el técnico estadístico preguntará qué grado de precisión se necesita, de modo que el tamaño de la muestra requerida pueda calcularse *antes de elaborar el resto del plan detallado* y mucho antes del trabajo de compilación de datos propiamente dicho.

Comentario N° 3: A menudo, en la práctica, la muestra tiene una "estructura". En el problema N° 2 se trataba de un ejemplo de muestreo en dos etapas, que tenía por objeto obtener un número suficiente de casos de niños pobres en escuelas en que la mayoría de ellos era de familias pudientes, y la combinación inversa. En este análisis, la precisión de cada casilla en una tabla de contingencia, u otro tipo de tabulación del análisis, puede establecerse por antici-

pado y la recolección de datos puede planificarse conforme a esas normas. En general, pues, el tamaño de la muestra dependerá del subgrupo más pequeño de la población (o personas, objetos, o acontecimientos) *sobre el cual deseamos determinar inferencias* y las conclusiones a las cuales se deba llegar con base en los datos compilados.

Comentario N° 4: Naturalmente, resulta absolutamente esencial formular una clara definición del grupo que se va a estudiar antes de diseñar una muestra o estudio basado en una muestra. (Esto es igualmente importante si ha de abarcar la totalidad del grupo; es decir, establecer una muestra del ciento por ciento, o muestra "de certidumbre"; de otro modo, no habría manera de identificar a los encuestados apropiados ni de controlar la compilación de información.) También es necesario elaborar el "modelo analítico" con sus respectivas comparaciones, verificaciones, etc., de modo que la muestra tenga la necesaria "estructura", es decir, que haya submuestras del tamaño necesario para que sirvan de base al análisis propuesto y a la interpretación de datos.

Comentario N° 5: La posibilidad de evitar juicios premeditados en cuanto a los datos, mediante muestras aleatorias y un control adecuado del acopio de datos, no debe confundirse con el aislamiento y verificación de factores, haciéndolos parte explícita del análisis y de la compilación de datos, como en el caso del análisis de la varianza en que se atribuyen ciertas influencias a determinados factores y su interacción (problema N° 4). En ese caso, teníamos factores definidos por anticipado, cuya influencia deseábamos medir. En el caso discutido inmediatamente antes, intentábamos simplemente evitar que el factor (conocido o desconocido) destruyera la validez del análisis.

Comentario N° 6: A esta altura de la exposición, debería quedar perfectamente claro que el diseño de un estudio estadístico requiere gran cantidad de habilidades y conocimientos especializados. En esto consiste el aporte del estadístico capacitado. Sin embargo, también es fundamental que el administrador de programas o el interesado en un estudio estadístico, sepa *cuáles son sus necesidades*: Qué tipos de grupos deben estudiarse, cómo deben clasificarse y dividirse; qué factores y relaciones deben examinarse con

el objeto de aplicar algunas de las técnicas discutidas anteriormente en este análisis de la técnica estadística. El administrador de programas también debe tener ideas bien definidas sobre el grado de precisión que requieren los datos; por ejemplo, hasta qué grado mínimo debe detectarse una diferencia entre grupos (posiblemente la diferencia entre niños que hayan recibido nuevos servicios o ventajas educacionales y otros que no los hayan recibido).

Comentario N° 7: A estas alturas también debería quedar claro que un estudio de alta calidad a menudo puede realizarse con solo mantener la muestra en el menor tamaño que sea compatible con las necesidades del estudio, y destinar los recursos disponibles a garantizar que la misma sea efectivamente aleatoria, que el trabajo de campo se efectúe correctamente, que las respuestas o la información tengan alto grado de validez y que todas las operaciones sean cuidadosamente controladas. Ciertos observadores consideran anormal que un estudio de muestreo (u otro tipo de compilación de datos) pueda ser mucho más válido y preciso que un estudio que abarque todos los individuos sobre quienes se desea establecer conclusiones; el problema de reunir datos relativos a tantos casos —con un presupuesto fijo— propicia todo tipo de *errores ajenos al muestreo*, que podrían haberse controlado (y que *probamos* se han controlado), con el mismo presupuesto y un estudio mejor diseñado.

