Nº 5/I

Arturo Nuñez del Prado B.

ESTADISTICA BASICA PARA PLANIFICACION /I

M. Angélica Lasach Kaiser.

CUADERNOS DEL INSTITUTO LATINOAMERICANO
DE PLANIFICACION ECONOMICA Y SOCIAL

Serie I - Núm. 5/I

Apuntes de clase

Arturo Nuñez del Prado B.

ESTADISTICA BASICA PARA PLANIFICACION

Primera Parte
Estadística Descriptiva

Santiago de Chile
1969

Primera impresión: julio de 1969

Texto: Unidad de Composición y Cuadros CEPAL/ILPES
Gráficos: Unidad de Dibujo CEPAL/ILPES
Impresión: Unidad de Reproducción de Documentos CEPAL/ILPES

INDICE

INTRODUCCION

Circula, tanto en castellano como en otros idiomas, una profusa bibliografía sobre estadística; la conocen, frecuentan y utilizan, provechosamente por cierto, los interesados. Por ello, quizás cabría preguntarse qué sentido tiene aumentar, con un nuevo trabajo, el número ya abundante de publicaciones sobre la materia. Sin embargo, las siguientes consideraciones parecen justificar el esfuerzo.

En primer término, la bibliografía sobre estadística, en su gran mayoría, presenta métodos estadísticos puros, ya sea por su tratamiento matemático o desde un punto de vista descriptivo, menos riguroso, aunque más accesible a los investigadores en general. Por ello no es fácil encontrar un texto que resuma un conjunto de métodos estadísticos necesarios para el investigador en materia de planificación. Uno de los propósitos de este trabajo consiste, precisamente, en realizar una selección de temas, los más útiles, que permitan al planificador disponer de un instrumental indispensable. Por otra parte, es necesario advertir que el texto fue concebido para estudiantes de cursos intensivos de planificación, donde a la estadística como asignatura se le concede un número reducido de clases y seminarios. Los diferentes temas presentados en el texto, se complementan con una serie de ejercicios de seminarios, a través de los cuales se pretende mostrar problemas y soluciones concretas sobre aspectos que, habitualmente, se presentan durante la realización de planes de desarrollo.

Tampoco puede dejar de considerarse la gran heterogeneidad de alumnos que asisten a estos cursos, pues el tema de la planificación económica y social, compete a muchas disciplinas. El tratamiento de los diferentes temas, en consecuencia, debe hacerse accesible a esta especial categoría de alumnado. La anterior observación se refiere principalmente a los métodos matemáticos utilizados. Para la comprensión de los planteamientos metodológicos ofrecidos, es necesario un manejo ágil del álgebra superior y una cabal comprensión de los conceptos básicos de geometría analítica y cálculo diferencial e integral.

Otro aspecto que se tuvo en cuenta al redactar este texto es que la asignatura estadística brinda un instrumental previo que permite un mejor tratamiento de otras asignaturas de los cursos de planificación. Se hizo, por lo tanto, un esfuerzo de compatibilización con las necesidades de instrumental que tienen asignaturas como análisis económico, contabilidad social, evaluación de proyectos, planificación, etc.

Los puntos anteriores constituyen un conjunto de condiciones, en función de las cuales se ha estructurado un curso especial y se redactó un texto básico; para exponer todos y cada uno de los temas presentados, se tuvieron en cuenta las mencionadas restricciones.

Este Cuaderno corresponde a una versión revisada de los apuntes de clase utilizados en los cursos que se dictan en el Instituto Latinoamericano de Planificación Económica y Social, cuya amplia acogida corroboran sus varias ediciones mimeografiadas. De todas maneras parece necesario advertir que el trabajo que se presenta en modo alguno pretende ser un manual de estadística económica. Tiene tan sólo un propósito didáctivo y constituye un texto de iniciación para el estudio de esta técnica cuantitativa. Está destinado a facilitar el dominio de un instrumental mínimo indispensable en materia de planificación, proporcionando un conjunto de conceptos que aseguren al planificador la posibilidad de percibir tanto las posibilidades como las limitaciones del empleo de indicadores estadísticos, y permitan interpretar cabalmente los estadígrafos de uso más frecuente. Interesa al mismo tiempo que el planificador pueda establecer, en estudios más profundos, relaciones de trabajo eficientes con estadísticos y econometristas.

Finalmente, es preciso destacar que este texto está constituido en buena parte por resúmenes y adaptaciones de temas tratados en otros, como los Apuntes de Estadística del profesor Pedro Vusković, el Curso General de Estadística del profesor Enrique Cansado, la Introducción a la Estadística de G.V. Yule y M.G. Kendall, y la Estadística General Aplicada de F. Croxton y D. Cowden. El objetivo perseguido no es la originalidad conceptual; antes bien, se pretende brindar un texto didáctico, que incluya un conjunto de temas íntimamente relacionados con la planificación, complementados con ejercicios que ilustren conceptos de manejo cotidiano en la práctica.

I. ESTADISTICA Y PLANIFICACION

A. Las necesidades de información

Tomar decisiones racionales, supone disponer de informaciones fieles, en cantidad suficiente, y con la oportunidad debida. Las decisiones erróneas se deben más a la falta de información que a deficientes evaluaciones de ésta. Cabe reconocer, desde un comienzo, que una buena proporción de las informaciones que debe manejar un planificador son de tipo cuantitativo. La estadística convencional presenta métodos que facilitan el análisis sobre variables cuantitativas; para el análisis de variables cualitativas, la estadística no paramétrica ya alcanzó un grado de desarrollo que permite tratamientos serios y de verdadera utilidad. Existe, por lo tanto, un cuerpo de conocimientos que posibilita el análisis tanto de variables cuantitativas como de las cualitativas. Sin embargo, las decisiones en planificación implican evaluaciones no estadísticas para ciertos aspectos del complejo problema. Conviene tener presente, por lo que antes se ha dicho, que la estadística es un instrumento útil que permite analizar una parte de los fenómenos que condicionan las decisiones en planificación, donde intervienen aspectos económicos, sociales y políticos con todas sus interacciones. Puede lograrse el conocimiento de la estructura de importaciones, la distribución del ingreso y la composición de fuerzas políticas, empleando instrumentos estadísticos; pero estimar la reacción de las clases populares ante una cierta tasa de crecimiento del consumo, exige una evaluación subjetiva e implica, en parte, juicios no cuantitativos.

Ha surgido una falsa controversia sobre las necesidades de información. Por una parte, reducir los problemas que aparecen en planificación a términos puramente cuantitativos, sería simplificar en exceso el problema; por otra, desconocer la utilidad de los instrumentos, significaría circunscribir la discusión a un marco muy general y peligrosamente confuso. Estas dos posiciones no constituyen alternativas; y tampoco es posible encontrar defensores de una u otra. Identificarse con alguna de estas posiciones extremas, significaría no comprender realmente qué significa la planificación. Por lo demás, tampoco parece posible disociar el planteamiento de problemas generales, de una necesaria cuantificación. Así, por ejemplo, proponer una cierta redistribución del ingreso, supone conocer con bastante detalle cuantitativo, la distribución existente y la redistribución deseada; señalar la necesidad de distribuir ingresos, es apenas indicar vagamente un problema y un objetivo, cuya factibilidad y profundidad no pueden juzgarse si se desconocen cuantitativamente los tramos de ingreso y las proporciones de la población que los capta. Este ejemplo en modo alguno pretende postular que las informaciones deben ser necesariamente precisas; en planificación se puede trabajar con estimaciones, que en general suponen aproximaciones. Tampoco es indispensable una precisión rigurosa; saber, por ejemplo, que el coeficiente de inversión es del 10 o del 12 por ciento no establece una significativa diferencia para calificar esta situación. Vale decir, se pueden tolerar desvíos razonables, mas es preciso tener conciencia de qué significa un orden de magnitud, un intervalo razonable o, en general, la utilidad de una aproximación. Es indispensable, cuando se trabaja con estimaciones, plantear un intervalo donde, con una probabilidad cercana al 100 por ciento, se encuentre el valor verdadero; ahora bien, este intervalo tendrá una amplitud razonable, siempre que las conclusiones no sean significativamente distintas en uno y otro extremo de dicho intervalo. Admitido en planificación este criterio puede permitirse, y a veces no hay otra alternativa, la cuantificación aproximada, sobre la base de estimadores. Juicios de valor y opiniones generales no bastan; es preciso que las justificaciones obedezcan a deducciones lógicas, identificando magnitudes, proporciones e indicadores en general, pues todos ellos ayudan, y a veces en forma insustituible, a calificar fenómenos y obtener conclusiones objetivas y consistentes.

Dentro de este contexto general la acumulación de informaciones y su uso racional es inherente a un proceso de planificación. A continuación, se detalla el empleo de los métodos estadísticos durante las diversas etapas de un proceso de planificación.

1. En el diagnóstico

Identificar los principales problemas de un sistema socioeconómico implica, por una parte, la investigación de una perspectiva histórica, la identificación de la estructura de poder, el comportamiento de estos grupos y los resultados que producen en las principales variables de evaluación: distribución del ingreso, estructura del comercio exterior, estructura y magnitud de la inversión, estructura del consumo, etc. Ahora bien, puede ser útil, aunque no suficiente, conformarse con impresiones cualitativas sobre los aspectos citados; esta clase de información, si bien admite cierto tipo de calificaciones, no permite una comprensión cabal del funcionamiento del sistema considerado. Un diagnóstico completo debería incluir una descripción de su funcionamiento.

Para responder a este punto, resulta casi innecesario insistir sobre la urgencia de contar con información básica y con métodos estadísticos que permitan tratar y evaluar dicha información; se necesitan tanto los análisis de corte transversal, como los de procesos dinámicos. En otras palabras, la estadística descriptiva, vale decir, los indicadores de posición, dispersión y asimetría hacen posible caracterizar situaciones en el tiempo desde diferetnes puntos de vista. Además el análisis de series cronológicas, los análisis de regresión y correlación, se utilizan para analizar el comportamiento y la interrelación de las variables en el tiempo. Como es evidente se trata de análisis parciales, por etapas, que necesariamente deben compatibilizarse para obtener conclusiones consistentes. Cada estadígrafo o indicador muestra un aspecto, una faceta del problema; es necesario disponer, pues, de un conjunto de indicadores, estáticos y dinámicos, para así poder analizar el problema desde ángulos diferentes que den un marco integral al estudio.

Sin embargo, aún disponiendo de indicadores estáticos y dinámicos, tampoco es posible dar por terminada esta etapa sin antes disponer de una descripción detallada del funcionamiento de la actividad socioeconómica; y esta descripción puede ser literal o matemática. Esta última forma tiene evidentes ventajas, en cuanto a consistencia, claridad, precisión y el planteamiento explícito de supuestos. La estadística juega un papel fundamental en la determinación de funciones y en la especificación de valores de los parámetros que intervienen en las diferentes relaciones.

2. En la prognosis

En las proyecciones de las variables socioeconómicas resulta particularmente importante la aplicación de métodos de regresión y correlación y las estimaciones por razón, proporción y elasticidad, cuando se acepta el cumplimiento de los supuestos implícitos en las tendencias históricas, y se admiten reacciones similares a las del pasado, frente a cambios en determinados factores cuyo comportamiento futuro pueda preverse, o sobre el cual puede influirse deliberadamente. Esta visión anticipada del futuro permite comprender la magnitud de los problemas en una dimensión potencial. En los diagnósticos suelen percibirse graves problemas, pero una proyección deja ver el empeoramiento de esas situaciones, y los efectos que tendrían, si no se adoptan decisiones que cambien el sentido de esas tendencias. Las extrapolaciones estadísticas son precisamente los instrumentos apropiados para establecer prognosis.

3. En la configuración de la imagen

Si por la imagen a largo plazo que se puede hacer de un país se entiende un conjunto de intenciones condicionadas por expectativas de variables no controlables, que refleje los cambios deseables para las situaciones percibidas en el diagnóstico, parece lícito convenir que dicha imagen no sólo debiera incluir características cualitativas, sino que admitiría y sería beneficioso que en la imagen se precisen hasta dónde la información permita caracterizaciones cuantitativas de importancia. Si en el diagnóstico, por ejemplo, se advierte que la estructura de exportaciones e importaciones es perniciosa, y por lo tanto debe ser modificada, postular que se debe cambiar esa situación en favor de una estructura de exportaciones que conceda mayor importancia al rubro manufacturas a expensas de una menor participación de materias primas, es sólo formular una intención general, que no puede ser calificada por su factibilidad ni por la importancia del objetivo y su compatibilidad con otros objetivos que plantea la imagen. Para cumplir con los requisitos mencionados, no se puede negar la necesidad de disponer de un conjunto de datos cuantitativos y cualitativos que pueden obtenerse aplicando métodos estadísticos idóneos. La imagen debería ser la culminación de un plan a largo plazo, y como tal, factible y consistente. ¿Cómo calificar factibilidad y consistencia en un plano muy general de ideas? No se pretende justificar la caracterización de una imagen bajo una maraña de datos que harían perder la visión central de las metas propuestas. Es a todas luces evidente que una imagen, por el plazo dentro del cual se la sitúa y por las informaciones que se disponen en el momento de configurarla, no puede abarcar detalles si se desea hacer un trabajo serio y honesto. Pero sí debe contener un conjunto de variables que permitan juzgar la audacia o la cautela en el planteamiento de objetivos y su compatibilidad intrínseca. El estudio de series cronológicas, las proyecciones por elasticidad, los modelos de regresión y correlación y, en general, los modelos de planificación, constituyen herramientas de enorme utilidad.

4. En la delimitación de la estrategia

Constituye la estrategia el vínculo entre diagnóstico e imagen; que en la práctica significa decidirse por alternativas referentes a medidas muy generales de política económica que conducen al logro de los objetivos planteados en la imagen.

Naturalmente existe una relación íntima entre diagnóstico, imagen y estrategia; y estos conceptos del proceso de planificación están mutuamente condicionados. A los efectos de su presentación se los considera por separado, pero eso no debe hacer suponer que se trata de elementos disociados; antes bien, supone un encadenamiento de acciones en el tiempo. Se dijo antes que la elección de una estrategia resulta de tomar en cuenta alternativas; supuesto esto, es necesario evaluar dichas alternativas en la medida que las informaciones lo permitan. Necesariamente es indispensable definir grandes proyectos ligados a las alternativas de estrategia e imagen planteadas. Es fundamental un mínimo de cuantificaciones, proyecciones y estimaciones que sustenten las evaluaciones. Cambios sustantivos y concretos, implícitos o explícitos, en la imagen y estrategia deberían evaluarse tanto desde puntos de vista cualitativos, como cuantitativos. Nuevamente parece oportuno insistir sobre el hecho que es indispensable un mínimo conocimiento de métodos estadísticos para enfrentar este complejo problema. Es evidente que la base de sustentación de imágenes y estrategias realistas está constituída por el acopio de informaciones respecto de perspectivas en el avance tecnológico, en las prospecciones de recursos naturales, en las variaciones de las estructuras sociales, económicas, institucionales, políticas, en la composición de la estructura de poder, etc. Supone, por lo tanto, mecanismos ágiles de captación de informaciones que permitan anticipar la toma de decisiones, de acuerdo a las expectativas que se tengan.

La obtención de informaciones durante un proceso de planificación, debería ser una tarea continua y no esporádica; y por otra parte, la información primaria que se obtenga debe ser depurada, clasificada, resumida y analizada, aplicando adecuadas técnicas estadísticas.

5. <u>En el plan a mediano plazo</u>

Consiste, por una parte, en la actualización de la imagen a un plazo que media entre los tres y diez años, con mucho mayores detalles y especificaciones. Por otra, implica un conjunto de decisiones de política económica mucho más concretas e individualizadas que en el plan a largo plazo; se definen proyectos sectoriales y regionales que dan contenido físico a la estrategia; se hace indispensable una evaluación más precisa de objetivos y metas. Las técnicas de proyección y los modelos de programación constituyen los instrumentos cuantitativos más utilizados durante esta etapa. Estimaciones por regresión y elasticidad, la utilización de matrices de insumo producto, los balances de materiales, etc., son instrumentos a los cuales nuevamente se apela, por lo general, en forma intensiva. La confección de una matriz de insumo producto, exige el manejo de una serie de métodos estadísticos; las actualizaciones de matrices obsoletas implican correcciones de coeficientes que suponen una racional utilización de números índice y de técnicas muestrales. Proyecciones a precios constantes y a precios variables que alcancen una estructura deseada, justifican una cabal comprensión de los temas señalados. Parece pues innecesario destacar las necesidades de estimación estadística de parámetros que suponen descripciones más detalladas del funcionamiento del sistema socioeconómico necesarias a mediano plazo.

6. <u>En el plan a corto plazo</u>

Para esta etapa es necesario concretar medidas específicas de política económica; pues parece que debería haber una superposición entre las intenciones y las actuaciones. Los mecanismos de evaluación que permitirán calificar la transformación de las intenciones en acciones, requieren una descripción muy detallada del funcionamiento del sistema económico. Es fundamental disponer, en un plano sectorial, institucional, y sociopolítico, más detallado, de funciones que expresan el comportamiento de variables trascendentes, por ejemplo funciones de producción, funciones de ingreso, funciones de consumo, funciones de precios, etc. Todo esto supone que se dispone de una cantidad de informaciones y se utilizan métodos estadísticos que permiten determinar los valores de los parámetros.

7. <u>En el control de avance y en la reformulación de los planes</u>

Sabido es que el proceso de planificación constituye una continua revisión de alternativas a la luz de las nuevas informaciones que se van obteniendo a medida que transcurre el tiempo. Por otra parte, es indispensable estar informado sobre la realidad de la actividad socioeconómica para compararla con las intenciones que reflejan los diferentes tipos de plan. Estos dos hechos, entre otros, obligan a realizar sondeos periódicos para ir percibiendo las posibles distorsiones y para cerciorarse del grado de avance en el cumplimiento de las metas y objetivos del plan. Además es necesario subrayar que el acopio de información debe ser oportuno; verificar hechos históricos siempre será interesante para una serie de propósitos; pero verificar hechos en el momento que se producen es indispensable para las tareas de planificación. El empleo de técnicas muestrales, aplicadas tanto a la estimación de variables cuantitativas como a la de variables cualitativas, por sus indudables ventajas, parece ser una herramienta verdaderamente útil, y esto sobre todo si se piensa en el costo y la oportunidad con que se entregan los resultados. Así, entre los estudios coyunturales que se realizan en Francia, figuran encuestas mensuales sobre el desarrollo de la actividad industrial, sobre las condiciones de vida de las familias, etc., informaciones éstas que se recogen con extraordinaria frecuencia y periodicidad. En los países latinoamericanos, donde las ideas de cambio y reforma, en lo social y en lo económico, implican tareas reflejadas en una cantidad de planes, es fundamental plantear métodos oportunos para captar información y sistemas que permitan evaluarla.

B. <u>Métodos de obtención de informaciones</u>

Para las distintas etapas del proceso de planificación enumeradas en páginas anteriores, es posible mencionar los siguientes métodos que permiten recoger la información necesaria.

1. <u>Censo</u>

Como es sabido constituye una indagación completa, sobre las variables que interesa investigar, de los elementos que componen una población claramente definida. El conocimiento censal de una población asegura la posibilidad de obtener datos fehacientes, siempre que no se cometan errores en la recopilación y en el tratamiento de la masa de datos. En general es muy difícil que un censo, sobre todo cuando la población es muy amplia y diversa, esté exento de alguno de los errores señalados. Mientras éstos no distorsionen significativamente las características reales de las poblaciones censadas. Pueden pasarse por alto, desde el punto de vista de la planificación, desviaciones razonables respecto de los valores verdaderos.

Determinan las desventajas más serias de este método, el elevado costo que significa trabajar con volúmenes de información muy grandes y la demora consiguiente en obtener resultados concretos. Sin embargo, son indispensables investigaciones censales, pese a estas desventajas, por lo menos cada cierto número de años, para posibilitar la utilización de otros métodos durante los períodos intermedios y para disponer periódicamente de informaciones completas.

2. Técnica muestral

Debe entenderse como una indagación parcial sobre las variables que interesa investigar, de los elementos que componen una población. Es parcial, puesto que se considera una fracción, una muestra, de la población; sin embargo esta fracción poblacional, debe ser calificada por su representatividad, es decir, debe asegurar que refleja, con alguna aproximación, las características poblacionales que interesa investigar. Este método, en cierta medida debe ser considerado como un complemento de las investigaciones censales; y complemento en un doble sentido: para intercalar estimaciones entre los períodos censales y para desglosar y agregar otras variables a las investigadas por métodos censales.

Sus grandes ventajas radican en el costo que, en general, es muy inferior al de un censo; en la oportunidad con que se entregan las estimaciones, y también en la posibilidad de realizar indagaciones exhaustivas sobre fenómenos concretos. Tal vez la mayor desventaja de esta técnica la determine la necesidad de trabajar con márgenes de probabilidad inferiores al 100 por ciento, es decir, sin la certeza absoluta que las estimaciones son válidas. Ahora bien, esta desventaja implica riesgos que, por lo general, tienen una probabilidad de ocurrencia inferior al 10 por ciento y muchas veces menores al 5 por ciento. Esta probabilidad puede ser tan pequeña como se quiera; pero su reducción tiene como contrapartida un crecimiento del tamaño de la muestra. El objetivo es trabajar con probabilidades pequeñas de error y con tamaños de muestra que no conviertan en prohibitiva la investigación por razones de costo y tiempo.

3. Estudios de casos típicos

Puede considerarse este método como un límite de muestras pequeñas dirigidas. Consiste en seleccionar algunos elementos representativos de grupos homogéneos de la población estudiada. El análisis de estos casos, que constituyen puntos importantes del abanico completo de la población, puede entregar informaciones que aunque incompletas, representen por lo menos un punto de partida para indagaciones más precisas. Este método debe ser interpretado como una investigación preliminar, como una prueba de factibilidad de posteriores investigaciones muestrales o censales. Sus ventajas en materia de costo y tiempo son evidentes. Y su gran desventaja estriba en el hecho que incorpora cierta dosis de arbitrariedad en la calificación de lo que es un caso típico o representativo; pero como se dijo, su principal aplicación responde a su factibilidad. Piénsese, por ejemplo, en diferentes alternativas de impuestos progresivos y exenciones; si se seleccionan algunos casos representativos: familias de un obrero no calificado, de uno calificado, de un empleado público, de un empleado particular, de un profesional, de un gerente, de un rentista, etc., en cada uno de estos casos podrá probarse cada una de las alternativas planteadas. Algunas de ellas serán fácilmente descalificadas y la discusión puede llegar a circunscribirse a muy pocas alternativas. Una pequeña muestra puede dilucidar la discusión cuando el número de alternativas se ha reducido. Es necesario admitir que este método supone algún conocimiento realista de la población para elegir los casos realmente representativos.

4. Experimentación numérica

El surgimiento y la utilización generalizada de computadores electrónicos permite la posibilidad de tanteos, pruebas de ensayo y error, con una velocidad extraordinaria, empleando un conjunto numeroso de alternativas. Para la planificación a corto plazo, disponer de una descripción detallada del funcionamiento de la actividad económica es una herramienta de innegable utilidad; ahora bien, las descripciones detalladas suponen que se dispone de una extraordinaria cantidad de información. Para ello se hace necesario destinar una buena cantidad de tiempo y esfuerzo a la recopilación de datos, aunque sean aproximados, que permitan alimentar el modelo que representa la formalización matemática de las ideas que se tengan respecto del funcionamiento del sistema económico. Las relaciones del modelo, de definición y comportamiento, implican una enorme cantidad de parámetros, muchos de ellos desconocidos; por ello, una alternativa podría ser el intento de reproducir la historia reciente a través del modelo.

El modelo se alimenta con los datos disponibles sobre variables exógenas y parámetros y se dan valores intuitivos respecto de los parámetros desconocidos. Con ese conjunto de datos, unos fieles y otros estimados, se trata de reproducir la historia, por ejemplo de los últimos dos años, cuyas variables exógenas y las principales variables de resultado son conocidas. De la comparación entre las variables de resultado efectivas y las variables de resultado dadas por el modelo, surgen ideas para reformular las ecuaciones de comportamiento y para modificar los valores intuitivos de los parámetros. Se pueden repetir muchas veces los experimentos hasta encontrar satisfactoria la reproducción que entrega el modelo. Una reproducción aproximada de las variables de resultado es el punto de partida para realizar pruebas de sensibilidad, las que consisten en modificar ligeramente los valores de cada uno de los parámetros y analizar su efecto sobre las variables de resultado; la utilización de coeficientes de elasticidad permite calificar los parámetros como críticos y neutros, según el valor de dichos coeficientes. Esta calificación da una pauta para realizar investigaciones estadísticas y econométricas que permitan estimar aquellos parámetros críticos con métodos más rigurosos y confiables.

Dada la interrelación que tienen las variables en un modelo de este tipo, también será posible estimar, por medio de la experimentación numérica, variables y parámetros que puedan despejarse de otras relaciones de la descripción. Evidentemente esto supone una gran confianza en el tipo de funciones elegidas y en los valores de los parámetros conocidos.

C. La disponibilidad de información en América Latina

Una apresurada generalización sobre las disponibilidades de información estadística en América Latina, corre el riesgo de no tener una validez total. La gran heterogeneidad de países que caben dentro del común denominador de subdesarrollados también se manifiesta en la disponibilidad de datos. Con todo, es posible exponer algunas consideraciones que poseen un cierto carácter general.

En primer lugar, es necesario aceptar como hecho evidente la insuficiencia de información básica sobre varios aspectos de importancia para la planificación. Pocos son los países que poseen informaciones sobre distribución de ingreso y de riqueza, y aparentemente no hay alguno que efectúe este tipo de investigaciones en forma continuada a nivel nacional.

La carencia de datos en una buena parte de los países latinoamericanos llega a ser realmente crítica; en algunos apenas se dispone de uno o dos índices de precios, y donde como agravante se notan características inflacionarias permanentes. Esta carencia de datos sobre precios ofrece un doble inconveniente: la dificultad de plantear políticas de precios y la imposibilidad de trabajar con valores reales, ya sea poder de compra o expresiones físicas; dos aspectos que interesan fundamentalmente a toda tarea de planificación. No es difícil seguir enumerando una serie de deficiencias en materia de datos: se desconocen estructuras de consumo, de inversión, de capital, de producción, de costos, etc.

En segundo lugar, una vez admitido el hecho que hay una aguda escasez de datos, tampoco se puede dejar de reconocer que aún así, no siempre se hace un aprovechamiento óptimo de la escasa información disponible. Además existen problemas de interpretación de datos, provocados por deficiencias de las publicaciones y por falta de entrenamiento en el análisis estadístico.

Sin desconocer que en general la información es escasa, conviene tener en cuenta que hay una mayor cantidad de datos de la que habitualmente se supone, pues también existen investigaciones cuyos resultados no revisten carácter oficial. Es indispensable, por tanto, realizar inventarios de las estadísticas no publicadas. Además, existe una gran cantidad de información en estado primario, cuyo traspaso a tarjetas perforadas o a cintas o discos magnéticos, bastaría para obtener clasificaciones o resúmenes que, a su vez, pueden transformar una masa de datos casi inútil en informaciones aprovechables para propósitos múltiples.

En tercer lugar, no se puede dejar de subrayar que al problema de la escasez, se agrega el del atraso con que se dan a conocer informaciones que pueden tener enorme importancia en un determinado momento, pero que después sólo son útiles para verificar hechos pretéritos, que ya pertenecen a la historia. La idea de oportunidad en materia de planificación adquiere una significativa importancia.

En cuarto lugar, parece útil un breve comentario sobre la calidad de las informaciones. Con frecuencia muchas estadísticas provienen de muestras insuficientes o fueron obtenidas a través de cuestionarios que, involuntaria o inadvertidamente, inducen a cierto tipo de respuestas. Un conjunto de informaciones socioeconómicas, difícilmente puede superar todas las pruebas de consistencia que pueden plantearse.

Finalmente, dada la exigua disponibilidad de recursos para enfrentar investigaciones estadísticas, puede observarse una curiosa asignación de prioridades de tareas. Los organismos públicos y privados, plantean muchas veces investigaciones en forma independiente, sin preocuparse por armonizar criterios de los usuarios ni por dispendiosas duplicaciones.