(Este sería el momento indicado para que el lector vuelva a la primera sección y revise los objetivos del seminario. Probablemente descubrirá nuevos significados importantes, a menos que haya tenido la ventaja de haber realizado estudios estadísticos profundos antes del seminario.)

Una última observación: el administrador de programas, en nuestro contexto actual, es el jefe, el patrón, la autoridad. El es quien debe decidir lo que es importante. El personal técnico tiene la responsabilidad de velar para que se obtenga la información necesaria si los recursos disponibles lo permiten.

Todo el proceso de elaboración de las preguntas que deberán contestarse, la elección de los grupos que deberán ser analizados, las cantidades y relaciones que han de me-

dirse, las interacciones que deberán medirse, la precisión que deberá lograrse, el plan de análisis total que deberá aplicarse, el tamaño y estructura de la muestra (Incluyendo la posibilidad de que abarque la totalidad del grupo) constituyen lo que se llama "*diseño experimental*". En este sentido, el "experimento" podría consistir simplemente en someter a algunos niños a una prueba de rendimiento en su trabajo escolar, o reunir información sobre cada niño, profesor y escuela. Sin embargo, también podría incluir la *selección aleatoria de casos* que deberán "atenderse" y casos que "no se atenderán". Si se intenta descubrir si un nuevo texto es mejor que otro actualmente en uso, sería necesario elegir, en forma aleatoria, a los niños, o mejor aún, las *clases* que utilizarán el nuevo texto, de modo que las inferencias puedan determinarse con respecto a *todos los niños* en ese sistema escolar particular. También será necesario evitar las influencias foráneas que puedan destruir la comparación, y sería altamente aconsejable medir la interacción profesor-texto a fin de determinar qué tipo de profesores trabaja bien (si se da el caso) con el nuevo texto y qué tipo de características y preparación deben tener los profesores para sacar partido de cualquier ventaja posible.

Problema N° 7: Aplicar la discusión de los principios estadísticos a la administración, concebida desde el punto de vista de los resultados, es decir, a la administración por objetivos, presupuesto de producción, estudios costo-eficacia y estudios costo-beneficio. La táctica clave consiste en definir y medir resultados —resultados de la educación en el contexto actual— y luego utilizar esta información cuantitativa y estructurada en todas las formas posibles para la planificación, preparación de presupuestos y en la administración de una labor educativa eficaz, real y provechosa.

Probablemente, la mayor innovación en administración pública desde la Segunda Guerra Mundial, ha sido el énfasis que se ha puesto en las técnicas rigurosas de administración, *basadas en la medición de la producción —resultados— y en la relación de los recursos —insumos— y las operaciones que permiten alcanzar la producción.*

A fin de comprender exactamente lo que esto significa, debemos establecer una distinción: se puede relacionar todo tipo de estudio con el "producto" o resultado de un programa, o con las *operaciones* que tienen como fin producir un resultado deseado. Por ejemplo, la mayoría de los estudios aún se centran en los insumos y operaciones y no en los resultados, en parte porque los estudios concebidos desde el punto de vista de los resultados son por lo general mucho más difíciles. Por consiguiente, el informe del administrador —o su solicitud para un presupuesto— indicará la cantidad de dólares, pesos, escudos o guaraníes invertidos; el número de niños que asisten a la escuela; el número de profesores pagados, etc. Todo esto constituye una contabilidad de recursos invertidos en un intento por educar a los niños, con algún sentido útil. También habrá logros en cuanto al proceso: contratación y capacitación de profesores, construcción de aulas, adquisición de textos escolares. Ninguno de estos logros mide la variación de las capacidades y actitudes que constituyen los objetivos del sistema educativo.

Es posible y útil estudiar estas operaciones en detalle. Se ha desarrollado toda una disciplina de matemáticas aplicadas —junto con la estadística— para estudiar las operaciones al servicio de la administración: la "ciencia de la administración". Esta disciplina contiene una amplia gama de técnicas que ni siquiera se han mencionado anteriormente: programación, control de inventarios, todo tipo de métodos para lograr un nivel óptimo con el objeto de determinar la mejor manera de llevar a cabo *una operación*, no de lograr un producto. Existen técnicas útiles, como los métodos de "*curso crítico*", que pueden contribuir a la realización más económica y rápida de una operación que deba llevarse a cabo de una sola vez y completarse en una fecha determinada. La aplicación de métodos de la "ciencia de la administración" en los sistemas educativos no está muy avanzada; debe prestársele mayor atención, pero éste no es el tema del seminario.

Entonces, ¿cómo podemos aplicar el estudio de resultados para una mejor administración de los sistemas educativos?

En primer lugar, debemos saber cuáles son los resultados del sistema educativo que se desean obtener, y debemos medir estos resultados.

El término *eficiencia* se utiliza a menudo para indicar el costo unitario del producto: por ejemplo, los fondos invertidos o el número de horas-profesor por alumno que pasa de cuarto a quinto grado. Cabe observar que esto difiere de la contabilidad de costos, que típicamente asigna costos a los insumos y distribuye éstos entre diferentes usos; en este caso relacionamos el costo con el *producto*, y el costo, a su vez, puede definirse para cada uno de los principales tipos de recursos utilizados en la obtención del producto.

La *eficacia* se define generalmente como una medida del éxito en el logro de un propósito definido (por ejemplo, el número de alumnos de cuarto grado que concluyan los estudios satisfactoriamente y son promovidos al grado siguiente). Podría parecer que la eficiencia y la eficacia representan, a menudo, objetivos incompatibles: es posible lograr reducciones de costo si solo los casos más fáciles (en otras palabras, los niños que pueden completar el

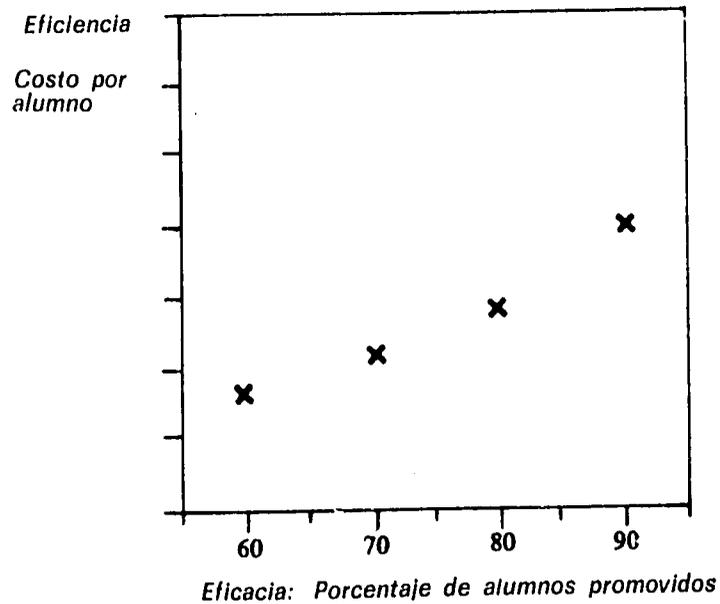
trabajo del cuarto grado sin dificultad) se benefician con el servicio. Es posible trazar un gráfico (Gráfico V-7) del costo promedio (eficiencia) de diversos grados de eficacia. Si los niños cuya enseñanza es difícil (es decir, que tienen problemas en aprender) deben completar su trabajo, requieren un esfuerzo extraordinario —recursos— a fin de alcanzar ese objetivo.