D. El control de la calidad de las informaciones estadísticas

Del aserto que en planificación no cabe exigir precisión rigurosa, no se sigue necesariamente que se pueda trabajar con cualquier calidad de información recogida; se advirtió ya que las desviaciones debían ser razonables y no exceder ciertos límites de tolerancia fuera de los cuales las conclusiones pueden ser ambiguas. Las siguientes consideraciones quizás puedan facilitar la comprobación de la calidad.

1. Hipótesis

Cuando se plantea una investigación, es fundamental establecer hipótesis previas acerca de los posibles valores que tendrían las variables investigadas. Confrontar los resultados efectivos con las hipótesis previas, constituye una primera prueba que ayuda en la evaluación. Si ambos tipos de datos apuntan en el mismo sentido, aumenta la confianza para admitirlos; si difieren en forma significativa será necesario averiguar la causa de las diferencias, por lo tanto habría que revisar las hipótesis previas y la metodología que se aplicó para obtener informaciones.

2. Consistencia

Es conveniente establecer pruebas de consistencia interna, dentro del conjunto de indicadores que se obtienen como resultado de la investigación, y también consistencia con otros antecedentes disponibles. Todo esto permite calificar mejor los datos con los cuales se trabaja. Recuérdese que, en planificación, uno de los puntos más delicados es lograr compatiblizaciones en distintos sentidos.

3. Representatividad

La evaluación de la representatividad de un indicador es parte de la etapa de control; un ejemplo concreto de falta de una evaluación adecuada, es la jerarquización del nivel de desarrollo de un país basándose sobre los ingresos por habitante.

En primer lugar, un promedio no siempre es representativo; para que lo sea, todos los valores tendrían que estar muy próximos a dicho promedio; el caso de los ingresos por habitante está muy lejos de acercarse a esta situación. En segundo lugar, hay un problema de cómputo de los ingresos reales: evaluaciones del autoconsumo y utilización de sistemas de deflactación muy diferentes entre los países, hacen todavía menos comparables los referidos promedios. A pesar de todo esto hay una tendencia a utilizarlos en las comparaciones internacionales sin especificar sus limitaciones. Cuando se manejan estadígrafos o indicadores que, en general, resumen de manera imperfecta una masa de datos, parece recomendable detenerse en el análisis de su representatividad y especificar qué faceta del problema puede abordarse a través de ellos y qué otros aspectos quedan al margen de tales indicaciones.

4. Sesgos

Cuando se recolectan informaciones por medio de encuestas, es particularmente importante registrar y eliminar, o por lo menos reducir, la posibilidad de trabajar con informaciones que contengan sesgos. Por sesgo se entiende una desviación sistemática de las características observadas respecto de las características reales, por consiguiente parecería casi innecesario subrayar el peligro que esto significa durante la etapa de obtención de conclusiones. Un mal diseño de cuestionarios puede, en cierto modo, inducir las respuestas pedidas. Los encuestadores, a veces, no mantienen una posición neutra e inclinan, en un cierto sentido, a quienes responden. Una encuesta sobre ingresos y consumos generalmente evidencia subestimaciones en el monto de las rentas y distorsiones en la estructura del gasto, sobre todo en los estratos de ingresos altos, y esto por razones fáciles de imaginar. La realización de confrontaciones utilizando muestras yuxtapuestas, permite comprobar la existencia de estos desvíos y, a veces, en algún sentido cuantificarlos. Con todo, tampoco puede desconocerse que el proceso de reducción de estos errores supone esfuerzos que encarecen sustancialmente la investigación.

E. Estadística para planificación

Respecto de las tareas inmediatas que parecen ineludibles, admitida la necesidad de establecer procesos continuos de planificación, se estima conveniente discutir las siguientes ideas:

1. Investigaciones simultáneas

No puede aguardarse que estén disponibles los datos para comenzar una serie de investigaciones socioeconómicas. Parecería más bien que la decisión de iniciar una investigación, debería adoptarse admitiendo dos criterios: por un lado, estructurar la parte conceptual y sustantiva de la investigación; y, por otra, simultáneamente, la investigación estadística pertinente. Muchas investigaciones se postergan por falta de información, cuando todo retardo causa un perjuicio que se magnifica a medida que transcurre el tiempo. Durante una primera etapa, los organismos de investigación, deberían incorporar unidades cuyo objetivo esencial fuese la captación de información, en particular centros de muestreo, para no depender de otros organismos en la disponibilidad oportuna de datos. En ese sentido, las investigaciones interdisciplinarias parecen ofrecer facilidades en el planteamiento de investigaciones simultáneas; desde luego, no es esta una solución integral del problema de la información; antes bien, se corre el riesgo de duplicar esfuerzos y malgastar recursos por falta de comunicación y coordinación entre los organismos. Sin embargo, como etapa transitoria y mientras se organizan planes nacionales de estadística, parecería que podrían evitarse así mayores postergaciones. Por lo demás no sería demasiado oneroso algún grado de duplicación, dada la exigüidad de los recursos destinados a la investigación estadística. En todo caso un mínimo de coordinación entre los diferentes organismos sería suficiente durante esta etapa transitoria.

2. Organización institucional

Un tema sobre el que se insistió mucho y sobre el que poco se hizo, es la adecuación de la estructura institucional que facilitaría el establecimiento de los procesos de planificación. En este sentido, y desde el punto de vista que se está considerando parece realmente conveniente establecer una estrecha relación entre las oficinas de estadística y los organismos de planificación. La cantidad de recursos disponibles para captar información obliga a una detenida jerarquización de las diferentes investigaciones. El análisis de las distintas publicaciones en materia estadística que dan a conocer los países revela que existe una apreciable cantidad de información, que aunque importante es susceptible de ser sustituida o postergada en beneficio de otra más urgente. Es posible que este tipo de consideración esté determinada en parte por algún interés específico en materia de planificación, pero en todo caso el diseño de un plan nacional de estadística supondría evaluar diferentes asignaciones de recursos, considerar los intereses de los usuarios públicos y privados, evitaría innecesarias duplicaciones, delimitaría las responsabilidades de los diferentes organismos para la entrega de la información y, lo que es más importante, establecería plazos concretos para dar por terminadas investigaciones que, muchas veces se arrastran durante lapsos prolongados.

3. Requisitos concretos de información estadística en materia de planificación

Parecería realmente fuera de lugar aquí elaborar, por interminable, una lista exhaustiva; sin embargo, enumerar los principales campos donde parece urgente concretar investigaciones, mostraría la enorme e impostergable tarea que se tiene por delante. (Se parte del supuesto que existen adecuados sistemas de contabilidad nacional: cuentas nacionales, esquemas de fuentes y usos de fondos, matrices de insumo producto, balanza de pagos, etc.).

a) Tal vez uno de los principales campos corresponda al análisis de la distribución del ingreso y de la riqueza. Distinguir entre asalariados y no asalariados supone dos categorías demasiado elementales y heterogéneas en exceso. Es preciso detallar escalas de ingreso que abarquen grupos homogéneos de personas; pero convéngase también que cualquier indagación sobre esta materia ofrece muchas dificultades. La veracidad de los datos sobre ingresos personales (sobre todo en el grupo no asalariado) debería quedar asegurada por investigaciones complementarias sobre la propiedad, el consumo, la tributación, etc. Estudios con controles cruzados son indispensables desde el punto de vista de la calidad de los datos.

   Tampoco parece exagerado admitir que informaciones sobre estos aspectos deberían constituir el punto de partida de la elaboración de cualquier plan que especifique acciones concretas.

b) Otro aspecto, muy ligado al anterior, es el reconocimiento de la estructura de preferencias de los consumidores según sus diferentes niveles de ingreso. La compatibilidad entre una estructura de producción y la demanda final, implica trabajar con coeficientes de elasticidad ingreso y elasticidad gasto a un nivel de desagregación bastante grande; la obtención de funciones consumo por tipos de productos homogéneos, permite dar coherencia a las variables. Por otra parte, investigaciones de este tipo permitirían verificar la representatividad de las ponderaciones efectuadas en los diferentes índices de precios al consumidor.

c) Dentro de las tareas urgentes en materia de recopilación estadística para la planificación, tiene especial importancia determinar el valor del capital por sectores, subsectores o ramas de actividad, que es, por cierto, un problema en extremo delicado. Requiere estudios muy prolijos confeccionar un inventario con especificación de antigüedad, y que al mismo tiempo evalúe el contenido.

d) Las estructuras de precios vigentes no siempre se ajustan a los objetivos sociales y económicos que plantean los planes. Estudiar la formación de los precios, sus componentes, la especulación y las duplicaciones en las distintas etapas del proceso de distribución y comercialización, son tareas de tipo estadístico que también deben tener prioridad en el futuro inmediato. La disponibilidad de índices de precios, en general, es insuficiente para cumplir con algún éxito tareas de cuantificación económica a través del tiempo; hay países, por ejemplo, donde sólo se dispone de un índice de los llamados de "costo de vida" confeccionado sobre una base obsoleta y para un sector escasamente representativo de la población.

e) Dentro de los países pueden delimitarse zonas geoeconómicas bastante diferenciadas, la necesidad de establecer determinado tipo de desagregación de la contabilidad social que tome en cuenta algún tipo de regionalización, permitiría dar, desde este punto de vista, una mayor coherencia a los objetivos perseguidos, a los proyectos elegidos y a las medidas de política económica.

f) El análisis de factibilidad necesario para imponer ciertas medidas de política económica, exige identificar los centros de decisión y la gravitación e influencia que tienen sobre la actividad económica; tales investigaciones tampoco son ajenas al empleo de métodos estadísticos. Lo que se dió en llamar "sociometría" apunta en este sentido.

g) Numerosas variables macroeconómicas se calculan tomando como referencia un período equivalente al año; sin embargo, una desagregación semestral o trimestral, permitiría adoptar medidas oportunas, si los hechos contingentes así lo requieren. Este tipo de desagregación temporal, tendría que basarse, naturalmente, sobre estimaciones algo elementales. Estimaciones sobre producto, consumo, inversión, a un nivel de detalle sectorial parecen factibles utilizando muestras e indicadores parciales. Los indicadores a corto plazo cobran actualmente una significativa importancia.

F. Plan Nacional de Estadística

Las anteriores consideraciones sobre escasez, oportunidad, duplicaciones, necesidad de establecer prioridades en materia de investigaciones estadísticas y la fijación de plazos para realizarlas, constituyen fundamentos sólidos para confeccionar planes de estadística a nivel nacional. La realización y puesta en marcha de un plan de esta naturaleza implica la necesidad de hacer un inventario crítico, en materias de calidad y oportunidad, de las informaciones estadísticas existentes, tanto de las publicadas como de las que se reservan para el uso interno de diferentes organismos. Disponer de un catálogo semejante significaría una considerable ayuda para los investigadores; permitiría realizar un diagnóstico eficiente sobre la disponibilidad de informaciones.

La confección de un plan de estadística supone determinar prioridades respecto de los diferentes tipos de información, de la periodicidad y oportunidad de su publicación. Una encuesta a los principales usuarios, actuales y potenciales, parecería un método adecuado para preparar una lista del conjunto más general de informaciones necesarias.

La frecuencia con que se requiere cada tipo de información, permitiría fijar prioridades a las diferentes investigaciones. Es evidente que los países empeñados en un proceso de planificación, tendrían que asignar recursos y especificar investigaciones, con una orientación clara en ese sentido.

A esta altura del trabajo se hace necesario repetir que la incorporación de técnicas muestrales para recoger información y utilizar computadores electrónicos en el procesamiento y cálculo de datos, conducen a un ahorro de tiempo y a veces de costo particularmente significativos.

La asignación de responsabilidades en cada una de las investigaciones programadas es un punto delicado durante la elaboración de un plan; en este sentido deberán precisarse las responsabilidades de las oficinas de estadística, como así también de cada uno de los demás organismos públicos, las universidades y centros de investigación y de las instituciones privadas.

La Oficina de Estadística dependiente del organismo de planificación, o por lo menos estrechamente vinculada a él, debería asumir la responsabilidad central de la coordinación administrativa y técnica del plan. Es de gran importancia, una vez asignadas las responsabilidades, discutir en profundidad los posibles métodos para captar información, el sistema de control, el tipo de clasificaciones, los indicadores resultantes y sus métodos de cómputo.

La fijación de plazos, entre el comienzo y el término de las investigaciones, es fundamental para garantizar el cumplimiento del programa de trabajo.

Estimaciones de costo, confección de presupuestos por investigación y el logro de financiamientos oportunos, son tareas que completan este conjunto de ideas que pretende ser un esquema primario de discusión.

Nada fácil es imponer la idea de trabajar con seriedad y perspectiva, cuando se programan tareas de investigación estadística, perfilando un verdadero plan. Los problemas a corto plazo, en general, postergan asignaciones de recursos para esta clase de trabajos; cuando por ejemplo existen presiones políticas, económicas y sociales, por captar fracciones de presupuestos exiguos, es difícil imponer un plan de estadística que demanda recursos nada desdeñables y que parece no ofrecer utilidad inmediata. Sólo una evaluación a largo plazo, y una comparación con otras asignaciones de recursos que no siempre puede defenderse, permiten favorecer esta propuesta.

Una de las tareas de los planificadores es, precisamente, llamar la atención sobre la urgencia de un plan de esta naturaleza, indispensable para cualquier proceso de planificación.

II. ESTADIGRAFOS DESCRIPTIVOS ESTATICOS

A. Distribuciones de frecuencia

1. Generalidades

Un conjunto de datos, o masa estadística, puede ser resumido y clasificado de acuerdo a criterios convenientes. Provengan las informaciones de censos o de muestras relativamente grandes, siempre serán útiles para el análisis, ya que difícilmente podrán obtenerse conclusiones válidas de una masa estadística no clasificada.

Los tipos de variables fundamentales, por lo menos para este trabajo, serán los siguientes:

a) variables cardinales: susceptibles de medición cuantitativa; y las que a su vez comprenden:

   i) continuas: variables que pueden tomar cualquier valor dentro de un intervalo (ingresos, estaturas, distancias, etc.)

   ii) discretas: variables que sólo toman algunos valores dentro de un intervalo (número de hijos por familia, número de accidentes de tránsito por día, etc.)

b) variables ordinales: sólo susceptibles de ordenación pero no de medición cuantitativa (grado de cultura de una persona: muy culta, regularmente culta, poco culta, inculta).

Para cada uno de estos tipos de variables, un conjunto de observaciones puede dar origen a una distribución de frecuencias; y ésta debe entenderse como un cuadro o tabla resumen de los datos originales.

En el caso de variables continuas será necesario fijar intervalos de frecuencia para llegar a un resumen efectivo de la información original. El punto medio de cada intervalo se denominará marca de clase y constituirá el valor representativo de cada intervalo. El número de observaciones que correspondan a cada intervalo se denominará frecuencias absolutas.

Una tabla de distribución de frecuencia para variable continua y sus símbolos correspondientes, se representa de la siguiente forma:

Ingresos de profesionales			Número de profesionales
Intervalos		Marcas de clase	Frecuencias absolutas
Y'_{i-1}	Y'_i	Y_i	n_i
Y'_0	Y'_1	Y_1	n_1
Y'_1	Y'_2	Y_2	n_2
Y'_2	Y'_3	Y_3	n_3
.	.	.	.
.	.	.	.
Y'_{m-1}	Y'_m	Y_m	n_m

donde: $Y_i = \dfrac{Y'_{i-1} + Y'_i}{2}$ marca de clase

$n = \sum_{i=1}^{m} n_i$: número de observaciones

$c_i = Y'_i - Y'_{i-1}$: amplitud del intervalo

Estas tablas pueden ser de amplitud constante o de amplitud variable, según los valores que tome c_i.

Cuando se trata de variable discreta o discontinua, la tabla de distribución de frecuencias adquiere la forma siguiente:

Y_i	n_i
Y_1	n_1
Y_2	n_2
Y_3	n_3
.	.
.	.
Y_m	n_m

Cabe destacar que cuando la variable adquiere numerosos valores distintos para abreviar el trabajo, con cierta arbitrariedad y con alguna pérdida de precisión, puede tratarse como una variable continua, formando intervalos de clase.

Por último, en el caso de variables no mensurables, dicha tabla adoptará una forma como la siguiente:

Variable	Frecuencia
Característica A	n_A
Característica B	n_B
.	.
.	.
Característica Z	n_Z

El lector advertirá que las tablas de distribución de frecuencias, facilitan enormemente el análisis. Es muy ventajoso disponer de informaciones clasificadas en intervalos o en valores específicos de la variable, ya que, de esta manera, es posible obtener conclusiones primarias acerca de la variable que se investiga.

Respecto de las frecuencias, es posible y generalmente útil, presentarlas en términos relativos, calculando la proporción que corresponde a cada intervalo o marca de clase sobre el total de observaciones. Se denominan frecuencias relativas, y se simbolizarán por h_i :

$$h_i = \frac{n_i}{n}$$

Tanto las frecuencias absolutas como las relativas son susceptibles de acumulación respecto de los intervalos o marcas de clase. Las frecuencias absolutas acumuladas se simbolizarán por N_i y se definen:

$$N_i = \sum_{i=1}^{i} n_i$$

Las frecuencias relativas acumuladas se simbolizarán por H_i y se definen:

$$H_i = \sum_{i=1}^{i} h_i$$

En general este tipo de frecuencias se acumulan en sentido creciente de la variable, y una frecuencia acumulada N_i indica el número de casos u observaciones donde la variable toma valores a lo sumo iguales a Y_i, en el caso de variable discreta, y a Y_i' en el caso de variable continua. Sin embargo, para ciertos análisis, también es necesario acumular en sentido inverso; de ahí que se hable de frecuencias acumuladas hacia arriba o hacia abajo.

2. Representación gráfica

En general, la representación gráfica de una tabla de distribución de frecuencias, permite percibir con mayor claridad algunas características de la masa de datos que se investiga; por ello resulta bastante más fácil transmitir conclusiones a personas no habituadas a la interpretación de distribuciones de frecuencia, cuando se utilizan gráficos estadísticos.

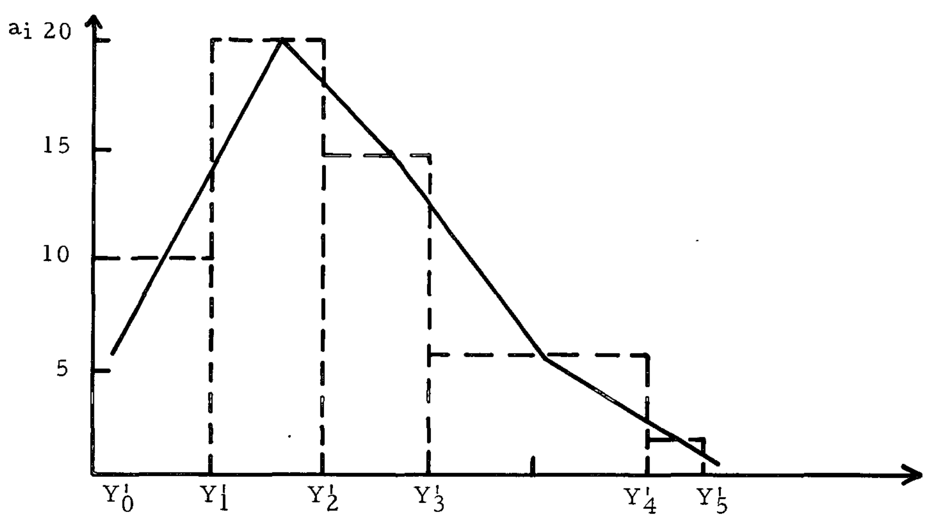
a) Representación gráfica de variable continua. Si se utiliza un par de ejes coordenados, en el eje de las abscisas se representará la variable estudiada, en tanto que en el eje de las ordenadas, se representará las frecuencias correspondientes. Recuérdese que en este tipo de variables, la frecuencia corresponde a un intervalo y por esto se representa mediante una superficie.

Con un ejemplo se ilustrarán estas ideas; admítase, en este sentido, la siguiente tabla correspondiente a las edades de los participantes de un curso de estadística:

Edades		Alumnos	Amplitud de Intervalo
Y'_{i-1}	Y'_i	n_i	C_i
18	22	10	4
22	26	20	4
26	30	16	4
30	38	12	8
38	40	1	2

Puesto que la amplitud más frecuente es 4, puede adoptársela como amplitud unitaria; así el cuarto intervalo tendrá dos veces la amplitud unitaria elegida y el quinto intervalo tendrá la mitad de dicha amplitud. La representación gráfica se hará de la siguiente manera:

Gráfico 1



En el gráfico, para calcular la altura de cada rectángulo, se planteó la relación de superficie siguiente:

$$\text{superficie} = \text{base} \times \text{altura}$$

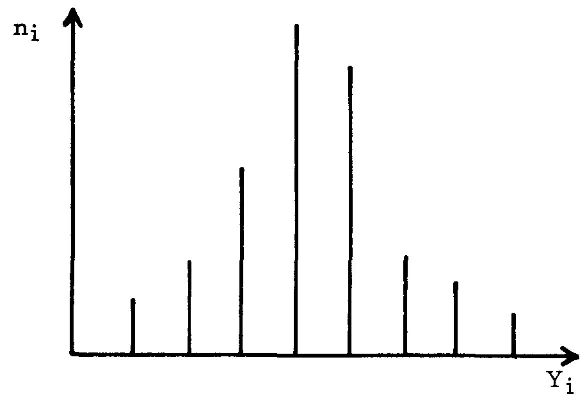
$$n_i = c'_i \times a_i$$

donde c'_i es la amplitud unitaria elegida con el objeto de diseñar un gráfico adecuado. Desde luego que también pudo haberse trabajado con las amplitudes originales, aunque habría sido algo más laborioso.

Este tipo de gráficos recibe el nombre de histogramas y la línea quebrada que une los puntos medios de los lados superiores de los rectángulos se denomina polígono de frecuencias.

b) Representación gráfica de variable discreta. En este caso la frecuencia correspondiente a cada valor de la variable estará representada por una barra vertical.

Gráfico 2



Naturalmente, se puede construir, en forma similar, gráficos que relacionen la variable con cualquiera de los tipos de frecuencias que se han visto, relativas, acumuladas, etc.

A continuación se presentará un ejemplo donde se seguirán todos los pasos necesarios para llegar a una tabla completa de distribución de frecuencias. Supóngase que se dispone de las siguientes informaciones acerca de los sueldos de los obreros de una fábrica (en dólares por mes):

68	48	53	73	100	80	40	55	65	95	85	35	110	120	60
90	70	40	80	100	70	50	55	70	65	45	80	60	90	50
55	60	30	110	110	90	70	60	45	65	80	85	90	68	72
50	40	45	90	105	108	35	45	50	70	82	84	66	38	48

Una de las primeras decisiones que deben adoptarse es determinar el número de intervalos que tendrá la tabla. Para ello es necesario considerar el objetivo que se persigue con el estudio de la variable, qué tipo de diferenciaciones o agrupamientos interesaría conocer; por otra parte, es indispensable determinar el recorrido de la variable, es decir, el menor y mayor valor entre los datos que se analizan. Por último, el número de observaciones, de manera que los diferentes intervalos tengan frecuencias en alguna medida significativas. Cuando existan valores escasos, muy alejados de lo que podría llamarse una concentración central, puede optarse por dejar los intervalos extremos abiertos. Supónganse que, tomando en cuenta las consideraciones anteriores, se decide clasificar los datos originales en 9 intervalos de amplitud constante. Dado que la diferencia entre los valores extremos (30 y 120) es de 90, la amplitud de los intervalos será igual a 10.

Y'_{i-1}	Y'_i	Observaciones	n_i	h_i	N_i	H_i	N_i^*	H_i^*	Y_i
30.0	40	~~////~~ //	7	7/60	7	7/60	60	60/60	35
40.1	50	~~////~~ ~~////~~	10	10/60	17	17/60	53	53/60	45
50.1	60	~~////~~ ///	8	8/60	25	25/60	43	43/60	55
60.1	70	~~////~~ ~~////~~ /	11	11/60	36	36/60	35	35/60	65
70.1	80	~~////~~ /	6	6/60	42	42/60	24	24/60	75
80.1	90	~~////~~ ////	9	9/60	51	51/60	18	18/60	85
90.1	100	///	3	3/60	54	54/60	9	9/60	95
100.1	110	~~////~~	5	5/60	59	59/60	6	6/60	105
110.1	120	/	1	1/60	60	60/60	1	1/60	115
			60	1					

* Acumuladas "hacia arriba".

Para evitar situaciones ambiguas, cuando se presentan observaciones con valores de la variable que corresponden a los límites de los intervalos, puede seguirse el criterio planteado en el ejemplo, es decir, agregar un decimal a la columna de límites inferiores, aunque esto tenga utilidad solamente para clasificar las observaciones, ya que en los cálculos que posteriormente se tratará, tales decimales son despreciados.

Con lo visto hasta el momento, es posible realizar los primeros análisis de un conjunto dado de datos. Tanto la representación gráfica como la tabulación de las distintas clases de frecuencias ayudan a ensayar los primeros juicios. Es necesario insistir sobre la necesidad de tomar en cuenta, en todas las decisiones respecto de la tabla de frecuencias, la naturaleza del fenómeno que se investiga; sobre todo en lo que se refiere al número de intervalos y a sus amplitudes: constantes o variables. Naturalmente que, una vez clasificados los datos originales, será preciso realizar análisis con mayor profundidad, utilizando los instrumentos estadísticos que se detallará en las próximas páginas.

B. Estadígrafos de tendencia central

Una vez conseguida la clasificación de los datos originales, cuyas características más esenciales se destacan, será preciso calcular un conjunto de indicadores que caractericen en forma algo más precisa, la distribución que se está estudiando. Interesa, en primer término, disponer de estadígrafos que representen valores centrales en torno de los cuales se agrupan las observaciones, en general se los designa como promedios, y son de extraordinaria utilidad tanto en el análisis de una distribución, como en la comparación entre distribuciones.

1. Media aritmética

Es sin duda, el estadígrafo más utilizado, sobre todo en la cuantificación de variables económicas. Se simbolizará por $\overline{Y}$ o $M[Y_i]$, y se definirá como:

$$\overline{Y} = M[Y_i] = \frac{\sum_{i=1}^{m} Y_i \, n_i}{n}$$

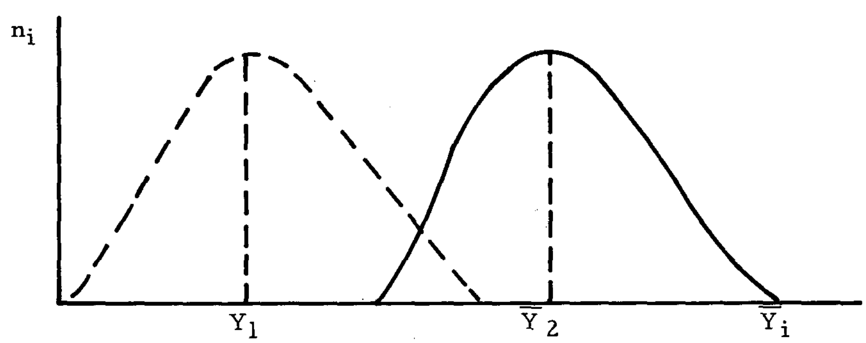
Puede observarse que a cada valor de la variable o marca de clase, se le atribuye una importancia o peso equivalente a la frecuencia absoluta correspondiente. Esta fórmula de cálculo es para datos agrupados en forma de una distribución de frecuencias. Cuando se desea calcular una media aritmética de datos no agrupados, todas las frecuencias absolutas serán iguales a la unidad, se simbolizará por $\overline{X}$ o $M[X_i]$ y se definirá como

$$\overline{X} = M[X_i] = \frac{\sum_{i=1}^{n} X_i}{n}$$

En general cuando el número de observaciones es relativamente grande conviene la agrupación de frecuencias en intervalos. Es evidente que cuando se resume un conjunto de datos en un número dado de intervalos, se pierde precisión; esta pérdida estará relacionada con la amplitud del intervalo, cuanto mayor sea ésta, menos preciso el cálculo. Por ello, en general, para un mismo conjunto de datos, la media aritmética obtenida de los datos originales, que es un cálculo exacto, diferirá de la obtenida de una tabla de distribución de frecuencias. La razón estriba en el supuesto de uniformidad de la distribución de frecuencias dentro de cada intervalo, supuesto que generalmente no se cumple; mas, en todo caso, esa pérdida de precisión está más que compensada por las ventajas que significa tener una tabla de frecuencias. Por lo demás en ciencias sociales, la precisión necesaria autoriza, dentro de ciertos límites, a desentenderse de una rigurosidad extrema.