La *administración por objetivos* es, en rigor, un método para determinar, por anticipado, los objetivos de *producción*

GRAFICO V-7

COSTO ESTIMADO POR ALUMNO Y PORCENTAJE DE ALUMNOS DE CUARTO GRADO PROMOVIDOS A QUINTO GRADO

(Ejemplo de una relación probable entre eficiencia y eficacia de los resultados educacionales)



que se desea alcanzar y luego realizar todos los esfuerzos posibles para lograr la finalidad prevista. También se utiliza como método para determinar objetivos de "proceso" (contratar 17 profesores más —medida de proceso— en lugar de promover 87 por ciento de los alumnos de quinto a sexto grado; que sería una medición de resultado, mientras que contratar 17 profesores más sería una medición de proceso). Frente a esta ambigüedad, es necesario determinar, en primer lugar, si la administración por objetivos, en un caso particular, se establece desde el punto de vista de la producción o si se supone una estructura menos precisa.

Los estudios sobre *costo-eficacia* constituyen una de las principales razones para medir el producto en primer lugar. Dichos estudios tratan de determinar las combinaciones de políticas y recursos que producen un resultado determinado al menor costo. Por consiguiente, si se desea crear una situación en la cual el 80 por ciento de los alumnos que ingresan al primer grado completará satisfactoriamente los ocho grados de escuela (elemental o primaria en algunos países, parte de los grados de secundaria en otros), se puede escoger entre muchas posibilidades de tecnología educativa: diversos tipos de material didáctico, normas para el reclutamiento y formación de profesores, escala de matrículas, etc. También existen políticas asociadas con todo tipo de tecnología: textos escolares gratuitos o suministrados por el alumno, sueldos de profesores, normas para la promoción de un grado a otro, etc. ¿Qué combinación permitirá obtener el puntaje más alto —niños que completan los ocho grados— por unidad de recursos invertidos?

La exploración sistemática de este asunto exige experimentación y evaluación —e innovación— más disciplinada que la realizada por la mayoría de los sistemas escolares hasta hoy. Tal vez uno de los resultados de este programa de capacitación de la Universidad de Georgetown consistirá en promover la experimentación, con propósitos más claros, sobre la mejor forma de alcanzar resultados significativos de los sistemas educativos entre varios países.

El análisis del costo-beneficio involucra otros problemas de gran dificultad técnica. Al medir la relación costo-eficacia, comenzamos con una medida convencional del re-

sultado; el único problema es cómo lograr el resultado al menor costo en recursos o insumos. El niño que ha completado el octavo grado tiene igual valor para la sociedad, la comunidad, su familia y para él mismo, independientemente de la forma en que se haya logrado el resultado.

En el análisis del costo-beneficio, tratamos de *comparar diferentes tipos de programas de servicio social*. Por lo tanto, tratamos de asignar un *valor* a diferentes servicios que pueden obtenerse con los mismos escasos recursos (fondos, personal). ¿Cómo comparar tres programas: uno que implique el logro de un nivel educacional, otro que involucre mejores servicios materiales y de salud para los niños, y otro que busca el desarrollo económico mediante empleos mejor remunerados y una mayor posibilidad de empleo de la población? Todos estos objetivos o metas son válidos; nadie podría negar su valor. Sin embargo, es inherente a las labores presupuestarias que se asignen recursos a ciertos programas, y, en el mismo proceso y por alguna razón, se excluyan otros o se les proporcionen menos recursos. Puede resultar desagradable para el ciudadano común y ser una decisión difícil para el gobierno, pero más educación (a cualquier nivel de costo-eficacia) significa menos servicios de salud para la madre y el niño y menos desarrollo económico. En la práctica, el problema no reside en elegir solamente entre los resultados con valores asignados —beneficios— para tres posibilidades, sino en elegir entre *todos los programas* o actividades que compiten por el mismo conjunto de fondos, personal y el apoyo del gobierno.

No hace falta recordar que esta parte de la administración, concebida desde el punto de vista de los resultados, se encuentra francamente subdesarrollada. Los estudios que se han hecho tienden a evaluar *sólo en términos del dinero*, que puede relacionarse con las mediciones de resultados especiales (no evaluadas) de los diversos programas: mayor recaudación de impuestos asociada con individuos que reciben más educación durante su vida; incremento del producto nacional bruto en relación con diversos incrementos del nivel educativo; evitar futuros costos en programas de bienestar para los trabajadores mejor capacitados; ahorro de gastos de hospital y otros gastos médicos mediante programas de medicina preventiva y tratamiento

preventivo a problemas médicos potencialmente graves. ¿Qué hace falta? Todo lo que no puede medirse por medio del dinero, especialmente con respecto a futuros gastos públicos adicionales, mediante diversos mecanismos fiscales. Esta omisión es de gran envergadura en todo sentido, excepto como medida fiscal de los gobiernos: en el programa de salud no se evalúa —no se establece ningún sistema de valores— la recompensa que representa la buena salud del individuo y la recompensa consiguiente para la familia, la vida más rica y plena de los individuos mejor educados, la satisfacción de una profesión u ocupación altamente productiva. En resumen, no contamos con un sistema válido para medir la *calidad de la vida*, ni siquiera medios convencionales útiles para efectuar esa medición. Por lo tanto, tenemos diferentes mediciones de producto para cada tipo de programa y para cada componente de cada programa, como un año adicional en la primaria y la secundaria.

En resumen, se ha preparado el camino —contando con suficiente disciplina intelectual y determinación— para los estudios de costo-eficacia, destinados a definir el "mejor" modo de obtener el producto óptimo en cada programa. La "mejor" solución será siempre la "mejor", independientemente del valor que se asigne al resultado. La ausencia de un sistema global de valores se deja sentir cuando se trata de comparar *diferentes tipos de programas*. Gran parte de la demanda de estudios de beneficios hace caso omiso de este problema, pues no se discuten en absoluto los aspectos técnicos de la valoración de resultados diferentes, de programas diferentes. Hay abundante literatura sobre "cómo producir el cambio", como si la renuencia personal o la oposición de las autoridades constituyeran el único obstáculo.

Otro aspecto del problema en la medición del valor de los resultados es la atención que se otorga a los problemas que *pueden* resolverse, tales como el tratamiento de valores en el futuro. Si se puede asignar un "valor" (salarios adicionales, recaudación tributaria adicional asociada al aumento de salarios, etc.), a un programa —por ejemplo, educación vocacional de cierto tipo— en cada año futuro, es posible aplicar principios bastante tipificados de análisis de financiamiento y técnica actuarial, a fin de obtener un "valor

actual" estimado de cada año, aunque la ganancia se obtenga dentro de muchos años. Al beneficio de cada año (medido) se le hace un "descuento" —reducción— ya que se estima que los pagos futuros valen menos que el ingreso recibido hoy. Tenemos, pues, la estimación de la parte del beneficio, que se convierte en una moneda corriente y deja de lado lo que muchos analistas podrían considerar un componente mucho mayor y difícil de medir.

Es de esperar que se puedan elaborar medios convencionales más satisfactorios para la valoración de los diversos aspectos de la "calidad de vida". Mientras tanto, habrá que aceptar la táctica de asignar valores monetarios a los beneficios cada vez que esto pueda efectuarse de manera razonable. Estos resultados serán considerados como una *exposición demasiado débil* de los verdaderos beneficios y a menudo irán acompañados de una lista de los factores (buena salud, empleo satisfactorio, vida intelectual y artística más plena, etc.) que no han sido considerados, dejando al lector, o a las autoridades encargadas del presupuesto, la tarea de decidir en qué grado esta exposición de la medición de beneficios es demasiado débil.

Sin duda, la administración, concebida desde el punto de vista de los resultados, ha hecho muchos aportes a su campo y, en particular, a la administración de programas públicos. Si se aplica con cierta profundidad de conocimiento y con un sentido de humildad, puede contribuir en gran medida a la importantísima tarea de adoptar decisiones de administración pública. Puede, en ciertas ocasiones, resultar una especie de monstruo fuera de control, si cae en manos de personas no preparadas o demasiado simplistas. En este sentido, es similar a otras técnicas utilizadas en la administración de programas públicos.