En el gráfico que a continuación se presenta, se tienen dos distribuciones de frecuencias (supuesto un gran número de intervalos pequeños) muy similares, y sin embargo con medias aritméticas muy distintas.

Gráfico 3



La media aritmética como estadígrafo de tendencia central, indica la posición de la distribución. Sobre este estadígrafo cabe advertir su alta sensibilidad a valores extremos de la variable. Un valor muy alejado de los valores centrales, aunque poco representativo por ser único, puede hacer variar significativamente el promedio. Por ello, cuando se está utilizando este indicador en un análisis, vale la pena advertir la representatividad de los valores extremos, y la influencia que éstos tienen sobre el resultado. Muchas veces se concluye que es preferible estratificar previamente los datos originales en dos o tres categorías, realizando cálculos de medias aritméticas en forma separada para cada grupo.

a) Propiedades. Se presentarán las propiedades más importantes de la media aritmética

i) Primera propiedad: La suma de las desviaciones ponderadas de los valores de la variable respecto de la media aritmética es cero

$$\sum_{i=1}^{m} (Y_i - \bar{Y})\, n_i = 0$$

$$\sum_{i=1}^{m} Y_i\, n_i - n\bar{Y} = 0$$

$$n\,\bar{Y} - n\bar{Y} = 0$$

ii) Segunda propiedad: La suma de los cuadrados de las desviaciones ponderadas de los valores de la variable es un mínimo, cuando se toman respecto de la media aritmética. Se iniciará la demostración tomando desviaciones respecto a un valor cualquiera P, para luego concluir que P forzosamente tendrá que ser la media aritmética.

$$\sum_{i=1}^{m} (Y_i - P)^2\, n_i \quad \text{mínimo}$$

La derivada de esa expresión respecto de P, se iguala a cero.

$$-2 \sum_{i=1}^{m} (Y_i - P)\, n_i = 0$$

$$\sum_{i=1}^{m} Y_i\, n_i - nP = 0$$

$$P = \frac{\sum_{i=1}^{m} Y_i\, n_i}{n}$$

Es seguro que igualando la primera derivada a cero se obtiene un mínimo, porque se llega a un valor concreto. Para que fuera un máximo, P tendría que ser infinito.

iii) Tercera propiedad: La media aritmética de una variable más (menos) una constante es igual a la media de la variable más (menos) la constante.

$$M\,[Y_i \pm K] = M\,[Y_i] \pm K$$

$$\sum_{i=1}^{m} \frac{(Y_i \pm K)\, n_i}{n} = M\,[Y_i] \pm K$$

$$\sum_{i=1}^{m} \frac{Y_i\ n_i}{n} \pm \frac{n\ K}{n} = M\ [Y_i] \pm K$$

$$M\ [Y_i] \pm K = M[Y_i] \pm K$$

iv) Cuarta propiedad: La media aritmética de una variable multiplicada (dividida) por una constante, es igual a la constante que multiplica (divide) a la media de la variable.

$$M\ [Y_i\ K] = K\ M\ [Y_i]$$

$$\sum_{i=1}^{m} \frac{Y_i\ K\ n_i}{n} = K\ M\ [Y_i]$$

$$K \sum_{i=1}^{m} \frac{Y_i\ n_i}{n} = K\ M\ [Y_i]$$

$$K\ M\ [Y_i] = K\ M\ [Y_i]$$

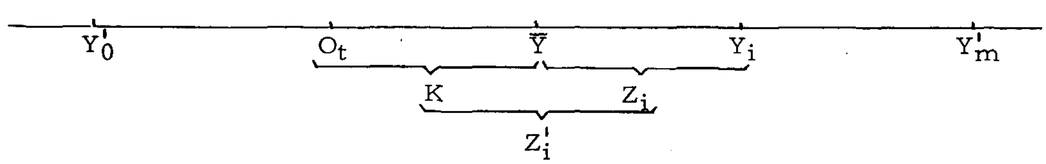
b) Métodos abreviados de cálculo. Se presentarán estos métodos no tanto por el ahorro de tiempo que puede significar su aplicación, como porque destacan algunos detalles sobre la media aritmética, y muestran procedimientos de trabajo con cambio de variable.

i) Primer método abreviado: Se trata de reducir la magnitud de la variable, en términos de desviaciones respecto de un origen de trabajo O_t arbitrariamente elegido. En cuanto a la elección de O_t vale la pena, para que el método sea realmente abreviado, tomar como origen de trabajo, un valor o marca de clase central de la distribución de frecuencias.

Se definirá esta variable reducida en la forma siguiente:

$$Z_i' = Y_i - O_t$$

Gráficamente se fijan los siguientes puntos dentro del recorrido de una variable:



Además, a las desviaciones respecto de la media aritmética, se simbolizará por Z_i, es decir:

$$Z_i = Y_i - \bar{Y}$$

La diferencia entre la media aritmética y el origen de trabajo arbitrariamente elegido se designará por K, es decir

$$K = \overline{Y} - O_t$$

Por este hecho, observando el gráfico, puede concluirse que

$$\overline{Y} = O_t + K$$

Se trata de encontrar una expresión para K, en función de desviaciones Z_i', para disponer de una fórmula de cálculo. En efecto,

$$Z_i' = Z + K \text{ (multiplicando por } n_i\text{)}$$

$$Z_i' n_i = Z_i n_i + K n_i \text{ (aplicando sumatoria)}$$

$$\sum_{i=1}^{m} Z_i' n_i = \sum_{i=1}^{m} Z_i n_i \quad nK$$

Recuérdese que en la primera propiedad de la media aritmética se demostró que:

$$\Sigma Z_i n_i = \Sigma (Y_i - \overline{Y}) n_i = 0$$

Luego

$$K = \frac{\Sigma Z_i' n_i}{n}$$

La fórmula de cálculo abreviado será

$$\overline{Y} = O_t + \frac{\Sigma Z_i' n_i}{n}$$

En la siguiente tabla de distribución de frecuencias se aplica este método:

Intervalo	Marca de clase	Frecuencia	Desviaciones	Desviaciones ponderadas
$Y_{i-1}' - Y_i'$	Y_i	n_i	$Z_i' = Y_i - O_t$ *	$Z_i' \; n_i$
0 - 8	4	8	-19	-152
8 - 20	14	10	-9	-90
20 - 26	23	30	0	0
26 - 30	28	9	5	45
30 - 40	35	3	12	36
		60		-161

* Se elige $O_t = 23$

$$\overline{Y} = 23 - \frac{161}{60} = 20.32$$

ii) Segundo método abreviado. Este método en general sólo se aplica con ventaja, cuando es constante la amplitud de los intervalos. Como en el anterior, se trata de trabajar en términos de desviaciones, pero además en este método dichas desviaciones se expresan en unidades de intervalo (dividiéndolas por C)

Esta nueva variable es en consecuencia,

$$Z''_i = \frac{Y_i - O_t}{C} = \frac{Z'_i}{C}$$

de donde

$$Z'_i = C Z''_i$$

En la fórmula del primer método se reemplaza Z'_i y se tiene

$$\overline{Y} = O_t + C \frac{\sum_{i=1}^{m} Z''_i n_i}{n}$$

Cabe destacar que este método permite obtener Z''_i en forma totalmente mecánica; basta fijar el origen de trabajo para completar todos los valores de Z''_i.

En el siguiente ejemplo podrá apreciarse las ventajas de esta forma de cálculo.

Intervalo	Marca de clase	Frecuencia	Desviaciones	Desviaciones ponderadas
$Y'_{i-1} - Y'_i$	Y_i	n_i	Z''_i	$Z''_i n_i$
2 - 6	4	20	-3	-60
6 - 10	8	40	-2	-80
10 - 14	12	50	-1	-50
14 - 18	16 *	90	0	0
18 - 22	20	60	1	60
22 - 26	24	40	2	80
		300		-50

* Eligiendo $O_t = 16$

Aplicando la fórmula del segundo método abreviado se tiene

$$\overline{Y} = O_t + C \frac{\sum_{i=1}^{m} Z''_i n_i}{n}$$

$$\overline{Y} = 16 - 4 \frac{50}{300} = 15.33$$

Obsérvese que una vez fijado el origen de trabajo las Z''_i se colocan en sucesión decreciente para las marcas de clase menores que O_t y en sucesión creciente para las marcas de clase mayores.

2. Mediana (M_e)

Se trata de otro estadígrafo de tendencia central de aplicación muy frecuente. Se define como el valor de la variable que supera a no más de la mitad de las observaciones y es superado por no más de la mitad de dichas observaciones. Es un estadígrafo menos sensible que la media aritmética ante valores extremos de la variable, y puede ser calculado aún en variables de tipo ordinal.

Cuando las observaciones no están agrupadas en forma de una tabla de distribución de frecuencias, su cómputo es en extremo sencillo. Basta disponer los valores en orden creciente y ubicar el valor central. Por ejemplo, supóngase que se tienen ordenados los siguientes valores de gastos en consumo de 7 familias (en dólares por mes).

40 - 47 - 60 - 70 - 78 - 80 - 90

La mediana será 70 dólares ya que este valor supera a 3 observaciones (40, 47 y 60) que no son más que la mitad (la mitad es 3.5) y a su vez es superada por 3 observaciones (78, 80 y 90) que tampoco son más de la mitad.

Cuando el número de observaciones es par, existen dos valores centrales que satisfacen la definición de mediana. Si en el ejemplo anterior se agrega una familia adicional se tiene:

40 - 47 - 60 - 70 - 78 - 80 - 90 - 180

En este caso tanto 70 como 78 son valores medianos. La mitad de las observaciones es 4 y 70 supera a 3 que no son más de la mitad y es superado por 4 que tampoco es más de la mitad, pues es exactamente la mitad. Igual cosa ocurre con 78; para evitar ambigüedades se toma como mediana en estos casos el punto medio entre los dos valores medianos. En el ejemplo, la mediana definitiva sería 74 dólares. Obsérvese la poca sensibilidad de este estadígrafo a los valores extremos. Al agregar la 8a. observación con un valor bastante más alto que el resto, la mediana ha experimentado apenas un ligero crecimiento; más todavía aunque en vez de 180 aquel valor hubiese sido de 5 000, la mediana siempre habría sido 74. En cambio no ocurre lo mismo para la media aritmética, que es sumamente sensible a ese tipo de valores extremos; intente el lector su cálculo en uno y otro ejemplo, y verificará una fuerte variabilidad.

Si los datos se agrupan en una tabla de frecuencias, el cálculo de la mediana implica computar previamente las frecuencias acumuladas.

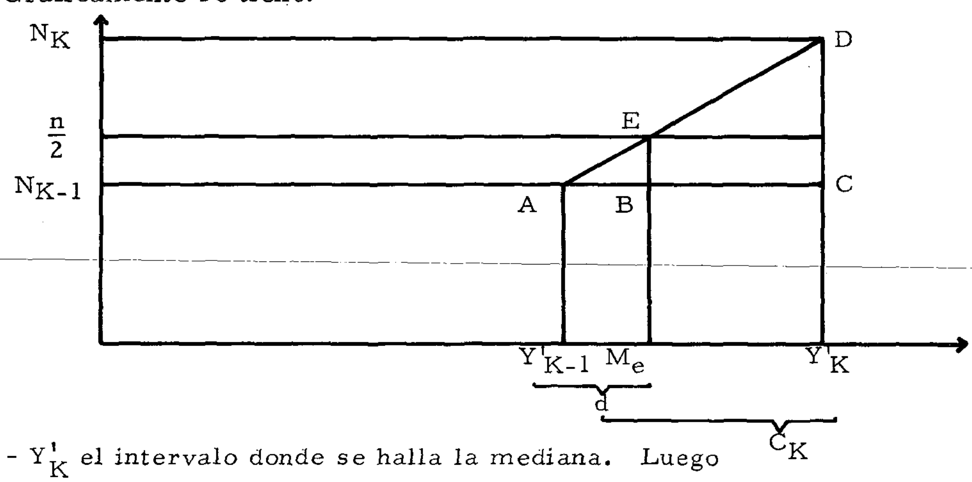
a) variable discreta, en este caso bastará con identificar la frecuencia acumulada que es inmediatamente mayor a la mitad de las observaciones. La mediana será aquel valor de la variable que corresponda a dicha frecuencia acumulada. Ejemplo:

Número de predios por persona Y_i	Número de propietarios n_i	Frecuencia acumulada N_i
1	200	200
2	160	360
3	150	510
4	100	610
5	80	690
6	40	730
7 y más	20	750
	750	

Siendo $\frac{n}{2}$ = 375, la menor frecuencia acumulada que supera este valor es 510, que corresponde al valor 3 de la variable, siendo éste el valor mediano. Dicho valor supera a 360 observaciones que no son más de la mitad y es superado por 260 que tampoco son más de la mitad, satisfaciendo la definición de mediana.

b) variable continua, en este caso el problema consiste en determinar un punto dentro del intervalo en que está comprendida la mediana. La identificación del intervalo donde se halla la mediana, es exactamente igual al caso de variable discreta: el intervalo será aquél que corresponda a la frecuencia acumulada inmediatamente superior a la mitad de las observaciones. Como se dijo, la preocupación consiste en fijar un punto dentro de ese intervalo, que corresponde a la mediana. Para ello se adoptará el supuesto que las observaciones se distribuyen linealmente dentro del mencionado intervalo.

Gráficamente se tiene:



Sea $Y'_{K-1} - Y'_K$ el intervalo donde se halla la mediana. Luego

$$M_e = Y'_{K-1} + d.$$

Será necesario encontrar una expresión para "d" en función de frecuencias que son los valores conocidos que se dispone y por los cuales está determinado este estadígrafo.

Por semejanza de triángulos se tiene que

$$\frac{AB}{BE} = \frac{AC}{CD}$$

Pero

$$AB = d$$

$$BE = \frac{n}{2} - N_{K-1}$$

$$AC = C_K \text{ (amplitud del intervalo K-ésimo)}$$

$$CD = N_K - N_{K-1} = n_K$$

Reemplazando, se tiene

$$\frac{d}{\frac{n}{2} - N_{K-1}} = \frac{C_K}{n_K}$$

de donde

$$d = C_K \frac{\frac{n}{2} - N_{K-1}}{n_K}$$

Luego

$$M_e = Y'_{K-1} + C_K \frac{(\frac{n}{2} - N_{K-1})}{n_K}$$

Ejemplo: La siguiente tabla muestra la distribución de los coeficientes producto-capital de 310 empresas industriales:

Coeficientes Y'_{i-1} - Y'_i	Empresas n_i	Frecuencias acumuladas N_i
0.15 - 0.20	40	40
0.20 - 0.30	80	120
0.30 - 0.42	100	220
0.42 - 0.50	60	280
0.50 - 0.70	30	310
	310	

$$\frac{n}{2} = \frac{310}{2} = 155$$

La menor frecuencia acumulada que supera a 155 es $N_K = 220$. Luego, $n_K = 100$,

$$Y'_{K-1} = 0.30 \text{ , } C_K = 0.12 \quad y \quad N_{K-1} = 120 \quad \text{(donde } K = 3\text{)}$$

Reemplazando estos valores en la fórmula

$$M_e = 0.30 + 0.12 \frac{155 - 120}{100} = 0.30 + 0.036 = 0.336$$

Como una extensión de este estadígrafo, será fácil ampliar el concepto a otros indicadores que dividen la masa de informaciones en otras proporciones y no sólo en mitades como lo hace la mediana.

Se tiene el caso de los cuartiles que dividen las observaciones en cuartas partes; así, el primer cuartil Q_1 es un valor de la variable que supera a no más de un cuarto de las observaciones, y es superado por no más de tres cuartos de ellas. Para identificar el intervalo donde se halla Q_1, habrá que determinar la frecuencia acumulada inmediatamente superior a $\frac{n}{4}$. El cálculo es similar al de la mediana:

$$Q_1 = Y'_{K-1} + C_K \frac{(\frac{n}{4} - N_{K-1})}{n_K}$$

Para el tercer cuartil, sucede otro tanto

$$Q_3 = Y'_{K-1} + C_K \frac{(3\frac{n}{4} - N_{K-1})}{n_K}$$

Para identificar el intervalo que comprende a Q_3, habrá que averiguar cuál es la frecuencia acumulada N_K, inmediatamente superior a $3\frac{n}{4}$; no es necesario detallar el segundo cuartil, porque coincide con la mediana. En forma similar se pueden encontrar estadígrafos que dividan al total de observaciones en décimas partes (deciles), en centésimas partes (percentiles), etc. Las fórmulas correspondientes pueden deducirse por analogía con las de los cuartiles.

Así, el 7° decil estará dado por

$$D_7 = Y'_{K-1} + C_K \frac{\frac{7n}{10} - N_{K-1}}{n_K}$$

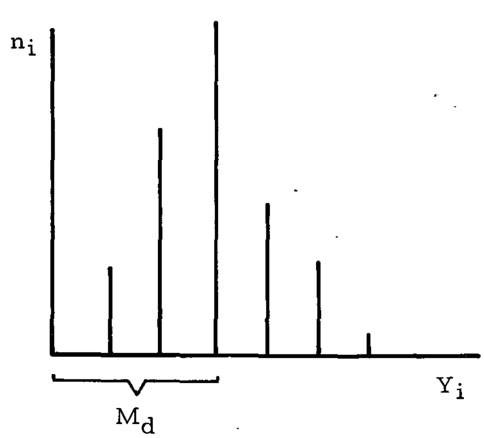
El 35 percentil estará dado por

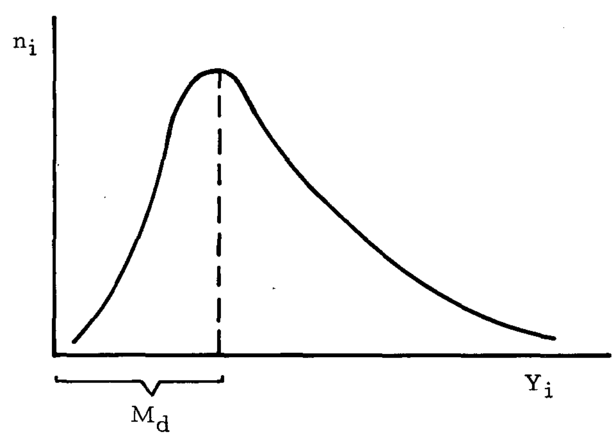
$$P_{35} = Y'_{K-1} + C_K \frac{\frac{35n}{100} - N_{K-1}}{n_K}$$

---

3. Moda o valor modal

Se trata de otro estadígrafo de tendencia central. Tiene un significado bastante preciso y es de extraordinaria utilidad, aunque inexplicablemente poco utilizado en estudios socioeconómicos. Se simbolizará por M_d y se definirá como aquel valor de la variable al que corresponde la máxima frecuencia. Es muy corriente que se confunda el valor modal con la frecuencia máxima; recuérdese que es un valor de la variable y por lo mismo se le representa en el eje de las abscisas. Está dado por la frecuencia máxima, pero no se trata de una frecuencia.





a) variable discreta, una vez agrupados los datos, es posible determinar inmediatamente el valor modal; bastará con fijar el valor de la variable que más se repite.

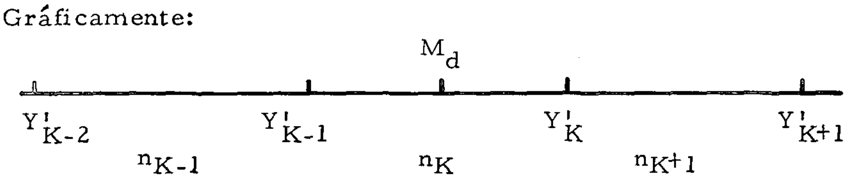
Ejemplo:

Número de cargas familiares Y_i	Número de familias n_i
0	80
1	120
2	210
3	380
4	180
5	60
6 o más	40
	1 070

La frecuencia máxima es 380, que corresponde al cuarto valor de la variable. El valor modal en consecuencia, es 3; este valor modal será tanto más representativo cuanto mayor sea la frecuencia máxima. Se presentarán algunos casos donde el valor modal pierde significación; es el caso donde hay varios valores de las variables que tienen frecuencias similares. Es así como se califica a las distribuciones de unimodales, bimodales, multimodales, etc.

b) variable continua, de la misma manera que en el cálculo de la mediana, primero es necesario determinar el intervalo donde se halla comprendido el valor modal; en este caso bastará ver cuál es el intervalo que tiene la frecuencia máxima. El paso siguiente es determinar un punto dentro de ese intervalo. Existen algunos criterios, un tanto arbitrarios, para deducir fórmulas del valor modal. Uno de esos criterios toma en cuenta la magnitud de las frecuencias de los intervalos contiguos (mayor y menor) al que comprende el valor modal; en otras palabras, dividirá al intervalo en partes inversamente proporcionales a las frecuencias de los intervalos contiguos.

Gráficamente:



$$\frac{M_d - Y'_{K-1}}{Y'_K - M_d} = \frac{n_{K+1}}{n_{K-1}}$$

La moda estará más cerca del intervalo contiguo que tenga mayor frecuencia.

Despejando M_d de la relación anterior, se tiene:

$$M_d\, n_{K-1} - Y'_{K-1}\, n_{K-1} = Y'_K\, n_{K+1} - M_d\, n_{K+1}$$

pero,

$$Y'_K = Y'_{K-1} + C_K$$

$$M_d\,[n_{K-1} + n_{K+1}] = Y'_{K-1}\,[n_{K+1} + n_{K-1}] + C_K\, n_{K+1}$$

$$M_d = Y'_{K-1} + \frac{C_K \quad n_{K+1}}{n_{K-1} + n_{K+1}}$$

Es necesario destacar que la deducción anterior, no toma en cuenta la amplitud de los intervalos contiguos; en caso de amplitudes muy diferentes, puede distorsionarse el valor de este estadígrafo.

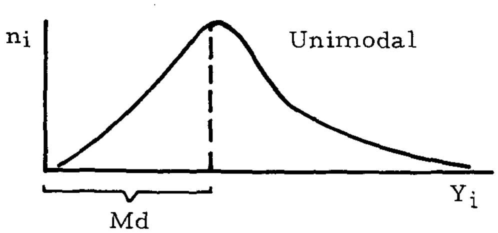
Ejemplo:

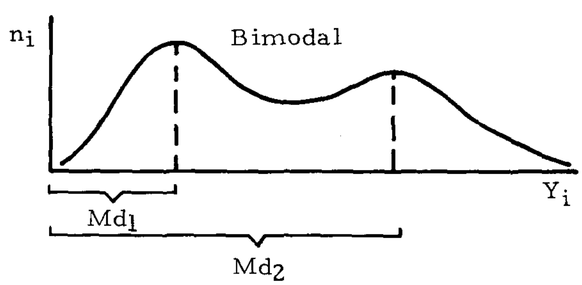
Ingresos de profesionales Y'_{i-1} - Y'_i	Profesionales n_i
0 - 20	25
20 - 40	45
40 - 60	80
60 - 80	60
80 - 100	40
100 - 120	15
120 y más	5

Inmediatamente se puede adelantar que la moda se encontrará en el tercer intervalo: 40 - 60. Aplicando la fórmula.

$$M_d = 40 + \frac{20\ (60)}{45 + 60} = 40 + \frac{1\ 200}{105} \approx 51.43$$

Si se supone una gran cantidad de intervalos pequeños para una cierta distribución, gráficamente el valor modal estaría así representado:





Este estadígrafo al igual que la mediana, puede determinarse para variables cualitativas, ya que basta con encontrar la frecuencia máxima. El siguiente ejemplo ilustra esta posibilidad:

Color de automóviles preferido por los clientes	Clientes encuestados
Blanco	18
Azul	22
Verde	40
Amarillo	25
Rojo	75

El "valor" modal, en este caso, es el rojo, ya que por tener la máxima frecuencia, es el preferido por los clientes.

Al iniciar el estudio de este estadígrafo, se decía que inexplicablemente era poco utilizado en los análisis. Es evidente que requiere, para su cómputo, más información que la media aritmética; ello podría explicar en forma parcial su poco uso, pero aun cuando se dispone de información, muchas veces se cree suficiente calcular por ejemplo un ingreso por habitante y quedarse con un análisis parcial. Sin duda, para caracterizar adecuadamente una distribución, se requiere una serie de estadígrafos cuyas indicaciones se complementen. Por ejemplo, saber que dos países tienen ingresos por habitante de 150 y 200 dólares al año, puede permitir cierto tipo de conclusiones, pero saber además que los valores modales de los ingresos anuales por habitante, son de 140 y 130 dólares respectivamente, permite obtener conclusiones bastante más objetivas. Naturalmente, son necesarios muchos otros antecedentes e indicadores que se presentarán a continuación, para realizar análisis más completos. Interesa destacar la necesidad de buscar un conjunto de indicadores que haga posible el análisis de las distintas facetas de un fenómeno sometido a estudio.

4. Media geométrica

Este estadígrafo se define como la raiz de orden n del producto de los n valores de la variable.

Cuando los datos no están agrupados, su fórmula de cálculo es

$$M_g = \sqrt[n]{X_1 \ X_2 \ X_3 \ldots\ldots\ldots\ldots\ldots X_n} = \sqrt[n]{\prod_{i=1}^{n} X_i}$$

Para fines prácticos es preferible calcular el logaritmo de la media geométrica y luego el antilogaritmo de ésta.

$$\log M_g = \frac{1}{n} \sum_{i=1}^{n} \log X_i$$

Si los datos aparecen agrupados, es decir, si las marcas de clase tienen frecuencias superiores a la unidad, se tendrá la siguiente fórmula.

$$M_g = \sqrt{\overbrace{Y_1 \ldots\ldots Y_1}^{n_1} \; \overbrace{Y_2 \ldots\ldots Y_2}^{n_2} \; \overbrace{Y_m \ldots\ldots Y_m}^{n_m}}$$

$$= \sqrt{Y_1^{n_1} \, Y_2^{n_2} \ldots\ldots Y_m^{n_m}}$$

$$M_g = \sqrt[n]{\prod_{i=1}^{m} Y_1 \, n_1}$$

$$\log M_g = \frac{1}{n} \sum_{i=1}^{m} (\log Y_i) \, n_i$$

El estadígrafo que se estudia, aparte del inconveniente que significa el engorro de su cálculo, está además limitado por el hecho que los valores de la variable deben ser positivos para que pueda ser interpretado. Si algún valor de la variable es cero, la media geométrica será cero; igualmente si aparece algún valor negativo el estadígrafo toma un valor imaginario. Pese a estos inconvenientes, para cierto tipo de variables, en especial las cronológicas, que sigan una tendencia exponencial, se hace indispensable su uso si se desea calcular valores intermedios, es decir, si se desea interpolar no linealmente. Por ejemplo, si cierta población en 1940 era de 2.5 millones y en 1960 alcanza a 4 millones, para calcular la población en 1955, sería indispensable el empleo de la media geométrica, si se admite el crecimiento exponencial de la población a una tasa constante.

$$P_{1950} = \sqrt{(2.5) \quad (4.0)} \doteq 3.15 \text{ millones}$$

$$P_{1955} = \sqrt{(3.15) \quad (4.0)} \doteq 3.55 \text{ millones}$$

5. Media armónica

El último de los estadígrafos de tendencia central que aquí se abordará, se define como el recíproco de la media aritmética de los valores recíprocos de la variable.