APENDICE A

Usos especiales

Como se indicó en el Prefacio, parte del material y los cuadros y gráficos de los seis seminarios pueden ser particularmente útiles para los grupos específicos con los que se han programado las discusiones.

Hasta cierto punto, las selecciones serán hechas a discreción del moderador de los seminarios y pueden depender, en parte, del tiempo disponible y de la cantidad de material que se incluya en el seminario.

Breve comentario sobre la selección de los seminarios para diferentes grupos

Generalmente, los Seminarios III y VI, además del material introductorio del Seminario I, son de mayor interés para maestros, funcionarios directamente involucrados en la administración escolar e instructores de institutos para formación del magisterio.

Los Seminarios II, IV, y VI son más apropiados para suministrar información básica para los altos funcionarios encargados de la política de planificación general, funcionarios de jerarquía intermedia de los ministerios de Educación y de Hacienda, e instructores de institutos de capacitación para planificadores y administradores de escuelas. El Seminario I suministra una discusión básica sobre conceptos y la multiplicidad de las mediciones de resultados. El Seminario V, de una manera general, ayuda a informar a los altos funcionarios de instituciones de planificación y ministerios de Educación y de Hacienda, sobre lo que pueden esperar de estos análisis.

Usos especiales del Seminario V

El material que se brinda en el Seminario V será de mayor utilidad para administradores y personas con cargos administrativos, encargados de presupuestos, planificación

y administración. También se beneficiarán las instituciones que capacitan personal para labores administrativas.

Este seminario no es apropiado para maestros, profesores de instituciones para la formación del magisterio o directores de escuelas.

GRAFICO V-1

PUNTAJES DE PRUEBAS APLICADAS A NIÑAS DE DIEZ AÑOS DE EDAD EN CUARTO GRADO

Número de alumnas

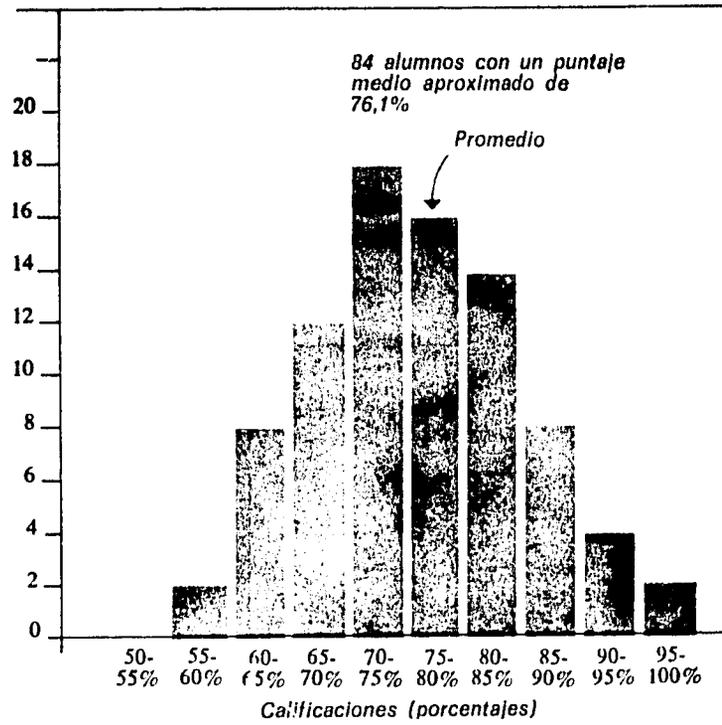
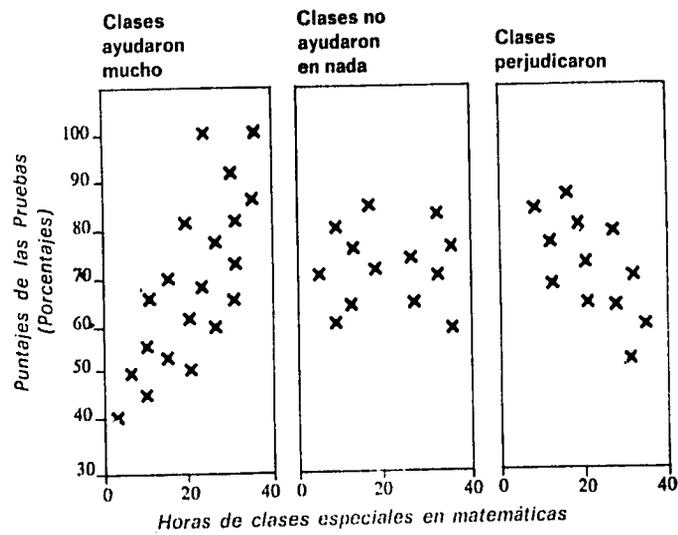