Para datos no agrupados:

$$M_h = \frac{1}{\dfrac{\Sigma \frac{1}{x_i}}{n}} = \frac{n}{\Sigma \frac{1}{x_i}}$$

Para datos agrupados:

$$\frac{n}{\Sigma \frac{n_i}{Y_i}}$$

Ejemplo: Un grupo de trabajadores construyen los primeros 120 metros de una avenida con una productividad de 12 metros diarios, en cambio los siguientes 120 metros lo hacen a razón de 18 metros por día. Se trata de determinar la productividad diaria durante todo el trabajo.

Si se decidiera calcular la media aritmética se tendría:

$$\bar{Y} = \frac{12 + 18}{2} = 15 \text{ metros diarios}$$

Por otra parte los primeros 120 metros requieren 10 días y los siguientes 120 metros 6.67 días; es decir, todo el trabajo lo harían en 16.67 días. Si la productividad diaria es de 15 metros, en los 16.67 días construirían un total de 250.05 metros, lo que es inconsistente, ya que el trabajo total es de sólo 240 metros. Si en cambio se utiliza la media armónica.

$$M_h = \frac{2}{\frac{1}{12} + \frac{1}{18}} = \frac{72}{3 + 2} = \frac{72}{5} = 14.4 \text{ metros}$$

Trabajando con una productividad media de 14.4 metros por día, en 16.67 días, se construirán 240 metros.

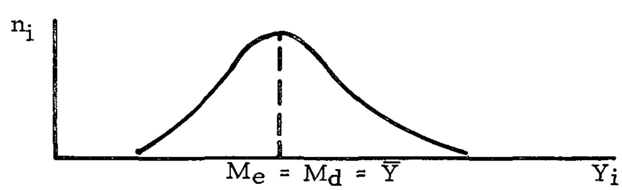
Como pudo advertirse, la media armónica se aplica cuando se presenta una relación inversa entre las variables implícitas; en el caso del ejemplo la relación inversa aparece entre la productividad y el tiempo.

e: espacio
p: productividad $\qquad e = p \times t \qquad p = e \cdot \frac{1}{t}$
t: tiempo

C. Evaluación de los estadígrafos de tendencia central

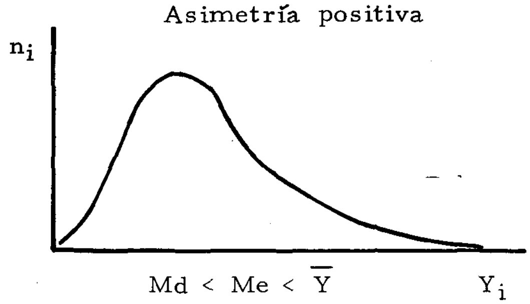
En más de una oportunidad se insistió sobre la necesidad de disponer de un conjunto de indicadores sobre la variable que se está estudiando. Los indicadores presentados, que en general se denominan de tendencia central, tienen definiciones precisas; por ello muestran aspectos particulares del fenómeno que se estudia. Se trata de un conjunto de estadígrafos complementarios; las conclusiones a que en último término den lugar, deberán ser producto de la consideración simultánea de los valores que alcanzan dichos indicadores.

Al analizar la bondad de cada uno de estos indicadores, es preciso tener presente el volumen de observaciones tomadas en cuenta para su cálculo y las limitaciones de cada uno de ellos. Siempre es conveniente complementar el análisis de las cifras, con una representación gráfica de la distribución de frecuencias de la variable; e interesa destacar la posición relativa de la media aritmética, la mediana y el valor modal. La posición relativa de estos estadígrafos depende de la forma de la distribución; de esta suerte, si la distribución es simétrica, es decir, si se observa perfecta simetría respecto de un eje central, los tres estadígrafos coinciden.

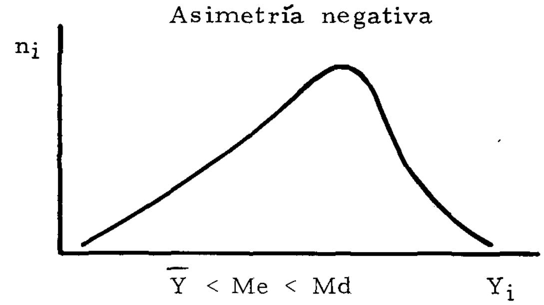


En el caso de distribuciones no simétricas, la posición relativa de los estadígrafos depende del tipo de asimetría. De esta manera, si la asimetría es positiva, es decir, si la distribución tiene su rama o manto más extendido hacia valores positivos de la variable, la moda será menor que la media aritmética. La mediana, por el hecho de dividir la masa de observaciones en dos partes, quedará comprendida entre ambas. Si la asimetría es negativa, es decir, cuando la distribución se extienda suavemente hacia valores negativos de la variable, la moda superará a la media aritmética, permaneciendo la mediana, y por la misma razón dada en el otro caso, comprendida entre ambos indicadores. Gráficamente

Asimetría positiva



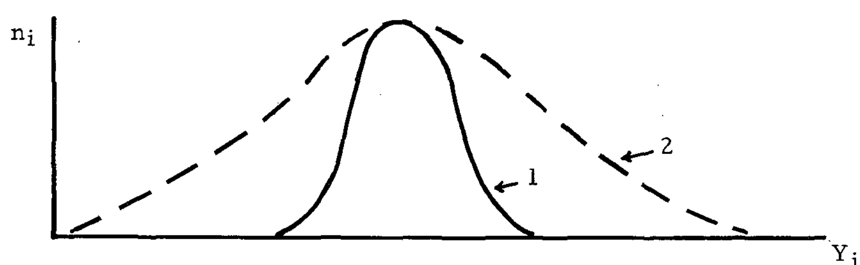
Asimetría negativa



Recuérdese que la media aritmética es un estadígrafo muy sensible a valores extremos de la variable, de allí que en un caso sea el mayor de los tres estadígrafos y en otro caso el menor de ellos. La moda, como el valor de la variable que más se repite, tiene en general una clara ubicación. Habría que agregar que su valor depende sobremanera de la amplitud de intervalo elegida y su representatividad sólo se garantiza cuando existe una clara concentración de frecuencias en un intervalo dado.

D. Estadígrafos de dispersión

Una vez caracterizada la distribución a través de estadígrafos de tendencia central y conociendo el tipo de asimetría, interesa tener indicaciones acerca del grado de heterogeneidad con que la variable se distribuye en un conjunto de observaciones. Dos distribuciones pueden tener iguales estadígrafos de tendencia central, sin embargo pueden mostrar grados de dispersión diferentes, como puede observarse en el gráfico que a continuación se muestra.



Evidentemente en la primera distribución (línea continua) los valores aparecen más concentrados en torno al eje central, en tanto que en la otra aparecen mucho más dispersos. Si ambas distribuciones representaran ingresos de dos poblaciones, se concluiría que en la primera distribución los ingresos son más homogéneos, mientras que en la segunda se observaría gran disparidad entre ingresos altos, medios y bajos.

Parecería innecesario destacar la importancia que tiene contar con indicadores que pudieran mostrar este tipo de características en una distribución; sobre todo en lo que se refiere a distribución de ingresos, ahora de tanta actualidad, es indispensable contar con indicaciones adecuadas en este sentido.

1. Recorrido de la variable

Cuando se aborda el problema de la dispersión, lo primero que se piensa es el campo de recorrido de la variable: la diferencia entre el mayor y menor valor de ella. Si bien brinda una primera idea acerca de la heterogeneidad, tiene el inconveniente que sólo toma en cuenta los dos valores extremos, descuidando el conjunto de valores intermedios. Puede suceder que uno de los valores extremos esté accidentalmente desplazado y no constituya por tanto un valor representativo; en este caso el recorrido sería exagerado y la dispersión aparecería distorsionada. Para iniciar el análisis es conveniente considerar el recorrido, pero en ningún caso es suficiente.

Para variable discreta $R = Y_m - Y_o$

Para variable continua $R = Y'_m - Y'_o$

2. Recorrido intercuartílico

Como una manera de subsanar el inconveniente de los valores extremos que presentaba el estadígrafo anterior, se define un nuevo indicador, que toma en cuenta el recorrido entre el primer y tercer cuartil.

$$D_q = Q_3 - Q_1$$

Si bien es cierto que este indicador representa un adelanto respecto del anterior, no lo es menos que siempre toma dos valores de la variable, dejando de lado el resto, y en consecuencia la influencia de valores extremos puede, aunque en menor medida, originar algún tipo de deformación en cuanto al grado de dispersión.

3. Varianza

Se define este estadígrafo en virtud de la propiedad de la media aritmética que minimiza la suma de las desviaciones al cuadrado. Se simbolizará por σ^2 ó $V[Y_i]$

a) Para datos no agrupados

$$V[X_i] = \sigma^2 = \frac{\Sigma (X_i - \overline{X})^2}{n} = \frac{\Sigma X_i^2}{n} - \overline{X}^2$$

b) Para datos agrupados

$$V[Y_i] = \sigma^2 = \frac{\Sigma Y_i - \overline{Y})^2 n_i}{n} = \frac{\Sigma Y_i^2 n_i}{n} - \overline{Y}^2$$

Si bien la varianza no tiene un fin per se sino que se utiliza en materias que se presentarán posteriormente, da origen a un estadígrafo que sí tiene utilidad e interpretación práctica. Se trata de la desviación típica o estándar que se define como la raíz cuadrada positiva de la varianza.

$$\sigma = +\sqrt{\sigma^2}$$

Mientras más dispersa sea la variable, mayor será la magnitud de la desviación típica puesto que mayores serán los desvíos respecto de la media aritmética, sin posibilidad de compensación de desvíos por tratarse de suma de cuadrados. Este estadígrafo se expresa en las mismas unidades de la variable estudiada, en tanto que la varianza se expresa en el cuadrado de la unidad de medida.

Como puede observarse en la fórmula, este indicador de dispersión toma en cuenta todos los valores de la variable con sus correspondientes frecuencias o pesos relativos, sin embargo, siempre es sensible a valores extremos. Por ello es conveniente, antes de calcular los estadígrafos, hacer un análisis previo de la tabla de distribución de frecuencias, para percibir la representatividad de valores extremos y sus posibles efectos sobre los valores de los estadígrafos.

c) Propiedades de la varianza

i) Primera propiedad: La varianza de una variable a la cual se le suma (resta) una constante, es igual a la varianza de la variable original.

$V[Y_i \pm K] = V[Y_i]$ Aplicando la definición de varianza a la variable $Y_i + K$, se tiene:

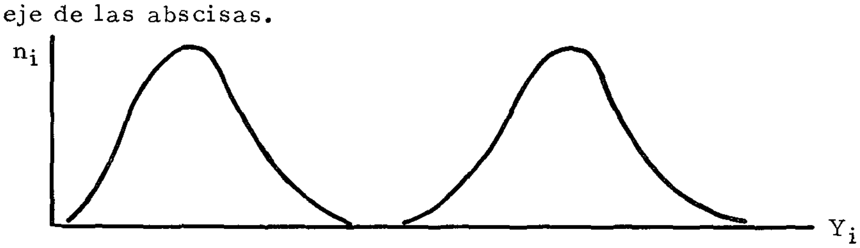
$$\frac{\sum_{i=1}^{m} (Y_i \pm K - M[Y_i + K])^2}{n} n_i = V[Y_i]$$

pero se vió que $M[Y_i + K] = K + M[Y_i]$

$$\frac{\sum_{i=1}^{m} (Y_i \pm K - M[Y_i] \pm K)^2}{n} n_i = V[Y_i]$$

$$\frac{\sum_{i=1}^{m} (Y_i - \overline{Y})^2 n_i}{n} = V[Y_i]$$

Gráficamente, las dos distribuciones tienen la misma varianza, pese a estar desplazadas en el eje de las abscisas.



ii) Segunda propiedad: La varianza del producto de una constante por una variable, es igual al cuadrado de la constante por la varianza de la variable.

$$V[K\ Y_i] = K^2\ V[Y_i]$$

$$\frac{\sum_{i=1}^{m} (K\ Y_i - K\ \overline{Y})^2 n_i}{n} = K^2\ V[Y_i]$$

$$\frac{\sum_{i=1}^{m}[K^2(Y_i-\overline{Y})^2]n_i}{n} = K^2\ V[Y_i]$$

$$\frac{K^2\sum_{i=1}^{m}(Y_i-\overline{Y})^2\ n_i}{n} = K^2\ V[Y_i]$$

$$K^2\ V[Y_i] = K^2\ V[Y_i]$$

d) Componentes de la varianza

En el caso que un conjunto de datos haya sido dividido previamente en grandes categorías o estratos, es posible desglosar la varianza en dos componentes muy útiles para el análisis. Admítase que una masa de datos ha sido dividida en L estratos; cada estrato tendrá una media aritmética, una varianza y un número de observaciones que expresa la importancia de cada uno de estos estratos. En este caso la variabilidad total puede deberse tanto a variabilidad dentro de cada estrato como a variabilidad entre los diferentes estratos.

i) Intervarianza: Estadígrafo que representa la variabilidad entre los estratos; se define como la varianza entre las medias de los estratos.

$$\sigma_b^2 = V[\overline{Y}_h] = \frac{\sum_{h=1}^{L}(\overline{Y}_h-\overline{Y})^2\ nh}{n}$$

donde $\overline{Y}_h$ es la media aritmética del estrato h
$\overline{Y}$ es la media aritmética general
n_h es el número de observaciones o tamaño de cada estrato

ii) Intravarianza: Estadígrafo que representa la variabilidad dentro de los estratos; se define como el promedio de las varianzas de los estratos.

$$\sigma_w^2 = M\ [\sigma_h^2] = \frac{\sum_{h=1}^{L}\sigma_h^2\ n_h}{n}$$

donde σ_h^2 es la varianza del estrato h, y n_h y n obedecen a las mismas definiciones del caso anterior.

Dado que los dos estadígrafos estudiados son partes componentes de la varianza, a continuación se presenta la correspondiente demostración.

$$\sigma^2 = \sigma_b^2 + \sigma_w^2$$

$$\frac{\sum_{h=1}^{L}\sum_{i=1}^{nh}(Y_{hi}-\overline{Y})^2}{n} = \frac{\sum_{h=1}^{L}(\overline{Y}_h-\overline{Y})^2\ n_h}{n} + \frac{\sum_{h=1}^{L}\sigma_h^2\ nh}{n}$$

pero $\sigma_h^2 = \dfrac{\Sigma (Y_{hi} - \bar{Y}_h)^2}{n_h}$; reemplazando

$$\sum_{h=1}^{L} \sum_{i=1}^{n_h} (Y_{hi} - \bar{Y})^2 = \sum_{h=1}^{L} (\bar{Y}_h - \bar{Y})^2 n_h + \sum_{h=1}^{L} \sum_{t=1}^{n_h} \frac{(Y_{hi} - \bar{Y}_h)^2 n_h}{n_h}$$

Elevando al cuadrado los correspondientes binomios

$$\sum_{h=1}^{L} \sum_{i=1}^{n_h} (Y_{hi}^2 - 2\bar{Y} Y_{hi} + \bar{Y}^2) = \sum_{h=1}^{L} (\bar{Y}_h^2 - 2\bar{Y} \bar{Y}_h + \bar{Y}^2) n_h +$$

$$+ \sum_{h=1}^{L} \sum_{i=1}^{n_h} (Y_{hi}^2 - 2\bar{Y}_h Y_{hi} + \bar{Y}_h^2)$$

Aplicando las propiedades de la sumatoria

$$\sum_{h=1}^{L} [\sum_{i=1}^{n_h} (Y_{hi}^2) - 2\bar{Y} n_h \bar{Y}_h + n_h \bar{Y}^2)] = \sum_{h=1}^{L} \bar{Y}_h^2 n_h - n Y^2 + \sum_{h=1}^{L} \sum_{i=1}^{n_h} [(Y_{hi}^2) - n_h \bar{Y}_n^2]$$

$$\sum_{h=1}^{L} \sum_{i=1}^{n_h} Y_{hi}^2 - n \bar{Y}^2 = \sum_{h=1}^{L} \bar{Y}_h^2 n_h - n \bar{Y}^2 + \sum_{h=1}^{L} \sum_{i=1}^{n_h} Y_{hi}^2 - \sum_{h=1}^{L} n_h Y_h^2$$

Simplificando términos queda

$$\sum_{h=1}^{L} \sum_{i=1}^{n_h} Y_{hi}^2 = \sum_{h=1}^{L} \sum_{t=1}^{n_h} Y_{hi}^2$$

Luego $\sigma^2 = \sigma_b^2 + \sigma_w^2$

Ejemplo: Piénsese en los sueldos y salarios pagados por una fábrica, que tienen una varianza de 137 600. Si las observaciones se clasifican por estratos: obreros, empleados administrativos, y técnicos, será posible analizar más a fondo la distribución de ingresos.

Las informaciones disponibles serían:

Estratos (h)	Tamaño estrato n_h	Media por estrato $\bar{Y}_h$	Varianza por estrato σ_h^2
Obreros	300	400	160 000
Empleados administrativos	100	400	160 000
Técnicos	100	500	40 000

La media aritmética general será:

$$\bar{Y} = \frac{\sum_{h=1}^{L} \bar{Y}_h n_h}{n} = \frac{2\ 100}{5} = 420$$

$$\sigma_w^2 = \frac{\sum_{h=1}^{L} \sigma_h^2 \ n_h}{n} = \frac{680\ 000}{5} = 136\ 000$$

$$\sigma_b^2 = \frac{\sum_{h=1}^{L} (\bar{Y}_n - \bar{Y})^2 n_h}{n} = \frac{8\ 000}{5} = 1\ 600$$

El cálculo de los anteriores estadígrafos permite concluir que la variabilidad se debe principalmente a heterogeneidad en las remuneraciones dentro de los estratos y no así a diferencias entre estratos; en otros términos las remuneraciones promedio de cada estrato, son bastante homogéneas ya que la intervarianza es pequeña, mientras que las remuneraciones dentro de cada estrato son muy heterogéneas puesto que la intravarianza es bastante grande. (Ver Anexo.)

e) Métodos abreviados de cálculo.

1) Primer método abreviado. Se trata de encontrar una fórmula que reduzca el volumen de operaciones; como en el caso de la media aritmética, la variable se expresará en términos de desvíos respecto de un origen de trabajo. Recuérdese que:

$$\sigma^2 = \frac{\sum_{i=1}^{m} (Y_i - \bar{Y})^2 n_i}{n} = \frac{\sum_{i=1}^{m} Z_i^2 n_i}{n}$$

ya que $Z_i = Y_i - \bar{Y}$

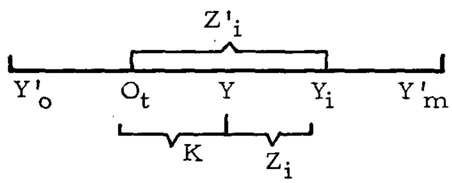
Si se desarrolla el cuadrado del binomio dentro de la sumatoria, se tiene:

$$\sigma^2 = \frac{\sum_{i=1}^{m} Y_i^2 n_i - 2\bar{Y} \sum_{i=1}^{m} Y_i n_i + n \bar{Y}^2}{n}$$

pero $\sum_{i=1}^{m} Y_i n_i = n\bar{Y}$, luego,

$$\sigma^2 = \frac{\sum_{i=1}^{m} Y_i^2 n_i}{n} - \bar{Y}^2$$

Observando el siguiente gráfico, resulta inmediata la deducción de la fórmula abreviada:



$$Z_i' = Z_i + K \qquad \text{Elevando al cuadrado}$$

$$Z_i'^2 = Z_i^2 + 2K\,Z_i + K^2 \qquad \text{Ponderando por } n_i$$

$$Z_i'^2\,n_i = Z_i^2\,n_i + 2K\,Z_i\,n_i + K^2\,n_i \qquad \text{Aplicando sumatoria}$$

$$\sum_{i=1}^{m} Z_i'^2\,n_i = \sum_{i=1}^{m} Z_i^2\,n_i + 2K\sum_{i=1}^{m} Z_i\,n_i + n\,K^2$$

Pero la primera propiedad de la media aritmética decía:

$$\sum_{i=1}^{m} Z_i\,n_i = 0 \;, \quad \text{luego}$$

$$\sum_{i=1}^{m} Z_i'^2\,n_i = \sum_{i=1}^{m} Z_i^2\,n_i + n\,K^2$$

pero

$$\sum_{i=1}^{m} Z_i^2\,n_i = n\,\sigma^2 \quad \text{y}$$

$$nK = \sum_{i=1}^{m} Z_i'\,n_i \quad \text{(en el cálculo abreviado de la media aritmética)}$$

Luego

$$\sigma^2 = \frac{\sum_{i=1}^{m} Z_i'^2\,n_i}{n} - \left(\frac{\sum_{i=1}^{m} Z_i'\,n_i}{n}\right)^2$$

ii) Segundo método abreviado. Consiste, como se recordará, en expresar las desviaciones en términos de unidades de intervalo (dividiendo por la amplitud del intervalo).

$$\frac{Y_i - O_t}{c} = \frac{Z_i'}{c} = Z_i''$$

es decir que $Z_i' = c\ Z_i''$ y reemplazando en la fórmula anterior se obtiene:

$$\sigma^2 = c^2 \left\{ \frac{\sum_{i=1}^{m} Z_i''^2\ n_i}{n} - \left(\frac{\Sigma Z_i''\ n_i}{n}\right)^2 \right\}$$

En el siguiente ejemplo se podrá comprobar el ahorro de tiempo que significa la aplicación de estas fórmulas.

En el caso del primer método abreviado:

Impuestos por contribuyente			Número de contribuyentes			
Y_{i-1}'	Y_i'	Y_i	n_i	Z_i'	$Z_i' n_i$	$Z_i'^2 n_i$
0	20	10	20	-40	-800	32 000
20	40	30	15	-20	-300	6 000
40	60	50*	10	0	0	0
60	80	70	8	20	160	3 200
80	100	90	5	40	200	8 000
			58		-740	49 200

* $O_t = 50.$

Reemplazando estos valores en la fórmula, se obtiene:

$$\sigma^2 = \frac{49\ 200}{58} - \left(\frac{-740}{58}\right) = 848.3 - 162.8 = 685.5$$

En el caso del segundo método abreviado se tiene:

Y_{i-1}'	Y_i'	Y_i	n_i	Z_i''	$Z_i''\ n_i$	$Z_i''^2 n_i$
0	20	10	20	-2	-40	80
20	40	30	15	-1	-15	15
40	60	50*	10	0	0	0
60	80	70	8	1	8	8
80	100	90	5	2	10	20
			58		-37	123

* $O_t = 50.$

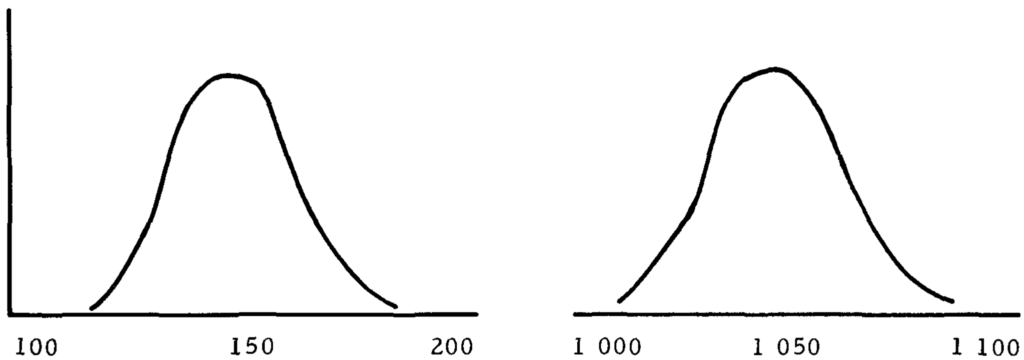
Aplicando la fórmula respectiva:

$$\sigma^2 = 20^2 \left\{ \frac{123}{58} - \left(\frac{-37}{58}\right)^2 \right\} = 400\{2.12 - 0.407\} = 685.5$$

4. Coeficiente de variabilidad

Tanto la varianza como la desviación típica tienen el inconveniente de los estadígrafos absolutos, ya que en el caso de indicaciones sobre dispersión, tiene mucha importancia no tomar en cuenta la posición de la distribución. Estos estadígrafos, sobre todo al comparar distribuciones, pueden deformar las conclusiones.

Obsérvese las dos distribuciones que aparecen a continuación.



Ambas distribuciones muestran la misma dispersión en torno a la media, es decir, tienen igual varianza y desviación típica; sin embargo, en términos relativos, una distribución donde el menor ingreso es 1 000 y el mayor es 1 100, es mucho más homogénea que otra distribución donde el menor ingreso es 100 y el mayor 200. En un caso la diferencia entre el mayor y menor ingreso es 10%, mientras que en el otro es de 100%.

Surge, por consiguiente, la necesidad de disponer de un estadígrafo que tome en cuenta la tendencia central de la distribución. Se define así el coeficiente de variabilidad, como también la razón entre la desviación típica y la media aritmética.

$$C.V = \frac{\sigma}{\bar{Y}}$$

En el ejemplo anterior, si ambas distribuciones tuvieran, por ejemplo, una desviación típica de 60, los coeficientes de variabilidad serían:

$$C.V_1 = \frac{\sigma 1}{\bar{Y}_1} = \frac{60}{150} = 0.4 = 40\%$$

$$C.V_2 = \frac{\sigma 2}{\bar{Y}_2} = \frac{60}{1\ 050} = 0.057 = 5.7\%$$

Estos estadígrafos permiten llegar a conclusiones más realistas y ciertas.

El coeficiente de variabilidad, o desviación típica relativa como también se le llama, puede tomar valores tan grandes como se quiera, ya que no hay una relación de dependencia limitante entre σ y $\bar{Y}$. Por otra parte, en el caso de una distribución donde la media aritmética fuera negativa no tiene sentido considerar el signo para calificar la dispersión. Por ello este estadígrafo podría definirse como el valor absoluto del cociente entre la desviación típica y la media aritmética.

Puesto que las propiedades de la media aritmética y la desviación típica ya fueron analizadas, las propiedades del coeficiente de variabilidad serán el resultado de las propiedades de los indicadores componentes.

Si bien es cierto que para calificar la dispersión de una distribución es más apropiado el coeficiente de variabilidad, de esto no debe deducirse que la varianza y la desviación típica carecen de utilidad; por el contrario, son muy útiles en el tratamiento de materias que se estudiarán posteriormente.

E. Utilización de indicadores de la programación

Con mucha frecuencia se escuchan quejas respecto de la escasez de informaciones estadísticas básicas para los países en vías de desarrollo. Admitida la deficiencia, es conveniente también reconocer que no siempre se hace un uso óptimo de la escasa información disponible. Si se reconoce como etapa primaria durante un proceso de planificación la posibilidad de realizar diagnósticos, será necesario destacar la enorme utilidad de la estadística descriptiva en lo que se refiere a caracterización de fenómenos en un momento dado. Por una parte, el diagnóstico requiere disponer de una perspectiva histórica sobre las variables estratégicas; por otra, esa misma perspectiva histórica debe complementarse con análisis cuantitativos en profundidad durante períodos que presenten cambios de orientación y/o ritmo en las tendencias observadas, aparte de una cuantificación detallada para el momento "cero" de un plan. Es para esos puntos que el instrumental de estadística descriptiva debe ser puesto a disposición del analista. Se ha insistido sobre la urgente necesidad de contar con un juego de indicadores para las principales variables; cada indicador mostrará una faceta de la variable estudiada, y un conjunto de ellos permitirá una adecuada e integral calificación de las variables que interesan para el diagnóstico.

En general, un conjunto de indicadores permite realizar análisis de consistencia; de ese modo puede llegarse a una primera evaluación acerca de la calidad y fidelidad de la información que se pretende utilizar.

TEMAS DE DISCUSION

Indique si las siguientes afirmaciones son verdaderas o falsas y justifique su opinión.

1. En toda distribución con media nula el percentil 38 es igual, en valor absoluto al percentil 62.

2. Las siguientes fórmulas dan resultados exactamente iguales para la media aritmética.

$$\frac{\sum_{i=1}^{n} X_i}{n}, \quad \frac{\sum_{i=1}^{m} Y_i n_i}{n}, \quad O_t + c \quad \frac{\sum Z_i'' n_i}{n}$$

3. La media armónica de una constante es igual al recíproco de la constante.

4. Los siguientes datos son consistentes.

$h_2 = 0.4 \qquad h_1 = 0.2 \qquad H_3 = 0.8 \qquad n = 50 \qquad n_4 = 5$

$\bar{Y} = 25 \qquad M_d = 30$

5. Para que la intervarianza fuese igual a la varianza, las medias de los estratos deberían ser iguales.

6. En una distribución asimétrica, el valor de la mediana coincide con el del segundo cuartil.

7. En una distribución normal tipificada el sexto percentil es igual a - 0.15.

8. Los siguientes datos son consistentes:

$$\sigma b^2 = 780$$
$$\bar{Y} = 150$$
$$CV = 20\%$$
$$\sigma_w^2 = 110$$

9. En una distribución cualquiera, se debe verificar que:

$$\bar{Y} + \sigma \leqslant Y'_m$$

10. Los siguientes datos son consistentes:

$m = 6 \qquad h_1 = 0.2 \qquad h_4 = 0.2$

$H_2 = 0.6 \qquad H_3 + H_4 = 1.9$

11. El número de intervalos en que se clasifica una masa de datos, depende de la cantidad de éstos.

12. La moda es un mejor indicador que la media aritmética, porque es poco sensible a valores extremos.

13. En una distribución normal, donde la mediana es 8 y la media cuadrática 10, el coeficiente de variabilidad será inferior al 50%.

14. Los siguientes datos son consistentes:

$$M_g = 0 \qquad\qquad M_d = -10$$

15. La mediana de una distribución es 50. Si se multiplican los valores de la variable por 1.6, el valor de la mediana será 80.

16. Los siguientes datos son consistentes para una población dividida en dos estratos de igual tamaño.

$$\bar{Y}_1 = 10 \qquad\qquad \bar{Y}_2 = 10$$

$$\sigma_1 = 8 \qquad\qquad \sigma_2 = 10$$

$\sigma^2 = 81$ (varianza para toda la población)

17. Si las edades de los alumnos siguen una distribución normal con media 28 y desviación típica 2, la proporción de alumnos menores a 23 años es del orden de 10 por ciento.

18. En el sector servicios el sueldo promedio es de 200 u.m. Si los varones constituyen el 70 por ciento de la población remunerada, es factible que su ingreso promedio mensual sea de 300 u.m.

19. Dada una población con una cierta varianza, la magnitud de la intervarianza y de la intravarianza depende del criterio de estratificación.

20. El valor modal está condicionado por el número de intervalos en que se tabule una masa de datos.

21. Los siguientes datos son consistentes en una distribución simétrica.

Recorrido intercuartílico:	80
Percentil N° 80:	200
Media aritmética:	160

22. La media geométrica de los valores de unavariable multiplicados por una constante, es igual a la media geométrica de la variable original.

23. Los siguientes datos son consistentes:

$$m = 5 \qquad \sum_{i=1}^{5} Y_i^2 n_i = 400 \qquad \overline{Y} = 5$$

$$h_5 = 0.1 \qquad H_3 = 0.5 \qquad n_4 = 8$$

24. El valor modal carece de adecuada representatividad, cuando las frecuencias de los intervalos contiguos al que contiene a este estadígrafo, representan fracciones pequeñas del total de observaciones.

25. Si la varianza de los ingresos de los obreros de una industria es 400, y la intravarianza y la intervarianza son iguales, estos componentes no alteran su participación relativa dentro de la varianza total si se reajusta a todos los obreros en un 20%.

26. La desviación típica de las utilidades reinvertidas de las sociedades anónimas de cierto país, es de 25 000 unidades monetarias. Si solamente se consideran las sociedades anónimas industriales, su desviación típica no podrá exceder de 25 000 unidades monetarias.

27. Al comparar las regiones A y B, se tiene que las desviaciones típicas de los ingresos familiares son de 600 y 450 unidades monetarias respectivamente. Se puede concluir en consecuencia, que en la región B el ingreso está más uniformemente distribuido.

PROBLEMAS PROPUESTOS

1. Elabore una distribución de frecuencias de 7 intervalos, de modo que el coeficiente de variabilidad sea de 80%.

2. El sector asalariado de una región ha sido dividido en dos estratos: obreros y empleados. Se trataba de analizar los efectos de una política de redistribución de ingresos. Los datos disponibles, antes y después de aplicar tal política, fueron los siguientes:

	Antes	Después
Coeficiente de variabilidad general	60%	50%
$\overline{Y}$ (obreros)	15	?
$\overline{Y}$ (empleados)	?	30
Proporción de obreros	0.6	0.6
σ_b^2	150	24
σ^2	225	144

Se le pide enumerar qué conclusiones le merece la política de redistribución tanto a nivel general como a nivel de estrato. Justifique sus opiniones, deduciendo aquellos estadígrafos que le parezcan pertinentes.

3. Compare el grado de heterogeneidad en los salarios de los obreros de la construcción en dos países.

Venezuela	Argentina
σ = 50 bolívares	σ = 50 000 pesos
Salario por habitante anual: 200 dólares	Salario por habitante anual: 180 dólares
Tipo de cambio: 5 bolívares por dólar	Tipo de cambio: 350 pesos por dólar

¿Qué podría concluir en cuanto a la distribución de ingresos en ambos países?

4. En cierta comunidad el impuesto total recaudado a 100 000 contribuyentes durante el año 1966 fue de 13 millones de unidades monetarias. El coeficiente de variabilidad de los tributos individuales es de 120%, se sabe que hay 60 000 contribuyentes obreros que cancelaron en conjunto 3 millones de unidades monetarias. Calcule la intravarianza de esa población estratificada en obreros y no obreros, y comente brevemente la homogeneidad de la distribución de los tributos.

5. El sector agrícola de un país se divide en dos estratos: cultivos intensivos y cultivos extensivos; respecto de los ingresos de los trabajadores del sector, se tienen los siguientes indicadores:

$$\bar{Y} = 100 \text{ u.m.} \qquad CV = 50\% \qquad \sigma_b^2 = 400$$

Se decide reajustar los ingresos de los trabajadores en 20% y darles además una bonificación de 30 u.m. a cada uno de ellos. Analice qué cambios en la distribución del ingreso provoca el reajuste.

EJERCICIOS

1. La inversión real anual de un grupo de industrias pesqueras se detalla a continuación:

Miles de dólares

10 - 12 - 8 - 40 - 6 - 8 - 10 - 30 - 2 - 8 - 6 - 14
16 - 20 - 25 - 28 - 30 - 26 - 30 - 4 - 6 - 10 - 18 - 17
13 - 17 - 21 - 7 - 6 - 8 - 14 - 7 - 15 - 19 - 27 - 22
0 - 14 - 6 - 8 - 9 - 11 - 13 - 15 - 18 - 20 - 30 - 60
12 - 6 - 5 - 5 - 6 - 8 - 7 - 12 - 15 - 36 - 39 - 52

Se pide:

a) Forme una tabla de distribución de frecuencias, con ocho intervalos de amplitud constante;

b) Calcule las frecuencias relativas;

c) Calcule las frecuencias absolutas y relativas acumuladas;

d) Represente gráficamente la distribución de frecuencias (histograma y polígono de frecuencias);

e) Calcule los siguientes estadígrafos:

i) Media aritmética (datos originales y métodos abreviados).
ii) Mediana;
iii) Moda;
iv) Media armónica.

2. Las distribuciones de ingreso de dos países son las siguientes:

País A

Ingresos anuales por hab.: dólares	Población remunerada
80 - 100	30 000
100 - 120	80 000
120 - 140	40 000
140 - 160	10 000
160 - 200	4 000
200 - 220	1 000

País B

Ingresos anuales por hab.: dólares	Población remunerada
60 - 90	10 000
90 - 120	20 000
120 - 150	50 000
150 - 180	20 000
180 - 210	15 000
210 - 240	10 000
240 - 270	4 000

a) Fundamente, empleando el cálculo de estadígrafos que crea conveniente, el grado de desarrollo comparado de ambos países. ¿Necesita además otros antecedentes para calificar a estos países con mayor rigor?

b) Calcule el ingreso promedio del 30% de la población de mayores ingresos en cada país. Compare los resultados y discuta sus conclusiones. Realice el mismo cálculo para el 40% de la población de menores ingresos.

c) ¿Cuál es el ingreso promedio de los dos países en conjunto?

d) Suponiendo que el país B devalúa su moneda en 30% y el país A en 5%, calcule los nuevos ingresos promedio en ambos países.

3. El ingreso por habitante de un país es de 310 dólares al año; su sector obrero, que constituye el 59 por ciento de la población, percibe 1/5 del ingreso total. Calcule el ingreso por habitante de este sector.

4. Un automovilista se dirige de Santiago a Talca, ciudades que distan 300 km entre sí; los primeros 200 km los recorre a una velocidad de 120 km/hora y el resto a 80 km/hora. Utilizando un estadígrafo de tendencia central, calcule la velocidad media.

5. La media aritmética entre dos números es 8 y su media geométrica 2. Calcule la media armónica.

6. Si la población de un país es al 31 de diciembre de 1950 de 5 800 000 personas, y al 31 de diciembre de 1960 de 7 200 000, calcule la población en 1955.

7. Si se tiene una distribución de frecuencias simétrica, con 6 intervalos de amplitud constante, y los siguientes datos:

$n = 150$ $\quad Y'_5 = 60$ $\quad n_2 = n_1 + 5$

$n_3 = 30$ $\quad Q_1 = 43,5$

calcule el 6° decil

8. La siguiente tabla de distribución de frecuencias representa los impuestos personales de un conjunto de profesionales.

Impuestos (u. m.) Y'_{i-1}	Y'_i	Profesionales n_i (miles)
0	20	30
20	40	25
40	60	15
60	80	13
80	100	12
100	120	5

Se le pide que:

a) Calcule la varianza por los tres métodos que conoce:

b) Calcule el coeficiente de variabilidad;

c) Si se cobra un impuesto adicional de 5 u.m. por persona, ¿cuál es la desviación típica?

d) Si se reajustan los impuestos por persona en 20%, ¿cuál es la varianza?

9. El C.V. de los ingresos de 200 empleados de una compañía es 57 por ciento. Después de reajustar, según ley, todos los sueldos en E° 11, este C.V. es ahora de 50 por ciento. Sin embargo, la gerencia fija un sueldo mínimo de E° 71. Antes del reajuste, había 35 personas que tenían un sueldo medio de E° 40 y todos ellos ganaban menos de E° 60; con la nueva política de la gerencia, sus sueldos serán elevados a E° 71. Determine la cantidad de dinero que necesitará mensualmente la compañía, para pagar los sueldos después de hacer efectivos los reajustes.

10. En una distribución de frecuencias se multiplican los valores de la variable por 3, y se obtiene una media aritmética de 54; si se suma 5 a los valores de la variable, se obtiene una media cuadrática de 24. Calcule el coeficiente de variabilidad de la distribución.

11. Se clasificó a los trabajadores de un mineral en dos categorías: mayores y menores de 25 años; y se extrajo la siguiente información:

	Número de obreros n_h	Productividad media $\bar{Y}_h$	Desviación típica σ_h
Mayores de 25 años	200	40	70
Menores de 25 años	300	60	40

Calcule la varianza de todos los obreros del mineral.

12. La media aritmética y la desviación típica de los porcentajes de reinversión de utilidades de las empresas constructoras son de 40% y 20% respectivamente. Si se supone que esta variable sigue aproximadamente una distribución normal, * se le pide que

a) Calcule cuál es el porcentaje de empresas que reinvierten más del 70%;
b) Calcule el porcentaje de empresas que reinvierten entre 20 y 50%;
c) Calcule cuál es la proporción de empresas que reinvierten cuando mucho 35% de sus utilidades.

13. Se sabe que los consumos familiares en bienes esenciales, están relacionados con los consumos familiares de energía eléctrica mediante la función:

$$C.Es. = 4.5\ C.En. - 5$$

Si la media aritmética y la desviación típica de los consumos de energía se estiman en E° 40 y E° 15 respectivamente, calcule el coeficiente de variabilidad de los consumos esenciales.

* Véase Anexo.

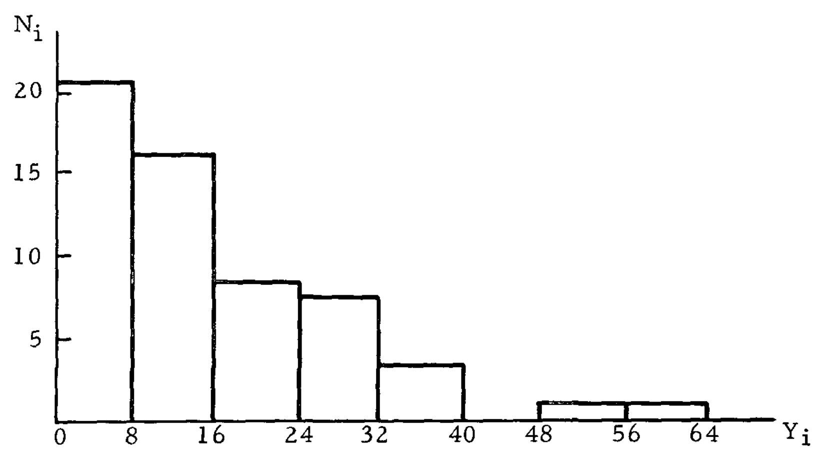
SOLUCION DE EJERCICIOS

1. Dado que el menor valor es 0, el máximo 60, y se especifican 8 intervalos de amplitud constante, se clasifican los datos en la siguiente forma; quedan así resueltas las partes a, b y c.

Y'_{i-1} - Y'_i	n_i	h_i	$H_i\downarrow$	$N_i\downarrow$	$H_i\uparrow$	$N_i\uparrow$
0 - 8	21	21/60	21/60	21	60/60	60
8.1 - 16	17	17/60	38/60	38	39/60	39
16.1 - 24	9	9/60	47/60	47	22/60	22
24.1 - 32	8	8/60	55/60	55	13/60	13
32.1 - 40	3	3/60	58/60	58	5/60	5
40.1 - 48	0	0	58/60	58	2/60	2
48.1 - 56	1	1/60	59/60	59	2/60	2
56.1 - 64	1	1/60	60/60	60	1/60	1
	60	1				

Obsérvese que la amplitud del primer intervalo, es ligeramente superior a la del resto.

d) La representación gráfica sería la siguiente:



e) Los valores de los estadígrafos serían los siguientes:

i) Media aritmética; existen varias alternativas

- datos originales

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{967}{60} = 16.12$$

- Marcas de clase

$$\bar{Y} = \frac{\sum_{i=1}^{m} Y_i n_i}{n} = \sum_{i=1}^{m} Y_i h_i = \frac{912}{60} = 15.2$$

Y_i	h_i	$Y_i h_i$
4	21/60	84/60
12	17/60	204/60
20	9/60	180/60
28	8/60	224/60
36	3/60	108/60
44	0	0
52	1/60	52/60
60	1/60	60/60
		912/60

- Métodos abreviados

Haciendo $O_t = 20$

Y_i	n_i	Z_i'	$Z_i' n_i$	Z_i''	$Z_i'' n_i$
4	21	-16	-336	-2	-42
12	17	-8	-136	-1	-17
20	9	0	0	0	0
28	8	8	64	1	8
36	3	16	48	2	6
44	0	24	0	3	0
52	1	32	32	4	4
60	1	40	40	5	5
	60		-288		-36

Por el primer método abreviado

$$\bar{Y} = O_t + \frac{\sum Z_i' n_i}{n} = 20 + \left(\frac{-288}{60}\right) = 20 - 4.8 = 15.2$$

Por el segundo método abreviado

$$\bar{Y} = O_t + c\, \frac{\sum Z_i'' n_i}{n} = 20 + 8 \left(\frac{-36}{60}\right) = 20 - 4.8 = 15.2$$

ii) Mediana:

$$M_e = Y_{k-1}' + C_k \; \frac{\frac{n}{2} - N_{k-1}}{n_k}$$

$\frac{n}{2} = \frac{60}{2} = 30.$ La menor frecuencia acumulada que supera este valor es $N_2 = 38$; luego en el segundo intervalo estará comprendida la mediana. (Véase tabla 1. 1.)

$$M_e = 8 + 8 \frac{30 - 21}{17} = 8 + \frac{72}{17} = 12.23$$

iii) Moda: Puesto que la frecuencia máxima corresponde al primer intervalo, se presenta un caso particular.

$$M_d = Y'_{k-1} + \frac{C_k \, n_{k+1}}{n_{k-1} + n_{k+1}} = 0 + \frac{8 \cdot 17}{17} = 8$$

Nótese que este estadígrafo adquiere el valor del límite superior del primer intervalo, porque no existe intervalo inferior.

iv) Media armónica

Y_i	4.00	12.00	20.00	28.00	36.00	44.00	52.00	60.00
n_i	21.00	17.00	9.00	8.00	3.00	0.00	1.00	1.00
n_i/Y_i	5.25	1.41	0.45	0.28	0.08	0.00	0.02	0.01

$$M_h = \frac{n}{\sum_{i=1}^{m} \frac{n_i}{Y_i}} = \frac{60}{7.50} = 8$$

2. a) Un primer indicador es el ingreso promedio (la población se expresa en miles de personas).

País A

$Y'_{i-1} - Y'_i$	Y_i	n_i	$Y_i \, n_i$
80 - 100	90	30	2 700
100 - 120	110	80	8 800
120 - 140	130	40	5 200
140 - 160	150	10	1 500
160 - 200	180	4	720
200 - 220	210	1	210
		165	19 130

País B

$Y'_{i-1} - Y'_i$	Y_i	n_i	Z''_i	$Z''_i \, n_i$
60 - 90	75	10	-2	-20
90 - 120	105	20	-1	-20
120 - 150	135	50	0	0
150 - 180	165	20	1	20
180 - 210	195	15	2	30
210 - 240	225	10	3	30
240 - 270	255	4	4	16
		129		56

$$\bar{Y}_A = \frac{\Sigma Y_i \, n_i}{n} = \frac{19\ 130}{165} = 115.94$$

$$\overline{Y}_B = O_t + C \frac{\Sigma Z''_i n_i}{n} = 135 + 30 \quad \frac{56}{129} \quad = 148.02$$

Un primer indicio que el país B tendría mayor grado de desarrollo, está dado por el mayor valor de su ingreso por habitante; sin embargo, es necesario calcular otros estadígrafos y conocer otro tipo de informaciones, para calificar aproximadamente ambas situaciones.

El cálculo de los valores modales ya indica que la diferencia entre ambos países, no parece tan importante como lo indicarían los ingresos promedios.

$$M_{d_A} = Y'_{k-1} + \frac{C_k \, n_{k+1}}{n_{k-1} + n_{k+1}} = 100 + \frac{20 \cdot 40}{70} = 111.4$$

$M_{d_B} = 135$ Corresponderá con la marca de clase que tiene la frecuencia máxima, puesto que los intervalos contiguos tienen igual frecuencia. No se puede pretender resumir una situación tan compleja en un par de estadígrafos; pues cada indicador constituye un análisis parcial de aquello que es susceptible de cuantificación.

b) Se trata de acumular las frecuencias hacia arriba, hasta completar el 30% de n. Como las frecuencias no coinciden perfectamente con esta cifra, se hace necesario dividir el intervalo de manera que la frecuencia acumulada $N_i \uparrow$ sea el 30% de n.

País A

$$n = 165$$
$$0.30\, n = 49.5$$

La frecuencia absoluta deberá ser 34.5 para satisfacer las condiciones del problema; 34.5 representa el 86.25% de 40 y en esta proporción se divide el intervalo, quedando

Y'_{i-1} - Y'_i	n_i	Y_i	$Y_i\, n_i$
122.75 - 140	34.5	131.38	4 532
140.00 - 160	10.0	150.00	1 500
160.00 - 200	4.0	180.00	720
200.00 - 220	1.0	210.00	210
	49.5		6 962

$$\overline{Y}_A = \frac{6\,962}{49.5} = 140.6$$

País B

Siguiendo el mismo proceso se llega a la siguiente tabla.

Y'_{i-1} - Y'_i	n_i	Y_i	$Y_i\, n_i$
165.5 - 180	9.7	172.7	1 675.2
180.0 - 210	15.0	195.0	2 925.0
210.0 - 240	10.0	225.0	2 250.0
240.0 - 270	4.0	255.0	1 020.0
	38.7		7 870.2

$$\bar{Y}_B = \frac{7\,870.2}{38.7} = 203.4$$

Para el 40% de la población de ingresos menores, el procedimiento es similar, se llega a las siguientes tablas estimadas.

País A

Y'_{i-1} - Y'_i	n_i	Y_i	$Y_i\, n_i$
80 - 100	30	90.0	2 700
100 - 109	36	104.5	3 762
	66		6 462

$$\bar{Y}_A = \frac{6\,462}{66} = 97.9$$

País B

Y'_{i-1} - Y'_i	n_i	Y_i	$Y_i\, n_i$
60 - 90	10.0	75.0	750.0
90 - 120	20.0	105.0	2 100.0
120 - 133	21.6	126.5	2 732.4
	51.6		5 582.4

$$\bar{Y}_B = \frac{5\,582.4}{51.6} = 108.2$$

c) $\bar{Y}\ (\text{Conjunto}) = \dfrac{n_A \bar{Y}_A + n_B \bar{Y}_B}{n_A + n_B} = \dfrac{(165)(115.9) + (129)(148.0)}{294} = 130$

d) La devaluación implica una tasa de cambio más alta, en moneda del país, por dólar. Luego, para el país A:

$$M\left[\frac{Y_i}{1.30}\right] = \frac{1}{1.30}\, M[Y_i] = \frac{1}{1.30}\, 115.94 = 89.2$$

Para el país B

$$M\left[\frac{Y_i}{1.05}\right] = \frac{1}{1.05} M[Y_i] = \frac{1}{1.05} \cdot 148.02 = 141.0$$

3. Se sabe que

$$\frac{\text{Ingreso total}}{\text{Población}} = \text{Ingreso por habitante}$$

$$\frac{Y}{P} = 310$$

Para el sector obrero, el ingreso por habitante será

$$\overline{Y}\text{ (obrero)} = \frac{1/5\ Y}{0.59\ P}$$

pero $Y = 310 \cdot P$, luego

$$\overline{Y}\text{ (obrero)} = \frac{1/5\ (310 \cdot P)}{0.59 \cdot P} = \frac{62}{0.59} = 105.08$$

4. El estadígrafo indicado es la media armónica

$$M_h = \frac{n}{\Sigma \frac{n_i}{Y_i}} = \frac{300}{\frac{200}{120} + \frac{100}{80}} = \frac{720}{7} = 103$$

5. Sean x e y los números:

$$\frac{x+y}{2} = 8$$

$$\sqrt{xy} = 2$$

$$M_h \frac{2}{\frac{1}{x} + \frac{1}{y}} = \frac{2}{\frac{x+y}{xy}} = \frac{2xy}{x+y} = \frac{xy}{\frac{x+y}{2}} = \frac{4}{8} = \frac{1}{2}$$

6. Si se supone un crecimiento exponencial a una tasa constante, es factible aplicar la media geométrica, que dará la población a mitad del período.

$$M_g = \sqrt{5.8 \cdot 7.2} = \sqrt{41.76} = 6.459$$

7. Los datos disponibles, pueden disponerse así en una tabla de frecuencias.

Y'_{i-1}	Y'_i	n_i 1/	N_i	Estadígrafos
Y'_0	Y'_1	n_1	N_1	
Y'_1	Y'_2	$n_1 + 5$	N_2	Q_1=43.5
Y'_2	Y'_3	30	75	
Y'_3	Y'_4	30	N_4	
Y'_4	60	$n_6 + 5$	N_5	
60	Y'_6	$\frac{6}{150}$	150	

Recuérdese que la distribución es simétrica y, por lo tanto,

$$n_1 = n_6 \text{ ; } n_2 = n_5 \text{ ; } n_3 = n_4 = 30$$

Se sabe que

$$n_1 + n_2 + n_3 = 75$$

$$n_1 + n_1 + 5 + 30 = 75$$

$$n_1 = 20 \qquad \text{luego}$$

$$n_2 = 25$$

Se trata de encontrar el valor de la amplitud del intervalo C_k. Para ello

$$Q_1 = Y'_1 + C_k \frac{\frac{n}{4} - N_{k-1}}{n_k}$$

Para determinar en qué intervalo se halla Q_1, debe verse cuál es la menor frecuencia acumulada que supera a $\frac{n}{4} = 37.5$. Se comprueba que es $N_2 = 45$; luego Q_1 estará dentro del segundo intervalo:

$$Q_1 = Y'_1 + C_k \frac{37.5 - 20}{25}$$

Pero

$$Y'_5 = Y'_1 + 4\ C_k = 60$$

Luego $Y'_1 = 60 - 4\ C_k$ y reemplazando en la fórmula de Q_1, se tiene:

$$43.5 = 60 - 4\ C_k + C_k \frac{17.5}{25} = 60 - 4\ C_k + 0.7\ C_k$$

$3.3\ C_k = 60 - 43.5 = 16.5$

$$C_k = \frac{16.5}{3.3} = 5$$

La distribución quedará así:

Y'_{i-1}	Y'_i	n_i	N_i
35	40	20	20
40	45	25	45
45	50	30	75
50	55	30	105
55	60	25	130
60	65	20	150

Una vez completada la tabla, se calcula inmediatamente el 6° decil.

$$D_6 = Y'_{k-1} + C_k \frac{\frac{6n}{10} - N_{k-1}}{n_k}$$

Para determinar k, será necesario calcular la menor frecuencia acumulada que supere a $\frac{6n}{10} = 90$. Se comprueba que es N_4.

$$D_6 = Y'_3 + 5\ \frac{90 - 75}{30} = 50 + \frac{75}{30} = 52.5$$

8. a) Una fórmula para el cálculo de la varianza es

$$\sigma^2 = \frac{\Sigma Y_i^2 n_i}{n} - \overline{Y}^2$$

Y_i	n_i	$Y_i n_i$	$Y_i^2 n_i$
10	30	300	3 000
30	25	750	22 500
50	15	750	37 500
70	13	910	63 700
90	12	1 080	97 200
110	5	550	60 500
		4 340	284 400

$$\overline{Y} = \frac{4\ 340}{100} = 43.4$$

$$\sigma^2 = \frac{284\ 400}{100} - (43.4)^2 = 2\ 844 - 1\ 884 \doteq 960$$

Utilizando métodos abreviados con $O_t = 50$ se tiene

Z'_i	$Z'_i n_i$	$Z'^2_i n_i$	Z''_i	$Z''_i n_i$	$Z''^2_i n_i$
-40	-1 200	48 000	-2	-60	120
-20	-500	10 000	-1	-25	25
0	0	0	0	0	0
20	260	5 200	1	13	13
40	480	19 200	2	24	48
60	300	18 000	3	15	45
	-660	100 400		-33	251

$$\sigma^2 = \frac{\Sigma Z'^2_i n_i}{n} - \left(\frac{\Sigma Z'_i n_i}{n}\right)^2 = \frac{100\ 400}{100} - 43.6 \doteq 960$$