GRAFICO V-2

RELACION ENTRE PUNTAJES EN MATEMATICAS Y HORAS DE CLASES ESPECIALES

(Ejemplos de diagramas de dispersión)

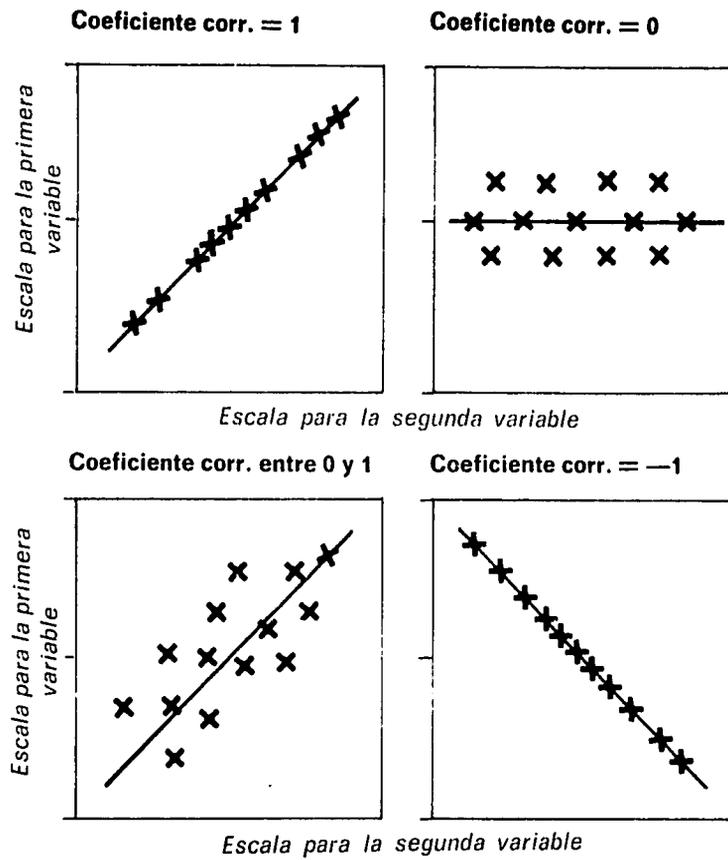


(Datos hipotéticos)

GRAFICO V-3

CORRELACION ENTRE DOS VARIABLES

(Muestran el coeficiente de correlación)



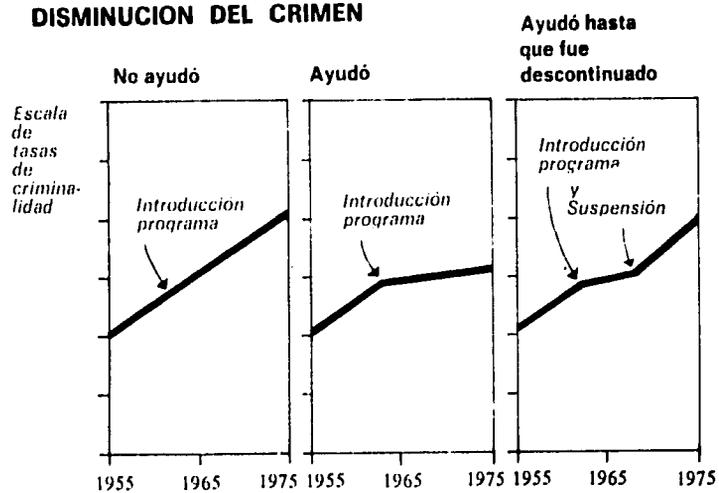
(Datos hipotéticos)