También

$$\sigma^2 = C^2 \left\{ \frac{\Sigma Z''^2_i n_i}{n} - \left(\frac{\Sigma Z''_i n_i}{n}\right)^2 \right\} = 400 \left\{ \frac{251}{100} - 0.11 \right\} \doteq 960$$

b) El C. V. es

$$C.V. = \frac{\sigma}{\bar{Y}} = \frac{\sqrt{960}}{43.4} = \frac{31}{43.4} = 71.4\%$$

c) Recuérdese que

$$V[Y_i + K] = V[Y_i]$$

$$V[Y_i + 5] = V[Y_i] = 960$$

Luego,

$$\sigma = 31$$

d) Recuérdese que

$$V[K\,Y_i] = K^2\ V[Y_i]$$

$$V[1.2\ Y_i] = 1.44\ V[Y_i]$$

$$= 1.44\ (960) = 1\ 382.4$$

9. Los datos son

$$\frac{\sigma}{\bar{Y}} = 0.57 \qquad (1)$$

$$\frac{\sigma}{\overline{Y} + 11} = 0.50 \qquad (2)$$

ya que

$$M[K + Y_i] = M[Y_i] + K$$

$$V[K + Y_i] = V[Y_i]$$

Luego,

$$\sigma = 0.57\ \overline{Y} \text{ reemplazando en } (2)$$

$$0.57\ \overline{Y}\ 0.50\ (\overline{Y} + 11)$$

$$0.07\ \overline{Y} = 5.5$$

$$\overline{Y} = 78.6 \qquad \text{(antes del reajuste)}$$

Además, esta media estaba compuesta por dos grupos: 35 personas con ingreso medio de 40 y 165 personas con ingreso de $\overline{Y}_1$, que se obtendrá de

$$\overline{Y} = \frac{35\ (40\ ) + 165\ (\overline{Y}_1)}{200}$$

$$78.6\ (200) = 1\ 400 + 165\ \overline{Y}_1$$

$$\overline{Y}_1 = \frac{15\ 720 - 1\ 400}{165} = \frac{14\ 360}{165} \doteq 87$$

Las nuevas medias aritméticas después de los reajustes serán:
El primer grupo de 35 personas tendrá un ingreso promedio de E° 71.
El segundo grupo de 165 personas tendrá un ingreso medio de E° 98 (87 + 11)
La cantidad de dinero necesaria será:

$$C.D. = 35\ (71) + 165\ (98) = 2\ 485 + 16\ 170 = 18\ 655$$

10. Los datos son:

$$M[3\ Y_i] = 54$$

$$M_c\ [Y_i + 5] = 24$$

Luego

$$M[3\ Y_i] = 3\ M[Y_1] = 54$$

$$M[Y_i] = 18$$

$$M_c\,[Y_i] = \left[\frac{\Sigma Y_i^2 n_i}{n}\right]^{\frac{1}{2}}$$

$$M_c\,[Y_i + 5] = \left[\frac{\Sigma (Y_i + 5)^2 n_i}{n}\right]^{\frac{1}{2}} = 24$$

$$\frac{\Sigma (Y_i^2 n_i + 10\, Y_i n_i + 25\, n_i)}{n} = 576$$

$$\frac{\Sigma Y_i^2 n_i}{n} + 10\,\frac{Y_i n_i}{n} + 25 = 576$$

$$\frac{\Sigma Y_i^2 n_i}{n} + 10\,(18) = 551$$

$$\frac{\Sigma Y_i^2 n_i}{n} = 371$$

Siendo

$$\sigma^2 = \frac{\Sigma Y_i^2 n_i}{n} - \bar{Y}^2 = 371 - 324 = 47$$

$$\sigma = 6.86$$

$$C.V. = \frac{6.86}{18} = 38.11\%$$

Aplicando las propiedades de la varianza, es posible llegar al mismo resultado:

$$V\,[Y_i] = V\,[Y_i + 5] = M_c^2 - (\bar{Y} + 5)^2 = 24^2 - 23^2 = 47$$

11. La descomposición de σ^2 en σ_b^2 y σ_w^2 es útil para llegar al valor de la varianza del conjunto

$$\sigma^2 = \frac{\Sigma (\bar{Y}_h - \bar{Y})^2 n_h}{n} + \frac{\Sigma \sigma_h^2 n_h}{n}$$

Será necesario calcular $\bar{Y}$:

$$\bar{Y} = \frac{\Sigma Y_h n_h}{n} = \frac{40\,(200) + 60\,(300)}{500} = 52$$

$$\sigma^2 = \frac{(40 - 52)^2\, 200 + (60 - 52)^2\, 300}{500} + \frac{4\,900\,(200) + 1\,600\,(300)}{500}$$

$$\sigma^2 = \frac{288 + 192}{5} + \frac{9\,800 + 4\,800}{5} = \frac{15\,080}{5} = 3\,016$$

12. Los datos son:

$$\overline{Y} = 0.4$$

$$\sigma = 0.2$$

a) Se tendrá que <u>estandarizar</u> el valor 0.7:

$$t_o = \frac{0.7 - 0.4}{0.2} = 1.5$$

Luego en la tabla se ve que

$$P(t_i \geqslant t_o) = P(t_i \geqslant 1.5) = 0.5 - 0.433 = 6.7\%$$

b) <u>Estandarizando</u> ambos valores se tiene

$$t_o = \frac{0.2 - 0.4}{0.2} = -1 \qquad t_1 = \frac{0.5 - 0.4}{0.2} = 0.5$$

$$P(t_o \leqslant t_i \leqslant t_1) = P(-1 \leqslant t_i \leqslant 0.5) = 0.341 + 0.195 = 53.6\%$$

c) <u>Estandarizando se tiene</u>

$$t_o = \frac{0.35 - 0.4}{0.2} = -0.25$$

$$P(t_i \leqslant t_o) = P(t_i \leqslant -0.25) = 40.1\%$$

13. Si la función es

C. Es. = 4.5 C. En. - 5 aplicando el operador media se tiene:

M [C. Es.] = 4.5 M [C. En.] - 5 reemplazando valores

M [C. Es.] = 4.5 [40] - 5 = 175

Aplicando el operador varianza:

V [C. Es.] = $(4.5)^2$ V [C. En.]

V C. Es. = 20.25 [225] = 4 556

El coeficiente de variabilidad será:

$$C.V. = \frac{+\sqrt{4\ 556}}{175} = \frac{67.5}{175} = 38.6\%$$

Anexo

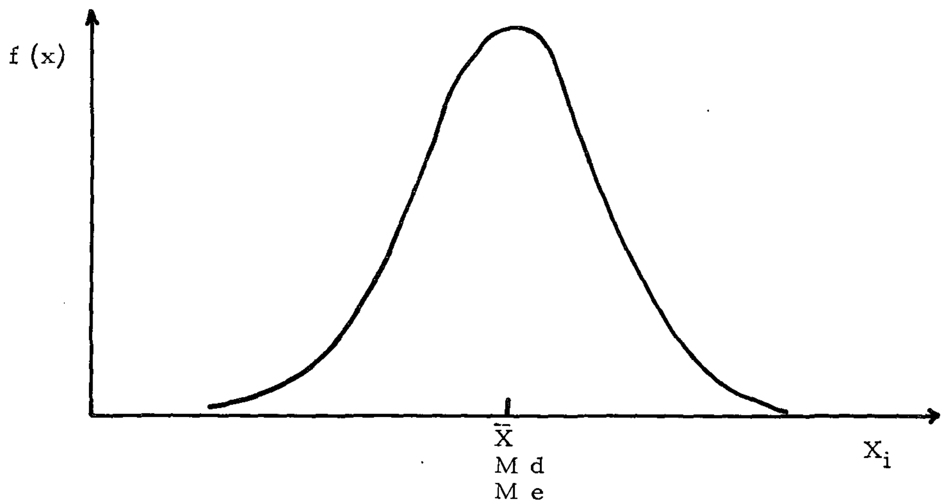
DISTRIBUCION NORMAL

Una de las más utilizadas, dentro de las distribuciones teóricas de variable continua, es la llamada distribución normal o de Gauss. Se trata de una distribución simétrica, unimodal y asintótica al eje de las abscisas. Numerosas variables que se analizan en la investigación socio-económica corresponden o se aproximan a la distribución normal.

La expresión matemática de esta distribución es la siguiente:

$$f(x) = c\, e^{-\frac{1}{2}\left(\frac{x - \bar{x}}{\sigma}\right)^2}$$

y la representación gráfica de una distribución normal particular, es la siguiente:



Si se desea que el área encerrada por la curva sea igual a 1, bastará con integrar la función entre $-\infty$ y $+\infty$, igualar el resultado a 1 y despejar el valor de c.

$$\int_{-\infty}^{\infty} f(x)\, dx = \sigma c \sqrt{2\pi} = 1$$

$$c = \frac{1}{\sigma\sqrt{2\pi}}$$

Luego, la expresión matemática de la distribución normal o función de densidad cuando el área que ella encierra es igual a la unidad, es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\; e^{-\frac{1}{2}\left(\frac{x - \bar{x}}{\sigma}\right)}$$

La anterior expresión muestra que una distribución normal está completamente determinada, cuando se específica el valor de su media aritmética y de su desviación estándar.

Si se obtiene la primera derivada:

$$f' = -\frac{1}{\sigma^2}\ (x - \bar{x}) \cdot f(x)$$

igualando a cero, resulta evidente que el único punto máximo se verifica para $x = \bar{x}$. Se trata entonces de una distribución unimodal donde coinciden la mediana, la moda, y la media aritmética.

Si se obtiene la segunda derivada:

$$f'' = -\frac{1}{\sigma^2}\ [1 - (\frac{x - \bar{x}}{\sigma})^2]\ f(x)$$

se encuentra dos puntos de inflexión igualando la anterior expresión a 0,

$$x = \bar{x} \pm \sigma$$

Geométricamente la desviación estándar es la distancia que hay entre el eje de simetría y un punto de inflexión.

Toda vez que se desee encontrar el área encerrada por la curva normal entre dos puntos cualesquiera, sería necesario integrar la respectiva función entre los límites deseados; existen tablas estándar donde ya se han realizado los cálculos de las integrales. Es necesario, para hacer uso de dichas tablas, expresar previamente la distribución particular que se tenga, en términos de una variable tipificada, definida de la siguiente manera:

$$t = \frac{x - \bar{x}}{\sigma}$$

donde esta nueva variable tiene media 0 y desviación típica 1. En efecto,

$$M[t] = M[\frac{x - \bar{x}}{\sigma}]$$

aplicando las propiedades de la media aritmética

$$M[t] = \frac{1}{\sigma}[M(x) - \bar{x}] = 0$$

Por otra parte,

$$V[t] = V[\frac{x - \bar{x}}{\sigma}]$$

aplicando propiedades de la varianza

$$V[t] = \frac{1}{\sigma^2}\ V[x] = \frac{\sigma^2}{\sigma^2} = 1$$

En consecuencia, el área encerrada en una distribución normal tipificada $N[0, 1]$ será

$$\frac{1}{\sqrt{2\pi}} \int_{-t}^{t} e^{-\frac{t^2}{2}} dt$$

ya que

$$t = \frac{x - \bar{x}}{}$$

$$dx = \sigma\, dt$$

En cualquier texto de estadística aparecen los valores tabulados de la distribución normal tipificada; una manera muy corriente de tabularla es la siguiente:

Variable tipificada	Area comprendida entre 0 y t
t	$\int_{o}^{t} f(t)\,dt$
0.00	0.000
0.05	0.020
0.10	0.040
0.15	0.060
0.20	0.079
0.25	0.099
.	.
.	.
.	.
.	.
.	.
.	.
.	.
0.50	0.195
0.60	0.226
0.80	0.288
0.90	0.316
1.00	0.341
1.20	0.385
1.50	0.433
1.60	0.445
2.00	0.477
2.50	0.494
3.00	0.499
∞	0.500

Es necesario tener cuidado en el uso de estas tablas y cerciorarse que los límites de la integral correspondan con lo que se desea. Así, hay tablas en las cuales se tabula el valor del área entre -t y +t, que corresponderán al doble de área de la tabla mostrada como ejemplo. Se decía que una vez especificados los valores de la media aritmética y la varianza, queda totalmente determinada la función.

Ejemplo: Admítase que las edades de los participantes de un curso siguen aproximadamente una distribución normal, con media 25 años y desviación típica 5. Con estos datos, será posible calcular la proporción de alumnos en cualquier tramo del eje de las abscisas.

Para encontrar la proporción de alumnos con más de 29 años de edad previamente se tendría que tipificar este valor. Sea $x_o = 29$

$$t_o = \frac{x_o - \bar{x}}{\sigma} = \frac{29 - 25}{5} = 0.8$$

Si se observa la tabla de páginas anteriores se tiene que el área entre 0 y 0.8 es de 0.288; luego la proporción de alumnos con más de 29 años será:

$$P(x_i > 29) = P(t_i > 0.8) = 1 - (0.50 + 0.288) = 0.212$$

Donde P indica proporción o probabilidad. Si se quisiera hallar la proporción de alumnos que tienen edades entre 22 y 26 años, sería necesario tipificar previamente cada uno de estos valores:

Sean $x_o = 22$ y $x_1 = 26$

$$t_o = \frac{22 - 25}{5} = -0.6 \qquad t_1 = \frac{26 - 25}{5} = 0.25$$

Recuérdese que la distribución es simétrica y que

$$P(t > t_1) = P(t < -t_1)$$

O sea que

$$P(o > t > -0.6) = P(o < t < 0.6) = 0.226$$

$$P(o < t < 0.25) = 0.099$$

Sumando las dos áreas se concluye que el 32.5% de los alumnos tienen edades comprendidas entre 22 y 26 años.

III. NUMEROS INDICE

A. El problema general

Cuando se define un número índice como un indicador de la tendencia central de un conjunto de elementos que generalmente se expresa como porcentaje, se advierte las limitaciones de todo estadígrafo.

El uso cotidiano que se hace de este instrumento obliga a plantear previamente algunas de sus limitaciones. Es muy útil no olvidar que se trata de un indicador que pretende reflejar el comportamiento de ciertas variables en forma aproximada; en consecuencia no se trata de una medición exacta. Por otra parte es necesario establecer que un número índice plantea una comparación, ya sea en el tiempo o en el espacio, respecto de un punto de referencia denominado base del índice.

A medida que se vayan introduciendo los distintos conceptos que se refieren al conjunto de los números índice, se profundizarán estos planteamientos primarios, ya que la experiencia aconseja que el estudiante tenga ciertas reservas a medida que avanza en este terreno, para evitar posteriormente una utilización indiscriminada y sin las aludidas reservas.

B. Clases de números índice

Fundamentalmente, dentro de la Estadística Económica, interesa disponer de indicadores sobre precios, cantidades y valores.

1 Un índice de precios será un indicador que refleje la variación de los precios de un conjunto de artículos entre dos momentos en el tiempo o dos puntos en el espacio; es el caso de un índice de costo de vida.

2 Un índice de cantidades será un indicador que refleje la variación en las cantidades de un conjunto de productos entre dos momentos en el tiempo o dos puntos en el espacio; por ejemplo, un índice de producción industrial.

3 Por último, un índice de valor indica la variación en el valor total de un conjunto de productos entre dos momentos en el tiempo o dos puntos en el espacio; ejemplo, índice de ventas comerciales.

C. Fórmulas de cálculo

Ocurre que, cuando se trata de analizar la variación, por ejemplo, en el precio de un sólo artículo, no es necesario un indicador especial, basta con expresar la variación en términos porcentuales.

Ejemplo:

Período	Precio del bien	Relación porcentual
1960	20	100
1961	25	125
1962	26	130
1963	30	150

Tomando como punto de referencia el precio del año 1960, y asignándole el valor 100, se calculan, por una regla de tres simple, los índices correspondientes a los otros períodos; así puede decirse que el precio del bien entre 1960 y 1965,[3] ha experimentado un alza de 50%. El cálculo de un índice, cualquiera sea éste, referido a un sólo bien, no necesita pues de un estadígrafo especial.

El problema de los números índice surge cuando se desea averiguar las variaciones de precios o cantidades de un conjunto de artículos; considérese el siguiente ejemplo:

Bienes	Precios 1960	Precios 1964
A (metro)	10	15
B (quintal)	100	140
C (litro)	50	60
D (tonelada)	8	6
	168	221

Para resumir el incremento en los precios de este conjunto de artículos, una primera solución consistiría en sumar los precios en ambos períodos y establecer la variación porcentual entre ambos agregados; de esa manera se llegaría a calcular un índice agregativo simple. Si se considera el año 1962 como base, se tiene:

1960	1964
$100 = \frac{168}{168} \cdot 100$	$131.5 = \frac{221}{168} \cdot 100$

Podría concluirse que el conjunto de estos precios ha variado en 31.5% durante el período. Sin embargo el método tiene dos serias limitaciones; por una parte, estará afectado por las unidades a que estén referidos los precios, y por otra, considera igualmente importante los cuatro bienes, cuando el bien B puede ser trigo y el bien D comino; no se discrimina, si se emplea este método, la importancia relativa de cada artículo. En cuanto a las unidades de medida, considérese el ejemplo anterior, y supóngase que el precio del bien B se refiere al quintal de trigo; si se tuviera el precio del kilo el resultado del índice sería distinto:

Bienes	Precios 1960	Precios 1964
A (metro)	10	15
B (kilo)	1	1.4
C (litro)	50	60
D (tonelada)	8	6
	69	82.4

Asignándole siempre a 1960 el valor 100 como base del índice, se tiene que, para 1964, el índice agregativo simple es ahora de 119.42 (100 · 82.4/69). Se llega a un resultado distinto sin que haya habido variación de precios alguna respecto del caso anterior, excepto que en ambos períodos se tomó un precio referido a una unidad distinta.

Por lo que toca a este problema, es posible evitarlo calculando precios relativos. Se asigna a los precios de cada uno de los artículos en el período base, el valor 100 y por regla de tres simple, se calculan los correspondientes al período para el cual interesa conocer el índice.

Si se observan los ejemplos anteriores, se tendrá:

Bienes	Precios 1960	1964
A	100	150
B	100	140
C	100	120
D	100	75

Obsérvese que en el bien B, sea cual fuere la unidad de medida, el incremento de su precio es de 40%. Pero aunque en este método denominado de cifras relativas se subsana el problema de las unidades, persisten otros problemas que exigen solución: en primer lugar, la elección de un indicador de tendencia central. Se tienen los precios relativos para 1964, pero es necesario resumirlos por medio de un estadígrafo de posición: media aritmética, mediana, etc. Todos ellos conducirán, en general, a resultados distintos. ¿Cuál de ellos admitir? Si bien en cada caso particular habrá un estadígrafo adecuado que satisfaga las necesidades de representatividad que se tenga, en general se utiliza la media aritmética, principalmente por la facilidad que implica su manejo algebraico. Es necesario cuidar que no hayan valores extremos que distorsionen el estadígrafo. En el último ejemplo, si se toma la media aritmética, el índice para 1960 será de 100 y el de 1964 de 121.25, es decir, el conjunto de artículos habrá experimentado un alza de 21.25% en el período. Obsérvese que la conclusión está referida al conjunto, pues se trata de un promedio. Cada bien en particular acusa variaciones dispares en sus precios, e incluso el bien D muestra decrecimiento.

En consecuencia, tomar cifras relativas salva el inconveniente de las unidades de medida, pero aún subsiste el problema de la ponderación, ya que a cada artículo debe asignársele la importancia debida.

Tomando como punto de partida los precios relativos se ensayarán algunos criterios de ponderación, para obtener los índices más usuales en la investigación económica.

Si los precios relativos $\frac{p_n}{p_o}$

donde p_n es el precio en el período dado y p_o el precio en el período base, se ponderan por los valores del año base: $p_o\ q_o$, se obtiene la conocida fórmula de Laspeyres para precios (IPL), es decir:

$$IPL = \frac{\sum \frac{p_n}{p_o} \cdot p_o\, q_o}{\sum p_o\, q_o} = \frac{\sum p_n\, q_o}{\sum p_o\, q_o}$$

La sumatoria se extiende a todos los artículos considerados en el índice. Como en todo promedio aritmético, se divide por la suma de las ponderaciones.

Un índice de precios de Laspeyres debe interpretarse como el nivel que alcanzan los precios en un año dado, respecto de un año base al que se asigna el valor 100, considerando las mismas cantidades del año base en ambos períodos; en otras palabras, se trata de percibir la variación en los precios de una canasta de productos elegidos en el año base y que permanece inalterada durante los períodos sucesivos.

Este índice, por lo tanto, tiene un significado bien concreto. El supuesto que la canasta de productos realmente no registre variaciones significativas, es otro problema; es el analista quien deberá determinar si el supuesto se cumple o no, y por lo tanto, juzgar la conveniencia de utilizar un índice de Laspeyres.

Por otra parte, si los precios relativos $\frac{p_n}{p_o}$, se ponderan por valores híbridos: $p_o\, q_n$, se tiene el índice de precios de Paasche (IPP), que también se utiliza con frecuencia:

$$IPP = \frac{\sum \frac{p_n}{p_o} \cdot p_o\, q_n}{\sum p_o\, q_n} = \frac{\sum p_n\, q_n}{\sum p_o\, q_n}$$

Obsérvese que ahora los precios están multiplicados por las cantidades del año que se calcula (q_n). Por este hecho un índice de precios de Paasche debe interpretarse como la variación de los precios de un conjunto de productos, suponiendo constantes las cantidades del año dado; en otros términos, la canasta de productos que se considera, es la del período que se calcula y se toma esta misma canasta para el año base.

Respecto de los índices de valor, por el significado simple que tienen, no requieren deducciones especiales, ya que son sencillamente el resultado de la división entre los valores del año que se calcula y el año base.

$$IV = \frac{\sum p_n\, q_n}{\sum p_o\, q_o}$$

A continuación se considerará un ejemplo donde se calcularán los índices presentados:

	Año 0		Año 1		Año 2	
Artículos	p	q	p	q	p	q
A	10	4	12	5	20	3
B	4	3	4	3	5	3
C	8	10	8	12	7	15
D	20	2	30	2	40	3

p: precio q: cantidad

No hace falta el cálculo en el año 0, base del índice, ya que coinciden precios y cantidades: $p_n = p_o$ y $q_n = q_o$

Para los índices de Laspeyres se tiene:

$$\text{IPL (año 1)} = \frac{\Sigma\, p_1\, q_o}{\Sigma\, p_o\, q_o} = \frac{48 + 12 + 80 + 60}{40 + 12 + 80 + 40} = \frac{200}{172} = 116.3$$

$$\text{IPL (año 2)} = \frac{\Sigma\, p_2\, q_o}{\Sigma\, p_o\, q_o} = \frac{80 + 15 + 70 + 80}{40 + 12 + 80 + 40} = \frac{245}{172} = 142.4$$

Para los índices de Paasche, suponiendo siempre el año 0 como base del índice:

$$\text{IPP (año 1)} = \frac{\Sigma\, p_1\, q_1}{\Sigma\, p_o\, q_1} = \frac{60 + 12 + 96 + 60}{50 + 12 + 96 + 40} = \frac{228}{198} = 115.2$$

$$\text{IPP (año 2)} = \frac{\Sigma\, p_2\, q_2}{\Sigma\, p_o\, q_2} = \frac{60 + 15 + 105 + 120}{30 + 12 + 120 + 60} = \frac{300}{222} = 135.1$$

El índice de valor:

$$\text{IV (año 1)} = \frac{\Sigma\, p_1\, q_1}{\Sigma\, p_o\, q_o} = \frac{60 + 12 + 96 + 60}{40 + 12 + 80 + 40} = \frac{228}{172} = 115.1$$

$$\text{IV (año 2)} = \frac{\Sigma\, p_2\, q_2}{\Sigma\, p_o\, q_o} = \frac{60 + 15 + 105 + 120}{40 + 12 + 80 + 40} = \frac{300}{172} = 174.4$$

En el siguiente cuadro puede apreciarse la diferencia entre las indicaciones de uno y otro índice

		Indices	
Años	IPL	IPP	IV
0	100.0	100.0	100.0
1	116.3	115.2	115.1
2	142.4	135.1	174.4

Puede apreciarse que el IPL, crece más que el IPP. En el primero se considera constante la canasta de productos del período base; en cambio, en el segundo, se considera constante la canasta de productos del período en que se calcula el índice. Por lo tanto ambos índices indican la variación promedio de los precios bajo supuestos diferentes; sin embargo, es muy frecuente confundir el significado de estos indicadores, porque no se consideran los supuestos de uno y otro.

Con referencia a los índices de cantidad, es necesario hacer el mismo tipo de consideraciones ya que se presentan problemas similares; unidades de medida, ponderaciones, etc.

Siguiendo el mismo criterio de los índices de precios, puede obtenerse el índice de cantidades de Laspeyres (IQL)

$$IQL = \frac{\Sigma \frac{q_n}{q_o} \cdot p_o q_o}{\Sigma q_o p_o} = \frac{\Sigma q_n p_o}{\Sigma q_o p_o}$$

El índice de cantidades de Paasche será (IQP)

$$IQP = \frac{\Sigma \frac{q_n}{q_o} \cdot q_o p_n}{\Sigma q_o p_n} = \frac{\Sigma q_n p_n}{\Sigma q_o p_n}$$

Mientras el IQL representa la variación en las cantidades suponiendo constantes los precios del período base, el IQP representa la variación de las cantidades suponiendo constantes los precios del período calculado. Nuevamente se insiste sobre la necesidad de no descuidar estos supuestos, cuando se interpreta un índice.

Las fórmulas presentadas son, como se dijo, las de uso más frecuente. Hay una cantidad extraordinaria de fórmulas de indices que se diferencian unas de otras según los factores de ponderación utilizados. Por razones de brevedad sólo se mostrarán las más conocidas

Marshall - Edgeworth para precios

$$IPM = \frac{\Sigma p_n (q_o + q_n)}{\Sigma p_o (q_o + q_n)}$$

Indice de precios de Keynes

$$IPK = \frac{\Sigma p_n (q_o \wedge q_n)}{\Sigma p_o (q_o \wedge q_n)}$$

donde el signo $\wedge$ es ínfimo y quiere indicar que se tome la *menor* de las cantidades que están a sus costados.

La llamada fórmula "ideal" de Fisher para precios, que es la media geométrica de los índices de Laspeyres y de Paasche.

$$IPF = \sqrt{IPL \cdot IPP} = \sqrt{\frac{\Sigma p_n q_o}{\Sigma p_o q_o} \cdot \frac{\Sigma p_n q_n}{\Sigma p_o q_n}}$$

En estas últimas fórmulas, bastará reemplazar p por q, para obtener las fórmulas correspondientes a índices de cantidad.

D. Pruebas sobre los números índice

Irving Fisher, quien plantea la fórmula por él llamada ideal, propuso pruebas para calificar los números índice.

1. Prueba de reversión de factores

La prueba se basa sobre un criterio de analogía: lo que es cierto para un producto, debiera ser cierto para un conjunto de ellos. Así, para cualquier artículo se tiene que: precio x cantidad = valor

Las fórmulas de Laspeyres y Paasche no satisfacen esta prueba, como puede verse a continuación.

IPL x IQL $\neq$ IV, en efecto

$$\frac{\Sigma p_n q_o}{\Sigma p_o q_o} \cdot \frac{\Sigma q_n p_o}{\Sigma q_o p_o} \neq \frac{\Sigma p_n q_n}{\Sigma p_o q_o}$$

IPP x IQP $\neq$ IV

$$\frac{\Sigma p_n q_n}{\Sigma p_o q_n} \cdot \frac{\Sigma q_n p_n}{\Sigma q_o p_n} \neq \frac{\Sigma p_n q_n}{\Sigma p_o q_o}$$

La fórmula de Fisher sí satisface esta prueba

IPF x IQF = IV

$$\left(\frac{\Sigma p_n q_o}{\Sigma p_o q_o} \cdot \frac{\Sigma p_n q_n}{\Sigma p_o q_n}\right)^{\frac{1}{2}} \left(\frac{\Sigma q_n p_o}{\Sigma q_o p_o} \cdot \frac{\Sigma q_n p_n}{\Sigma q_o p_n}\right)^{\frac{1}{2}} = \frac{\Sigma p_n q_n}{\Sigma p_o q_o}$$

Simplificando términos semejantes se tiene:

$$\frac{\Sigma p_n q_n}{\Sigma p_o q_o} = \frac{\Sigma p_n q_n}{\Sigma p_o q_o}$$

Sin embargo, esta prueba también se satisface con una combinación de índices de precios de Laspeyres y cantidades de Paasche o viceversa, como se comprueba a continuación:

IPL x IQP = IV

$$\frac{\Sigma p_n q_o}{\Sigma p_o q_o} \cdot \frac{\Sigma q_n p_n}{\Sigma q_o p_n} = \frac{\Sigma p_n q_n}{\Sigma p_o q_o}$$

IPP x IQL = IV

$$\frac{\Sigma p_n q_n}{\Sigma p_o q_n} \cdot \frac{\Sigma q_n p_o}{\Sigma q_o p_o} = \frac{\Sigma p_n q_n}{\Sigma p_o q_o}$$

Estas dos últimas relaciones merecen retenerse, por que se utilizan con frecuencia.

2. Pruebas de reversión temporal

Nuevamente el criterio de analogía; si el precio de un producto es en el período "a" de 40 y en el período "b" de 50, en el primer período se observa que el precio es el 80% del que se da en el período "b", y en éste el 125% del precio del período "a". Lógicamente el producto de estos porcentajes debe dar 1, es decir:

$$\frac{p_n}{p_o} \times \frac{p_o}{p_n} = 1$$

Esta prueba, no la cumplen las fórmulas de Laspeyres y Paasche.

En efecto: $$IPL_{b,a} \times IPL_{a,b} \neq 1$$

donde el primer subíndice indica el período que se calcula y el segundo el período base.

$$\frac{\Sigma p_b q_a}{\Sigma p_a q_a} \cdot \frac{\Sigma p_a q_b}{\Sigma p_b q_b} \neq 1$$

$$IPP_{b,a} \cdot IPP_{a,b} \neq 1$$

$$\frac{\Sigma p_b q_b}{\Sigma p_a q_b} \cdot \frac{\Sigma p_a q_a}{\Sigma p_b q_a} \neq 1$$

La fórmula de Fisher, sí cumple la prueba

$$IPF_{b,a} \times IPF_{a,b} = 1$$

$$\left(\frac{\Sigma p_b q_a}{\Sigma p_a q_a} \cdot \frac{\Sigma p_b q_b}{\Sigma p_a q_b}\right)^{\frac{1}{2}} \left(\frac{\Sigma p_a q_b}{\Sigma p_b q_b} \cdot \frac{\Sigma p_a q_a}{\Sigma p_b q_a}\right)^{\frac{1}{2}} = 1$$

3. Prueba circular

Si el precio de un producto es de 10 en el primer período, de 12 en el segundo y de 18 en el tercero, se comprueba que en el segundo período el precio es 120% del precio en el primero y en el tercer período 150% del precio en el segundo. En consecuencia el precio del tercer período es 180% del registrado en el primero:

$$(120\%) \quad (150\%) \quad = \quad 180\%$$

La prueba circular no la cumple ninguna de las fórmulas analizadas. Suponiendo tres períodos para IPL se tiene:

$$IPL_{3,2}\, IPL_{2,1} \neq IPL_{3,1}$$

donde el primer subíndice indica el período que se calcula y el segundo el período base.

$$\frac{\Sigma p_3 q_2}{\Sigma p_2 q_2} \cdot \frac{\Sigma p_2 q_1}{\Sigma p_1 q_1} \neq \frac{\Sigma p_3 q_1}{\Sigma p_1 q_1}$$

Esta relación sólo se cumpliría en el caso que $q_1 = q_2 = q_3$

Sin embargo, la relación planteada se cumple en forma aproximada, cuando no existen diferencias significativas en las cantidades durante los distintos períodos.

El lector podrá comprobar que ni la fórmula de Paasche ni la de Fisher, cumplen la prueba circular, pudiendo seguir el mismo proceso de comprobación realizado para la fórmula de Laspeyres.

Respecto de estas pruebas, es conveniente aclarar que la existencia de índices que no las cumplan no justifica dejarlos de lado; interesa más el significado concreto del índice, teniendo en cuenta sus alcances y limitaciones. Es así como la fórmula ideal de Fisher, que satisface las pruebas por él planteadas, no es susceptible de ser claramente interpretada, ya que se trata de la combinación de dos índices que si por separado adquieren cabal significado, al combinarlos ofrecen dificultades para su interpretación.

E. Base de un número índice

Al definir un número índice se ha destacado que se trata de una comparación de dos momentos en el tiempo o dos puntos en el espacio. El momento o punto con respecto al cual se establece la comparación recibe el nombre de base de un índice y se le asigna el valor 100, para analizar las variaciones porcentuales. Respecto de la elección del período base hay que tener siempre presente el objetivo que se persigue con el índice; en general se estima que el período base debe ser un período normal. Cabe preguntarse qué se entiende por normalidad en estos casos, cuando en los países en desarrollo los cambios son muy frecuentes y la anormalidad es un denominador común. Tal vez al definir el período base fuese más sensato pensar en un período durante el cual no existan accidentes o cambios violentos. Por lo demás, será necesario cambiar la base del índice cuando los supuestos planteados pierdan validez a medida que pasa el tiempo; es el caso de los índices de costo de vida, cuya base debe modificarse toda vez que la estructura de consumo presente cambios significativos con respecto de la admitida en el período base.

Sobre este mismo asunto, será necesario distinguir dos tipos de base: base fija y base variable. Los índices de base fija son aquellos que mantienen como base un período fijo de referencia, en tanto que los índices de base variable son aquellos que tienen como base el período inmediatamente anterior. Con un índice de base fija puede calcularse el correspondiente de base variable y viceversa; los resultados, en general, diferirán de los que se obtendrían a partir de los datos originales, ya que las fórmulas usuales no cumplen con la prueba circular.

Ejemplo: Supóngase que el índice de Laspeyres para los precios de los materiales de construcción sea el siguiente:

	Indice base 1960 = 100
1960	100
1961	110
1962	120
1963	144
1964	180
1965	190

El correspondiente índice de base variable sería:

	Indice de base variable
1960	-
1961	110.0
1962	109.1
1963	120.0
1964	125.0
1965	105.5

Otra operación que es muy usual respecto de los índices de base fija es la del empalme; se trata, como indica su nombre, de empalmar índices con base distinta. Obsérvese el siguiente ejemplo, donde se tiene un índice para el período 1956-1959 con base 1956, y otro índice de 1959 a 1962 con base 1959, y se pretende tener una serie para todo el período 1956-1962.

Años	Indice base 1956 = 100	Indice base 1959 = 100
1956	100	
1957	120	
1958	150	
1959	180	100
1960		110
1961		132
1962		150

Mediante una sencilla regla de tres, puede completarse cualquiera de las dos series, para tener el movimiento del índice durante todo el período.

Años	Indice base 1956 = 100	Indice base 1959 = 100
1956	100.0	55.6
1957	120.0	66.7
1958	150.0	83.3
1959	180.0	100.0
1960	198.0	110.0
1961	237.6	132.0
1962	270.0	150.0

Debe advertirse que este tipo de empalmes, significa sólo una aproximación que puede ser muy defectuosa, dependiendo de la similitud de las bases y de sus supuestos. En todo caso, siempre es posible recurrir a estos empalmes toda vez que se tenga conciencia de sus limitaciones.

F. Utilización de los números índice

Un número índice indica la evolución de precios, cantidades y valores, para un conjunto de productos. Prestan en consecuencia la utilidad inmediata de reflejar la tendencia de los cambios y ritmos de los conceptos señalados. Ese sólo hecho ya justifica su cómputo y su periódica utilización en la investigación socio-económica. Sin embargo prestan además otros servicios sobre los que es conveniente hacer algunos comentarios.

1) La deflactación*

Las alteraciones en los sistemas y niveles de precios que se presentan dentro de la actividad económica, originan dificultades en la comparación de valores monetarios que corresponden a períodos de tiempo distanciados. No es mucho lo que se puede deducir de la comparación de valores nominales, es decir, de valores expresados en unidades monetarias de distinto poder adquisitivo. Para poder llegar a conclusiones válidas acerca del comportamiento de una variable que represente "valor", será necesario expresar los montos monetarios nominales, en unidades homogéneas; esta transformación recibe el nombre de deflactación, y con ella se pretende eliminar, exclusivamente, el efecto de alteraciones en los precios.

El proceso de deflactación exige disponer de un índice deflactor, es decir, de un indicador que proporcione una pauta de las alteraciones en los precios que tengan relación con la variable que se pretende deflactar. Es de fundamental importancia recordar que no existe un índice deflactor único; cada variable, en rigor, debería tener un deflactor adecuado. La disponibilidad de sólo un reducido número de índices de precios, hace que éstos sean utilizados indiscriminadamente para una serie de propósitos; aunque este hecho puede adolecer de errores conceptuales, muchas veces se justifica el procedimiento, porque interesa conocer un orden de magnitud antes que un valor exacto, siempre que se tenga conciencia de las limitaciones del método. En todo caso, aún dentro de las escasas disponibilidades de índices de precios, es posible llegar a soluciones aceptables, ya sea eligiendo racionalmente un índice de precios como deflactor o combinando y ponderando en forma adecuada dos o más índices; este último procedimiento, si bien puede no conducir a soluciones ideales, por lo menos puede representar una disminución de las posibles distorsiones.

Antes de profundizar este problema de la deflactación, cuando se desea transformar unidades monetarias heterogéneas (unidades de cada período) en unidades monetarias homogéneas (unidades del período base), y permitir de este modo la comparación en el tiempo, el primer recurso al cual se apela, es expresar los montos monetarios nominales en unidades de moneda extranjera de valor más o menos estable; dólares, libras, etc. Mas sobre este procedimiento caben algunas objeciones. Los Gobiernos tienen instrumentos que les permiten fijar los tipos de cambio con las monedas extranjeras en forma arbitraria (arbiraria para los fines que aquí se comenta) y que en general no representan las alteraciones en los niveles de precios. En Chile hay una experiencia reciente; en el transcurso de menos de dos años, la unidad monetaria chilena llego a sufrir una fuerte devaluación, (E° 1 053 en noviembre de 1962 a E° 3 400 en abril de 1963, por dólar norteamericano tipo de cambio corredor). Si un valor dado en escudos se

---

* Se empleará el neologismo "deflactación" para indicar la metodología de transformación de valores expresados en precios corrientes a valores en precios constantes; el propósito es limitar el empleo de "deflación", para caracterizar con esta última expresión el fenómeno económico opuesto a la inflación.

expresase en dólares en ambas fechas, podría concluirse que entre los dos períodos mencionados dicho valor se redujo a la tercera parte; si bien esto puede ser cierto para una persona que va a gastar sus ingresos en EE. UU., no lo es para quien efectúa sus desembolsos en Chile, porque durante ese período ningún índice de precios propiamente tal ha sufrido un aumento de 200 por ciento.

Por otra parte, una segunda objeción a este procedimiento consiste en el hecho que aún las economías más estables pueden sufrir cierto grado de inflación. Los dos argumentos expuestos descalifican, en general, la conversión a unidades monetarias extranjeras como una alternativa de la deflactación. La mecánica de la deflactación implica dividir los montos monetarios nominales por el índice de precios elegido como deflactor adecuado; y su explicación podrá encontrarse en la siguiente regla de tres. Si en el año n se tiene un valor nominal VN_n y un índice de precios IP_n, ¿cuál sería este valor expresado en unidades monetarias de igual poder adquisitivo que las del año base? En otros términos, ¿cuál sería este valor si el índice de precios no hubiera variado?

El planteamiento queda así reducido:

$$\begin{array}{cc} IP_n & . \ VN_n \\ 100 & x \end{array}$$

$$x = \frac{VN_n}{IP_n} \; . \; 100 = \text{Valor real}$$

Desde otro punto de vista, se justifica la deflactación pensando en los componentes de un valor: precio por cantidad

$$\text{Indice de precios} \; x \; \text{cantidad} = \text{Valor (100)}$$

$$\text{Cantidad} = \frac{\text{Valor (100)}}{\text{Indice de precios}} = \text{Valor real}$$

Ahora bien, la evolución de las cantidades en el tiempo está libre, por así decirlo, de influencias monetarias directas, muestra la evolución física, real, de una serie, que es precisamente lo que se pretende con la deflactación.

Los valores reales así obtenidos, están expresados en unidades monetarias que tienen un poder adquisitivo correspondiente al año base del índice deflactor.

Con referencia a este último planteamiento, es necesario distinguir dos tipos de base. Se llamará base propiamente tal al período que corresponde al diseño del índice, donde se perfila la muestra, se establecen las ponderaciones, etc. Por otra parte, se utilizará la expresión "base aritmética" para referirse a cualquier período, que por transformación lineal, se le haya asignado el valor 100. Los cambios aritméticos de base implican hasta cierto punto una arbitrariedad que puede originar algún tipo de perturbaciones al establecer las comparaciones. Sus efectos serán tanto mayores, cuanto más distante sea la base propiamente tal del índice deflactor. Cuando hay consenso que los supuestos de la construcción y diseño del índice siguen siendo válidos, los cambios aritméticos de base no introducirán deformaciones significativas; por ello, antes de deflactar será necesario decidir de qué año serán las unidades monetarias en que interesa expresar los valores reales. Es recomendable, si se cumple la condición de constancia de los supuestos de la construcción del índice, expresar los valores nominales en unidades monetarias del año más reciente, lo que se consigue mediante un índice deflactor que tenga como base el último año en el sentido cronológico. La utilidad de la sugerencia anotada, radica en el hecho que el investigador tiene una visión reciente y tal vez objetiva del sistema de precios imperante, por lo que es probable que la obtención de conclusiones quede facilitada. Es necesario destacar que un proceso de deflactación conduce a valores reales que pueden tener dos interpretaciones: una expresión física o un poder de compra. Una variable monetaria está compuesta por una suma de valores del tipo $\Sigma p_n q_n$; si esta serie se deflacta por un índice de precios de los productos considerados en la serie nominal, el resultado será una expresión física de la serie. En efecto, utilizando un índice deflactor de Paasche, se tiene:

$$\frac{\text{Valor nominal}}{\text{IPP}} = \frac{\Sigma p_n q_n}{\dfrac{\Sigma p_n q_n}{\Sigma p_o q_n}} = \Sigma p_o q_n$$

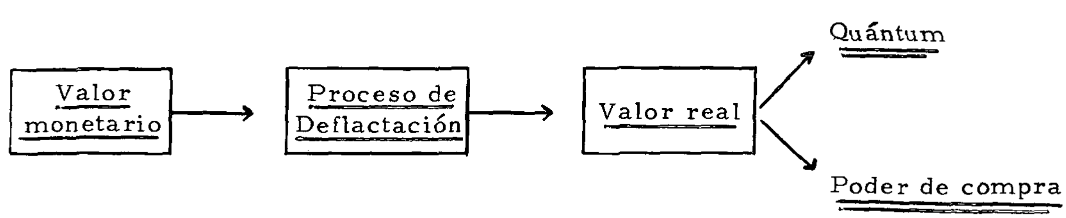
El resultado es evidentemente un quantum, es decir, cantidades del período n, valorizadas a precios del período base. Si se hubiera deflactado por un índice de precios de Laspeyres, se tendría:

$$\frac{\text{Valor nominal}}{\text{IPL}} = \frac{\Sigma p_n q_n}{\dfrac{\Sigma p_n q_o}{\Sigma p_o q_o}} = \text{IQP} \cdot \Sigma p_o q_o$$

Resultado que equivale a proyectar un valor en el período base, a través de un índice de cantidades de Paasche. También representa una evolución física de la serie, aunque con connotaciones diferentes al caso anterior. Cabe hacer notar que cuando se desea llegar a un expresión física rigurosa, sólo existe un tipo de deflactor adecuado: el que contienen los productos incluidos en la serie nominal.

Por otra parte, cuando se quiere obtener un poder de compra, es necesario especificar qué uso se dará a un monto monetario; de esta manera, si el sueldo de un empleado en diferentes períodos se deflacta por un índice de precios al consumidor, el resultado sería un poder de compra en términos de la canasta de productos elegida en el índice deflactor. Si el sueldo de dicho empleado se deflacta por índice de valores bursátiles, el resultado será un poder de compra en términos de acciones y bonos. El uso que se dará a un monto monetario es el que determina el tipo de poder de compra resultante.

Resumiendo esquemáticamente se tiene:



A continuación y a título de ejemplo, se presenta los resultados de un proceso de deflactación:

Años	Sueldo de un empleado (u. m. de c/año) A	Indice de precios al consumidor (base 1963 = 100) B	Sueldo real (u. m. de 1963) $\frac{A}{B} \cdot 100$
1958	400	36.5	1 096
1959	480	50.6	946
1960	600	56.5	1 062
1961	680	60.9	1 117
1962	720	69.3	1 039
1963	900	100.0	900

La deflactación presentada implica haber elegido como deflactor el índice de precios al consumidor. Tal vez de los valores reales no pueda deducirse, en forma categórica, que el empleado haya sufrido una merma de esa magnitud en su poder de compra; lo que hubiera ocurrido si dicho empleado gastara todo su ingreso en la forma que lo hace el empleado típico elegido como padrón en el índice de precios al consumidor. Si dicho empleado sólo gasta en mantención el 60 por ciento de sus ingresos y el 40 por ciento restante lo dedica, por ejemplo, a la construcción de una vivienda, la deflactación se realizaría en dos partes: la primera parte (60%) se deflactaría por el índice de precios al consumidor y el saldo debería deflactarse por un índice de precios de insumos de la construcción; de esta manera se obtendría la expresión real de su poder de compra, en términos del uso que dará a sus ingresos.

Cuando se ha realizado una deflactación, es necesario tener presente que los valores reales así obtenidos son simplemente aproximaciones, que serán tanto mejores cuanto más representativo sea el índice de los precios que se cancelarán con el monto monetario. Esto no quiere decir que sea indispensable "construir" índices si no se dispone de los más adecuados; sólo se pretende destacar la necesidad de seleccionar los índices disponibles, y en todo caso, tener presente las limitaciones que acuse una deflactación obligada por un índice que no sea el mejor. Es responsabilidad de los organismos autorizados y de los principales usuarios la elaboración de índices de precios de la actividad económica.

La deflactación, mecánicamente es un asunto trivial, pero la elección de deflactores adecuados requiere mucha atención. Véase por ejemplo qué sucede con los índices de producción y ventas industriales de Chile.* Observando las cifras siguientes debe admitirse que no existe la concordancia que debería existir, tampoco se encuentra razones que expliquen la diferencia; aunque es muy probable que el desajuste radique en la calidad del deflactor utilizado, antes que en las variaciones de existencias.

* César Molestina, "Los índices de producción industrial manufacturera y ventas reales industriales. Un comentario acerca de su comportamiento", mayo de 1963.

Período		Indice de producción industrial Base 1959 = 100	Indice de ventas industriales reales Base 1959 = 100
1959		100.0	100.0
1960		103.5	98.3
1961		106.1	110.6
1962		113.8	127.7
1963			
	Enero	108.4	120.9
	Abril	122.2	129.8
	Agosto	112.8	132.1

De la simple observación de estas series surgen las contradicciones anotadas. Si se admite que el índice de producción industrial es un indicador confiable, el índice de ventas sería el que adolece de defectos. Estos defectos pueden deberse a dos causas no excluyentes: que el índice de ventas nominales no sea lo suficientemente acertado, por errores inherentes o ajenos al muestreo, o que el deflactor utilizado no tiene la relación apropiada con los precios de los bienes industriales, o una combinación de ambas causas.

2 El deflactor implícito del producto bruto

No es necesario citar la imperiosa necesidad que hay de disponer de un sistema de contabilidad social expresada en precios constantes; en este sentido se tratará principalmente la deflactación del producto bruto.

Es importante aclarar la idea de deflactor implícito. Este aparece como resultado de un proceso de deflactación cuando se trata de cumplir con alguna restricción; se presenta principalmente cuando una variable global ha sido desglosada en componentes, con el objeto de identificar un índice deflactor adecuado, con cada componente. En estos casos la restricción aparece como suma de componentes que reproducen la variable global; es necesario agregar que el deflactor implícito se origina en el hecho que la restricción (muchas veces definición) debe ser satisfecha tanto en valores nominales o corrientes como en valores reales o constantes; entonces se dice que el proceso de deflactación es coherente. El deflactor resultante de una deflactación coherente se denomina implícito, porque justamente está implícito en el cumplimiento de la restricción.

Supóngase que X, Y y Z representan valores sujetos a la restricción:

$$X + Y + Z = W$$

restricción que se satisface en valores corrientes. Si además se desea que esta restricción sea satisfecha en valores constantes, se tiene:

$$\overline{X} + \overline{Y} + \overline{Z} = \overline{W}$$

donde la barra sobre el símbolo se utiliza para indicar que se trata de valores reales, es decir, que

$$\bar{X} = \frac{X}{IPX}$$

$$\bar{Y} = \frac{Y}{IPY}$$

$$\bar{Z} = \frac{Z}{IPZ}$$

La suma de $\bar{X} + \bar{Y} + \bar{Z}$ reproduce el valor real de W, es decir, $\bar{W}$. ¿Cuál deberá ser el deflactor de W para que el valor real resultante $\bar{\bar{W}}$ coincida con la suma de $\bar{X} + \bar{Y} + \bar{Z}$? En efecto, se tiene que

$$\frac{X}{IPX} + \frac{Y}{IPY} + \frac{Z}{IPZ} = \frac{W}{IPW}$$

Basta con despejar IPW de la anterior relación:

$$IPW = \frac{W}{\frac{1}{IPX} \cdot X + \frac{1}{IPY} \cdot Y + \frac{1}{IPZ} \cdot Z}$$

Se comprueba que el deflactor implícito IPW es el promedio armónico de los deflactores componentes, donde los factores de ponderación son justamente los valores nominales; otra forma de calcular el deflactor implícito es comparar $\bar{\bar{W}}$ que se obtiene como suma de $\bar{X} + \bar{Y} + \bar{Z}$ con W.

$$IPW = \frac{W}{\bar{\bar{W}}}$$

Para determinar el deflactor implícito del producto bruto, habrá que pensar previamente en alguna definición. Si la restricción es del siguiente tipo:

$$P = C + I + E - M$$

se determinarán los deflactores para consumo, inversión, exportaciones e importaciones y se obtendrá como suma el producto real ($\bar{P}$).

$$\bar{P} = \frac{C}{IPC} + \frac{I}{IPI} + \frac{E}{IPE} - \frac{M}{IPM}$$

El deflactor implícito del producto (DI), según esta restricción, estará dado por la relación

$$DI = \frac{P}{\bar{\bar{P}}}$$

Puede verificarse al mismo tiempo que el deflactor implícito es equivalente a la media armónica de los deflactores componentes.

$$DI = \frac{P}{\frac{1}{IPC} \cdot C + \frac{1}{IPI} \cdot I + \frac{1}{IPE} \cdot E - \frac{1}{IPM} \cdot M}$$

Si la restricción fuera de otro tipo, por ejemplo que la suma de los valores agregados sectoriales reproducen el producto bruto, se tendrá otro deflactor implícito sujeto a la restricción propuesta. Habría que discutir previamente la posibilidad de encontrar deflactores adecuados para el valor agregado; una alternativa, por cierto muy discutible, es deflactar el valor agregado de cada sector, por los precios de los bienes que el sector produce. El resultado sería un poder de compra de los valores agregados sectoriales, en términos de los bienes que produce cada sector. Puede argumentarse en favor de este método, sosteniendo que la contrapartida física del valor agregado es justamente la cantidad de bienes producidos, sin embargo, la justificación es en extremo débil. En todo caso, en los cursos de Contabilidad Social se muestran las ventajas y desventajas de este y otros métodos de deflactación. Para ilustrar, en seguida se detalla este procedimiento. Si se desea llegar a expresiones reales, cada rama de actividad debería deflactarse por un índice de precios relacionado directamente con dichas ramas de actividad. Así el valor agregado por el sector industrial, debería deflactarse por un índice de precios de bienes industriales; el valor agregado por el sector agropecuario, se deflactaría por un índice de precios de bienes agropecuarios. De esta manera se obtendrían valores reales en cada rama que representaría una aproximación a la evolución física de lo producido por cada sector. Sumando los valores reales de todas las ramas de actividad, se tendría el producto bruto en términos reales; y comparando los productos brutos en valores nominales y reales se obtiene el llamado deflactor implícito del producto, que no es otra cosa que un índice general promedio de los precios que rigen en la actividad económica.

Sean:

PN_{ij} el producto nominal del sector "i" en el año "j".

IP_{ij} el índice de precios del sector "i" en el año "j".

PR_{ij} el producto real del sector "i" en el año "j".

De la deflactación resulta

$$PR_{ij} = \frac{PN_{ij}}{IP_{ij}} \cdot 100$$

Si se suman todos los productos reales por sector en un año cualquiera "j", se tiene el producto bruto real en el año "j".

$$PR_j = \sum_{i=1}^{L} PR_{ij} = \sum_{i=1}^{L} \frac{PN_{ij}}{IP_{ij}} \cdot 100$$

donde i = 1, 2 ... L representa el número de sectores considerados. Por otra parte, el producto bruto real PR_j se obtiene deflactando el producto bruto nominal PN_j, por el deflactor implícito DI_j, conceptos todos referidos a un período j, es decir:

$$PR_j = \frac{PN_j}{DI_j} \cdot 100$$

Reemplazando en la igualdad anterior, se tiene

$$\frac{PN_j}{DI_j} \cdot 100 = \sum_{i=1}^{L} \frac{PN_{ij}}{IP_{ij}} \cdot 100$$

Despejando DI_j

$$DI_j = \frac{PN_j}{\sum_{i=1}^{L} \frac{PN_{ij}}{IP_{ij}}}$$

que justamente corresponde con la definición de media armónica de los deflactores sectoriales, considerando como ponderaciones los productos nominales de cada sector, ya que

$$PN_j = \sum_{i=1}^{L} PN_{ij}$$

A continuación se presentará un ejemplo para ilustrar el proceso de deflactación de los productos sectoriales.

Supóngase que la actividad económica ha sido dividida en cuatro sectores para los cuales se dispone de las siguientes informaciones:

Producto nominal

(unidades monetarias corrientes)

Sectores	1960	1961	1962
Minería	200	300	400
Agricultura	300	350	400
Industria	200	250	300
Servicios	500	600	700
Producto bruto nominal	1 200	1 500	1 800

Indices de precios (base 1960 = 100)

Productos mineros	100	130	150
Productos agrícolas	100	110	120
Productos industriales	100	120	130
Servicios	100	110	120

Deflactando el producto de cada sector, por el índice de precios correspondiente, se tendrá los productos reales de cada sector:

Producto real

(unidades monetarias constantes de 1960)

Sectores	1960	1961	1962
Agricultura	200	230.8	266.7
Minería	300	318.2	333.3
Industria	200	208.3	230.8
Servicios	500	545.5	583.3
Producto bruto real	1 200	1 302.8	1 414.1

Para calcular el deflactor implícito, recuérdese que

$$DI_j = \frac{PN_j}{PR_j} \cdot 100$$

	1960	1961	1962
Deflactor implícito	100	115.1	127.3

Como se demostró, estos valores equivalen a los promedios armónicos de los índices deflactores.

Con respecto a la deflactación del producto bruto es conveniente advertir que puede prestarse a presentaciones caprichosas, según sea la base de los índices deflactores. En otras palabras, el producto bruto real, mostrará distintas tasas de crecimiento según sea la base de los deflactores. Aparentemente esto no tendría por qué ocurrir, ya que un cambio aritmético de la base del índice deflactor no debería afectar las variaciones reales en el producto bruto. Mediante un ejemplo que pretende exagerar la situación antes que representar un hecho real, se ilustra este planteamiento.

Para abreviar, supóngase que la actividad económica ha sido dividida en dos sectores: primario y secundario. Los datos son los siguientes:

Producto bruto nominal

(u.m. corrientes)

	1956	1962
Sector primario	1 000	1 600
Sector secundario	1 000	1 800

Indices deflactores (base 1956 = 100)

Sector primario	100	110
Sector secundario	100	300

Si se calcula el producto bruto real, en precios constantes de 1956, se tiene:

Producto bruto nominal

(u.m. constantes)

	1956	1962
Sector primario	1 000	1 450
Sector secundario	1 000	600
	2 000	2 050

Entre ambos años el crecimiento es 2.5 por ciento, si se computa el producto en precios de 1956. Si aritméticamente se cambia la base de los deflactores y en consecuencia el producto se computa a precios de 1962, se llega a una variación porcentual diametralmente opuesta.