GRAFICO V-4

EFFECTOS DE LOS PROGRAMAS DE PREVENCIÓN DEL CRIMEN

(Ilustración del análisis de tendencias)

INTRODUCCION DE UN PROGRAMA PARA DISMINUCION DEL CRIMEN

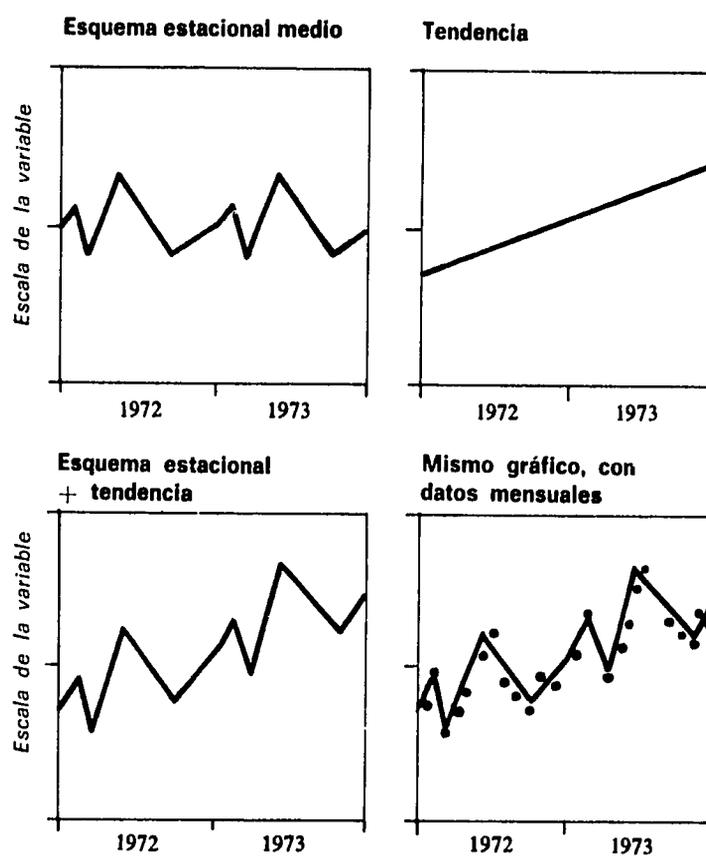


(Datos hipotéticos)

GRAFICO V-5

VARIACIONES A TRAVES DEL TIEMPO

(Ilustración de los componentes de una serie cronológica)



(Datos hipotéticos)

CUADRO V-6

**PUNTAJES DE PRUEBAS ESCOLARES
EN DOS COMUNIDADES**

(Muestra la combinación de influencias explicativas y la sensibilidad del análisis estadístico)

Escuelas de comunidades en que el nivel de ingresos familiares es:

Niños cuyos padres han alcanzado un nivel educacional:

	Alto	Bajo
Alto	214: 83%	2: *
Bajo	4: *	307: 72%
Alto	50: 83%	200: 81%
Bajo	200: 75%	50: 72%

Clave: Cada casilla indica el número de niños sometidos a prueba y el puntaje medio obtenido.

(Datos hipotéticos)

Muestra demasiado pequeña; promedio no es significativo.

GRAFICO V-7

**COSTO ESTIMADO POR ALUMNO Y
PORCENTAJE DE ALUMNOS DE CUARTO
GRADO PROMOVIDOS A QUINTO GRADO**

(Ejemplo de una relación probable entre eficiencia y eficacia
de los resultados educativos)