Los nuevos índices deflactores (con base en 1962) serán los siguientes:

Indices deflactores (base 1962 = 100)

Sector primario	91.0	100
Sector secundario	33.3	100

Deflactando los productos nominales, por los índices anteriores se tiene:

Producto bruto real

(u.m. constantes)

Sector primario	1 100	1 600
Sector secundario	3 000	1 800
	4 100	3 400

Puede observarse un decrecimiento de 17 por ciento en el mismo período. Nótese que para los sectores, considerados en forma aislada, no ocurre esta incompatibilidad, la que sólo se presenta cuando se comparan sumas de sectores con crecimientos nominales distintos y cuyos respectivos deflactores también muestran variaciones desiguales. Los resultados serán tanto más contradictorios, en uno y otro caso, cuanto más distintos sean los índices deflactores y mayores sus variaciones. Lo que ocurre es que se está sumando valores que no son rigurosamente homogéneos, dados los cambios aritméticos de la base de los índices. Los resultados a que se ha llegado no deben sorprender si se piensa en la limitación de los mencionados cambios de base. Por otra parte la estructura del producto bruto nominal es distinta en ambos períodos. En el primer caso los dos índices deflactores son iguales a 100 en 1956 cuando tanto el Sector Primario como el Secundario aportan un 50 por ciento del producto bruto cada uno. En el segundo caso los dos índices deflactores son iguales a 100 en 1962 y el Sector Primario aporta con un 47 por ciento al producto y con 53 por ciento el Sector Secundario. Por esta razón no se presentan las inconsistencias anotadas para los sectores considerados en forma aislada, y sí, en cambio, se presentan al comparar sumas deflactadas por índices distintos. La inconsistencia será tanto mayor cuanto más distintas sean las variaciones de los deflactores.

3 La proyección sobre la base de índices de cantidad

Un procedimiento corrientemente utilizado para proyectar el producto bruto real por sectores, es el basado sobre índices de cantidad. Consiste en hacer variar el producto de un período dado conforme al índice de producción correspondiente. De esta manera, si se sabe que para 1962, el producto generado por la Industria fue de 5 000 unidades monetarias de ese año y se dispone del índice de producción industrial para los años 1962 a 1965, se puede suponer que habrá correspondencia entre las variaciones de dicho Producto sectorial y del índice de cantidad.

Años	Producto industrial (precios de 1962)	Indice de producción industrial (base 1962 = 100)	Estimación del producto industrial (precios de 1962)
1962	5 000	100	5 000
1963	-	198	5 400
1964	-	120	6 000
1965	-	135	6 750

Este tipo de estimaciones, para breves períodos de tiempo parece justificado, pero no hay que perder de vista que el índice de producción muestra más bien las variaciones de la producción global antes que las del producto (valor agregado); la metodología de estimación presentada tiene como supuesto la constancia en los coeficientes de insumo producto, constancia que puede no ser real en una época caracterizada por tanto cambio tecnológico.

G. Indices de comercio exterior

Para abordar el intercambio de bienes y servicios con el resto del mundo, es indispensable cuantificar indicadores acerca de precios, cantidades, valores, etc., relacionados con exportaciones e importaciones; el tipo de problemas que en estos casos debe enfrentarse, son los inherentes a los números índice, más algunas precauciones que deben tomarse por circunstancias que atañen al comercio exterior.

1. Indices de precios

En general es necesario homogeneizar los datos básicos, ya que no siempre están sujetos a criterios uniformes de valuación: precios CIF, FOB, precios expresados en monedas diferentes, valores nominales de retorno u otro tipo de valorización, en qué momento se considera que la importación o exportación está consumada: en tránsito, en aduana, etc.

Para el cálculo de índices de precios, puede utilizarse fórmulas de Laspeyres y Paasche. Sin embargo, en caso de utilizar la fórmula de Laspeyres, si algún producto deja de negociarse, implica suponer que su precio bajó a cero, lo que constituye un evidente falseamiento. Habrá que tener la precaución de tomar en cuenta para el cálculo de los productos sólo los negociados tanto durante el período base como durante el período dado. Por otra parte, no es necesario adoptar esta precaución, si se aplica la fórmula de Paasche, ya que en ésta se elimina automáticamente el producto que deja de negociarse y no quedará comprendido en ninguno de los factores $p_n\ q_n$ ni $p_n\ q_o$. Por eso, para determinar índices de precios se suele utilizar fórmulas de Paasche. En estadísticas de comercio exterior, se acostumbra designar a estos índices como índices de valor unitario, por el tipo de informaciones que se tienen. No es el precio y la cantidad de un artículo, si no el precio promedio o valor unitario que correspondería a un artículo proveniente de una partida de artículos no totalmente homogéneos. De este modo, si la importación de 100 automóviles de distintas marcas representa un valor CIF de 400 000 unidades monetarias, el precio promedio por automovil es de 4 000 u.m. En este sentido es que se prefiere designar a estos índices como índices de valor unitario en vez de precios, aunque metodológicamente no hay diferencias.

2. Indices de cantidad

Nuevamente cabe hacer aquí referencia a la denominación especial de índices de quantum que se les da en comercio exterior por razones similares a las anotadas en el punto anterior.

En este caso, si un producto deja de importarse o exportarse, quiere decir que la cantidad negociada bajó a cero, lo que queda ahora bien reflejado en la fórmula de Laspeyres. Por esa razón se acostumbra utilizarla en el cálculo de índices de quantum de comercio exterior. Además, utilizar fórmulas de Laspeyres o Paasche para quantum y valor unitario, respectivamente, garantiza el cumplimiento de la prueba de reversión de factores planteada por Fisher. Por ello, basta conocer el valor total de exportaciones o importaciones y uno cualquiera de los dos índices mencionados, para deducir inmediatamente el otro. Recuérdese que IPP x IQL = IV ; IPL x IQP = IV.

El trabajo que requiere el cálculo de estos índices induce a la utilización de muestras, en vez de indagaciones totales.

En el caso de índices de precios se supone que los precios de los artículos incluidos en la muestra, representan los precios de los no incluidos. Naturalmente, será necesario calcular los errores de muestreo correspondientes para hacer un uso racional del indicador. Por lo que a los índices de quantum se refiere, el hecho que en todos los períodos se disponga del valor total de las importaciones, permite cuantificar la representatividad que tenga la muestra utilizada en cada período mediante el cociente.

$$\frac{\Sigma\, p_n\, q_n \text{ (para la muestra)}}{\Sigma\, p_n\, q_n \text{ (para el total)}} = \text{Representatividad}$$

Posteriormente, si se desea calcular un índice de quantum de Laspeyres, se cuantifica $\Sigma q_n\ p_o$ en la muestra, y para proyectar al total, se divide por la representatividad (que implica la aplicación de una regla de tres simple). De esa manera se tiene la estimación del numerador de la fórmula para el total poblacional; en cuanto al denominador no hay problema alguno, ya que para cada período se dispone del total de importaciones y exportaciones. Si se desea calcular un índice de quantum de Paasche, será necesario ajustar al 100 por ciento la expresión del denominador de la fórmula, calculada con datos muestrales.

H. Algunos indicadores económicos

Se presentarán algunos indicadores de uso muy frecuente en la literatura económica. Si bien es cierto que la mayoría de ellos deberían ser analizados con mucha más profundidad desde el ángulo de la Contabilidad Social, es conveniente examinarlos desde el punto de vista estadístico.

1 Indice de la relación de términos del intercambio

Se define como el cociente entre el índice de precios de exportaciones y el índice de precios de importaciones, referidos ambos a la misma base. En consecuencia, este estadígrafo indica la evolución en el tiempo de la relación de precios registrada durante el período base; en otros términos, representa variaciones en la capacidad de compra de un volumen de exportaciones. Responde a la siguiente interrogante: ¿en que proporción ha aumentado o disminuido el número de unidades que deben exportarse para financiar la importación del mismo volumen de productos que se introducía al país durante el año base? Se trata de un concepto estrictamente relativo; no se puede concluir que la relación de intercambio sea buena o mala, sino mejor o peor que la del año base. El índice de la relación de precios de intercambio queda entonces definido como:

$$\text{IRI} = \frac{\text{Indice de precios de exportaciones}}{\text{Indice de precios de importaciones}} = \frac{\text{IPX}}{\text{IPM}}$$

2 Producto e ingreso bruto

Cuando estos conceptos se consideran en valores constantes, ambas magnitudes no coinciden. El producto real es una medida del valor de los bienes y servicios debidos al esfuerzo productivo interno. El ingreso real, está relacionado, además, con el intercambio con el exterior, puesto que parte de la producción de un país se exporta y parte de los insumos y bienes finales deben adquirirse en el exterior; por consiguiente estas transacciones con el exterior afectan, por la relación de precios de intercambio, la disponibilidad de bienes y servicios que satisfacen la demanda de la población. A medida que se deteriora la relación de intercambio, hay una transferencia al exterior de una parte del esfuerzo productivo interno. En valores constantes, la diferencia entre ambos conceptos, recibe el nombre de efecto de la relación de precios del intercambio, luego:

$$\boxed{\text{Ingreso bruto} = \text{Producto bruto} \pm \text{Efecto de la relación de precios de intercambio}}$$

En símbolos: $$\text{IB} = \text{PB} \pm \text{EFI}$$

donde

Efecto de la relación de precios del intercambio = Poder de compra de exportaciones - Quantum de exportaciones

Es decir $$\text{EFI} = \text{PEX} - \text{QX}$$

Además

Poder de compra de exportaciones = Quantum de exportaciones x Indice de la relación de intercambio

$$PEX = QX\ (IRI)$$

Dado que el producto del quantum de exportaciones y el índice de precios de exportaciones es equivalente al valor corriente de las exportaciones, el Poder de compra también puede definirse como

$$PEX = \frac{\text{Valor corriente de las exportaciones}}{\text{Indice de precios de importaciones}}$$

en otros términos, el valor nominal o corriente de las exportaciones queda deflactado por el índice de precios de las importaciones. Por otra parte el quantum de las exportaciones es el valor de las exportaciones a precios constantes

$$QX = \Sigma\, q_n\, p_o$$

donde la sumatoria se extiende a todos los rubros de exportación, es decir

$$QX = \frac{\text{Valor corriente de exportaciones}}{\text{Indice de precios de exportaciones}}$$

reemplazando las expresiones correspondientes en la fórmula de EFI se tiene

$$EFI = QX\ (IRI) - QX = QX\ [IRI - 1.00]$$

A continuación se cuantificarán estos conceptos mediante un ejemplo, para destacar su importancia. Supóngase que se tienen las siguientes informaciones:

	1960	1962	1964
a) Producto bruto (u.m. corrientes)	1 000	1 200	1 500
b) Deflactor implícito (base 1960 = 100)	100	115	140
c) Indice de precios de exportaciones	100	110	110
d) Indice de precios de importaciones	100	120	130
e) Valor corriente de las exportaciones	200	240	280

Para calcular el EFI, se tiene (efecto de la relación de Precios de Intercambio)

		1960	1962	1964
f) PB real (u.m. de 1960) (a/b)		1 000	1 043	1 071
g) IRI (base 1960 = 100) (c/d)		100	92	85
h) QX (u.m. de 1960) (e/c)		200	218	255
i) QX· IRI = PEX	(h.g)	200	201	217
j) EFI	(i-h)	-	-17	-38
k) YB real	(f-j)	1 000	1 026	1 033

3 Capacidad para importar

Cuando se piensa en términos de valores corrientes, la capacidad para importar no es otra cosa que el valor total de las exportaciones, más el ingreso neto de capitales extranjeros.

La capacidad para importar a precios constantes estará dada por

	Existencia de oro y divisas	EOD
+	Quantum de exportaciones	QX
+	Radicación de capitales extranjeros	RC
+	Efecto de la relación de precios del intercambio	EFI
	Capacidad total de pagos sobre el exterior	
-	Remesas de utilidades e intereses	
-	Salida de capitales extranjeros	
	Capacidad para importar	CPI

La capacidad para importar a precios constantes, también puede determinarse deflactando la capacidad para importar a precios corrientes por el índice de precios de importaciones. La capacidad para importar a precios constantes es:

$$CPI = QX + QX\,[\,IRI - 1.00\,] + RC \text{ (neta a precios constantes)}$$

$$CPI = QX\,[IRI] + \frac{RC \text{ (neta a precios corrientes)}}{IPM}$$

$$CPI = QX \cdot \frac{IPX}{IPM} + \frac{RC}{IPM}$$

$$CPI = \frac{\text{Valor corriente de las exportaciones} + \text{RC (neta a precios corrientes)}}{\text{Indice de precios de importaciones}}$$

$$CPI = \frac{\text{Capacidad para importar a precios corrientes}}{\text{Indice de precios de importaciones}}$$

4 Tipo de cambio de paridad

Como una aplicación de números índice, es conveniente tratar la forma de convertir monedas de distintos países a unidades homogéneas con propósitos de comparación. Utilizar para ello los tipos de cambio oficiales puede originar distorsiones serias, en la medida que exista más de un tipo de cambio (discriminación de áreas cambiarias), o que el tipo de cambio único esté sobre o subvaluado.

Una alternativa teórica consistiría en la determinación de una canasta de productos, resultando el tipo de cambio entre dos monedas la relación de valores que sería necesario gastar para adquirir dicha canasta. Este método tiene inconveniente de tipo práctico; la determinación de los artículos que conformarían la canasta significa un serio problema, sobre todo si se piensa en la enorme variación de las preferencias de los consumidores en los distintos países. Sin embargo, con algunas restricciones, este método es utilizado en la práctica.

Un método que puede ser utilizado con alguna ventaja es la proyección de un tipo de cambio base, en un período en el que no se haya advertido sobre o subvaluaciones monetarias, sin cambios múltiples y, en general, sin accidentes que descalifiquen a ese período base.

La aludida proyección se efectúa sobre la base de las modificaciones en la relación de precios de los dos países cuyo tipo de cambio se pretende determinar.

Sean los países A y B; se desea estimar el número de unidades monetarias de A, por unidad monetaria de B. El tipo de cambio de paridad en un año cualquiera n, estará dado por

$$\text{Tipo de cambio de paridad año n} = \text{Tipo de cambio año base} \cdot \frac{\text{Deflactor implícito de A}}{\text{Deflactor implícito de B}}$$

Ejemplo: Supóngase que el tipo de cambio base en el año 0, fue de 5 u.m. de A por 1 u.m. de B, siendo los deflactores implícitos los siguientes:

Año	D.I. país A	D.I. país B
0	100	100
1	120	110
2	150	120
3	200	140
4	300	150

El tipo de cambio de paridad para los años siguientes será:

Año	Relación de deflactores DI_A/DI_B	Tipo de cambio de paridad
0	100.0	5.00
1	109.1	5.45
2	125.0	6.25
3	142.9	7.14
4	200.0	10.00

Evidentemente el método da estimaciones que serán tanto más confiables en la medida que el tipo de cambio base haya sido elegido adecuadamente, y los deflactores implícitos sean representativos de las variaciones de precios en ambos países.

5. Transferencias implícitas

El hecho que en la actividad económica se presenten transacciones intersectoriales donde cada sector tiene precios distintos, origina cierto tipo de transferencias de producto de un sector a otro, implícitas en las transacciones que efectúan cuando se contabilizan a precios constantes.

Los sectores cuyos precios crecen menos que el promedio general de precios representado por el deflactor implícito, están transfiriendo parte de su producto hacia aquellos sectores que tienen precios que crecen más que el promedio de precios.

Evidentemente que la suma algebraica de las transferencias será nula, por cuanto la ganancia de unos sectores tiene como contrapartida la pérdida de otros.

Se definen las transferencias implícitas para cada sector, como la diferencia entre la producción valorizada a precios del sector, y esa misma producción valorizada a precios promedios representados, como se dijo, por el deflactor implícito:

Transferencias implicitas = Producción a precios corrientes - Producción a precios constantes x Deflactor implícito

en símbolos $$T\ Imp = p_n\, q_n - p_o\, q_n\, (DI)$$

donde n y o indican el período que se calcula y el período base respectivamente.

$$\sum_{h=1}^{L} T\ Imp\,(h) = \Sigma\, p_n\, q_n - \Sigma\, p_o\, q_n\, (DI) = 0$$

Donde L es el número de sectores considerados, h = 1, 2, 3, ... L

$$\sum_{h=1}^{L} T\ Imp\,(h) = \text{Producción nominal} - \text{Producción real} \times \frac{\text{Producción nominal}}{\text{Producción real}} = 0$$

Ejemplo:

Sectores	Producción nominal año 0	Producción nominal año 1	Indice de cantidad base 0 = 100	Producción real	Indice de precios base 0 = 100
1	500	1 080	120	600	180
2	600	800	110	660	121
3	300	350	100	300	117
4	400	420	90	360	117
		2 650		1 920	

El deflactor implícito para el año 1 será:

$$DI = \frac{\text{Producción nominal año 1}}{\text{Producción real año 1}} = \frac{2\ 650}{1\ 920} = 138$$

Las transferencias implícitas por sector:

Sectores	Producción nominal $p_1\, q_1$	Producción real (DI) $p_o\, q_1\, (DI)$	Transferencias implícitas
1	1 080	828	+252
2	800	911	-111
3	350	414	-64
4	420	497	-77
	2 650	2 650	

I. Etapas de la construcción de números índice

Es interesante e ilustrativo seguir cada uno de los pasos para la confección de indicadores acerca de precios o cantidades. Los pasos que se enumerarán estarán referidos al diseño de un índice de costo de vida por constituir éste un caso bastante general y complejo.

1. Objetivo del índice

Es fundamental calificar con toda precisión el objetivo del indicador. En el caso de un índice de costo de vida, es necesario determinar a quiénes estará referido: si a la población en su conjunto, a una ciudad, a profesionales, a obreros, a campesinos, etc. De esto dependerá qué tipo de artículos conformarán la muestra de productos que componen el índice. Es imprescindible no descuidar el objetivo básico: los usos principales que tendrá el indicador.

2. Determinación de la estructura de consumo

Es conveniente, cuando se calcula un índice de costo de vida, clasificar los gastos: alimentación, vestuario, vivienda, varios, etc. De esta manera es posible calcular índices, por separado, para cada componente. Este desglose aparte que permite realizar análisis de las variaciones de precios en forma más detallada, es muy útil en la selección de deflactores adecuados. Estos índices desglosados pueden combinarse para obtener el índice general. Las ponderaciones de cada componente se establecen mediante una muestra. Se selecciona una muestra de unidades familiares para el caso del costo de vida, obtenida de la población para la cual se confecciona el índice, por ejemplo la población de obreros y empleados de una ciudad. Esta muestra, en lo posible, debe ser aleatoria y de un tamaño que asegure fidelidad y garantice confianza. En cada una de las unidades elegidas para la muestra, se lleva un registro de los gastos en bienes y servicios por tipo de bien o servicio, donde se recopilan los precios pagados y las cantidades consumidas. Vale la pena llevar este registro durante un tiempo prolongado, por lo general un año, para que permita captar las variaciones en los consumos durante las diferentes estaciones. De esta manera es posible llegar a obtener ponderaciones por componentes y por tipos de bienes y servicios.

3. Selección de artículos

En las anotaciones que se hacen en los registros o "libretas de consumo" acerca de los distintos bienes que se consumen, en general aparece una gran cantidad de productos, aunque muchos de ellos responden a consumos esporádicos o accidentales. Es preciso seleccionar los productos más importantes, de consumo habitual y representativos de las preferencias de los consumidores. Para seleccionar estos productos y servicios no es conveniente un método aleatorio, pues es preferible decidir qué productos serán considerados en el índice, tomando en cuenta el volumen del gasto efectuado en cada bien o servicio. En esa forma se llegará a establecer una lista de 100 a 300 productos cuya importancia relativa o ponderación se tiene tabulada.

4. Formas de valuación

Otra etapa importante durante el proceso que permite confeccionar índices de precios, es la decisión acerca de los tipos de precios que se considerará. Sabido es que los precios varían enormemente según el lugar donde se adquieren los productos; por otra parte es corriente que para artículos considerados de primera necesidad, hayan precios oficiales y precios reales con grandes diferencias entre sí. Para el caso de índices de costo de vida los precios debieran tomarse en los lugares donde el consumidor adquiere los productos: almacenes, ferias, etc. Sobre el particular es necesario tomar decisiones previas en forma categórica.

5. Variaciones de calidad de los bienes y servicios

Otro aspecto fundamental es la especificación precisa, hasta donde sea posible, de la calidad de los productos, de manera que se pueda controlar a través del tiempo la invariabilidad de sus características. Es muy frecuente, sobre todo en los artículos controlados, que se "incrementen" los precios reales, permaneciendo fijos los precios nominales, por una disminución de la calidad de los productos.

6. Base del índice

En general cuando se planteó el tema se advirtió sobre la necesidad de elegir como base un período donde estuvieran ausentes las modificaciones y circunstancias violentas. Ahora debe agregarse que, en el caso de índices de costo de vida, el período base debe ser modificado, toda vez que la estructura de consumo haya cambiado significativamente. En los tiempos actuales, cuando las innovaciones tecnológicas cambian con extraordinaria rapidez, determinando modificaciones en las preferencias de los consumidores, la base de un índice de costo de vida no debería tener una antigüedad superior a los 6 a 8 años.

Sin embargo, en las fórmulas de cálculo, para evitar cambios de base muy seguidos, modificaciones que significan altos costos, se contempla la posibilidad de introducir nuevos artículos, toda vez que se registren cambios en las preferencias de los consumidores.

7. Elección de los métodos de cálculo

En general es necesario tomar decisiones sobre la fórmula de cálculo; corrientemente, cuando se trata de índices de costo de vida se acostumbra utilizar fórmulas de Laspeyres, porque implica considerar constantes las ponderaciones del período base. La utilización de una fórmula de Paasche significaría un esfuerzo muy grande, ya que en cada período (generalmente cada mes) habría que volver a ponderar "hacia atrás", aparte que sería necesario calcular periódicamente estas ponderaciones a medida que pasa el tiempo.

La decisión acerca de qué fórmula utilizar estará condicionada, en primer lugar, al objetivo que se persigue con el índice y a la factibilidad de su diseño desde el punto de vista del costo y la oportunidad con que se entreguen los resultados.

La Dirección de Estadística y Censos de algunos países calcula el índice de costo de vida sobre la base de una modificación a la fórmula de Laspeyres.

$$I_i = I_{i-1} \left(\frac{\Sigma\, q_o\, p_{i-1} \left(\frac{p_i}{p_{i-1}}\right)}{\Sigma\, q_o\, p_{i-1}} \right)$$

Esta fórmula implica el cálculo del índice en un período i, sobre la base del índice en el período anterior, afectándolo por la variación que acusen los precios. La fórmula tiene la ventaja que se pueden introducir nuevos artículos según sus especificaciones, o cambiar la fuente de información.

TEMAS DE DISCUSION

Indique si las siguientes afirmaciones son ciertas o falsas, y justifique su opinión.

1. Poder de compra, valor real y quantum físico, son conceptos equivalentes.

2. La relación de intercambio es desfavorable para los países subdesarrollados, porque los precios de sus importaciones son mayores que los precios de sus exportaciones.

3. La fórmula de Laspeyres no proporciona indicaciones correctas, cuando se trata de averiguar la variación de los valores unitarios de las exportaciones.

4. Deflactar una serie de valores nominales por un índice de precios de Laspeyres, equivale a proyectar el valor del año base ($\Sigma p_o q_o$) por un índice de cantidades de Paasche.

5. Los siguientes datos son factibles. (Los índices tienen base 1960.)

   $DI_{1967} = 150$

   Transferencias implícitas del sector primario al resto de la economía 200. Indice de precios en 1967 del sector primario 140.

6. El deflactor implícito es único, independiente de la restricción planteada,

7. Las pruebas planteadas por Fischer significan una seria limitación a la utilización de índices de Laspeyres y Paasche.

8. Los siguientes datos son consistentes

   EFI = -200 QX = 4 000 IPX = 150 IRI = 80

9. Si los índices de precios de Laspeyres y Paasche coinciden en valor numérico, para cierto período, quiere decir que la estructura de ponderaciones es exactamente igual a la del período base.

10. Los índices de precios de Laspeyres no pueden tomar valores negativos.

11. El índice de valor debe ser siempre mayor que el índice de precios.

12. Los siguientes datos son consistentes.

    PEX = 3 000 QX = 2 400 IRI = 80

13. Si el producto nominal coincide en tres períodos con el producto real, quiere decir que los precios de todos y cada uno de los bienes y servicios han permanecido invariables en los tres períodos.

14. En un período caracterizado por una completa estabilidad de precios, los índices de Laspeyres y Paasche para precios, coincidirán necesariamente en valor numérico.

15. Si una economía se divide en dos sectores: primario y secundario; y el índice de precios de productos primarios coincide con el deflactor implícito, quiere decir que las transferencias implícitas de ingresos, son nulas para ambos sectores.

16. La magnitud del efecto de la relación de precios del intercambio, es independiente del período elegido como base de los índices deflactores.

17. Para que el índice de producción industrial refleje cambios reales, debiera ser deflactado por un índice de precios de bienes industriales.

18. El índice de precios de Laspeyres es el que más se presta para el cálculo de variaciones en el costo de vida, principalmente porque toma como ponderaciones, las cantidades de un período base considerado como normal.

19. Los índices de base variable, tienen la desventaja de que no puede establecerse comparaciones entre ellos.

20. Si el índice de la relación de precios del intercambio es inferior a 100, está reflejando una situación desfavorable para el país.

21. Si un obrero en 1964 tiene un sueldo superior en 30% al de 1963 y por otra parte, el índice de precios al consumidor para 1963 con base en 1964 es de 70, quiere decir que la situación del obrero en cuanto a poder de compra no ha variado.

PROBLEMAS PROPUESTOS

1. Una fábrica produce 2 bienes (A y B). Si se tienen los siguientes datos incompletos.

Año	Ventas totales en u. m. corrientes	Bien "A" Precio	Bien "A" Cantidad	Bien "B" Precio	Bien "B" Cantidad
1961	5 000	10	400	100	?
1966	9 000	15	?	400	?

Calcule un índice de cantidad de las ventas de esta fábrica para 1966, con base 1961.

2. Dadas las siguientes informaciones:

Producto nominal del país A en 1968	6 000
Producto nominal del país B en 1968	180 000
Unidades monetarias de A por una unidad monetaria de B en 1962	10
Producto real de A en 1968 en u. m. de 1964	4 000
Producto real de B en 1968 en u. m. de 1962	120 000

El índice general de precios con base variable para el país A, se presenta a continuación:

	País A
1961	-
1962	110
1963	120
1964	110

Se le pide estimar el tipo de cambio de paridad en 1968. Señale brevemente las limitaciones del método.

3. Una economía se divide en tres sectores: primario, secundario y terciario; y se dispone de los siguientes datos:

	Producto bruto (u.m.corrientes)	
	1964	1967
Sector primario	2 000	4 000
Sector secundario	5 000	6 000
Sector terciario	3 000	5 000

	Producto bruto (u. m. constantes de 1964)	
	1964	1967
Sector primario	2 000	2 500
Sector secundario	5 000	5 000
Sector terciario	3 000	4 000

Se pide determinar las transferencias implícitas de ingreso entre los sectores.

4. Los índices de precios y producción industrial han sido los siguientes:

Indices	1958	1959	1960	1965
Precios	80	100	150	300
Producción	*100*	*110*	*120*	*125*

Si se sabe que el producto de la industria fue en 1960 de 800 millones de u. m. se pide:

a) Estimar el producto nominal para 1965
b) Estimar el producto real para 1959, en u. m. de 1965
c) Señalar las limitaciones de los métodos utilizados.

EJERCICIOS

1. Para los artículos A, B, C y D, se tienen los siguientes precios y cantidades en los años que se indican:

Años	A		B		C		D	
	p	q	p	q	p	q	p	q
1959	10	12	4	15	1	10	30	10
1960	12	12	4	15	1	15	30	15
1961	15	20	5	10	2	20	50	20
1962	20	20	5	10	2	30	50	20
1963	30	30	6	15	2	50	60	20

Calcule para todo el período:

a) Un índice de precios de Laspeyres con base 1959;

b) Un índice de cantidad de Paasche con base 1959; y

c) Un índice de valor.

2. El índice de base variable de los precios de los productos agropecuarios muestra el siguiente comportamiento:

Años	Indice
1957	-
1958	104
1959	102
1960	108
1961	120
1962	150
1963	130

Calcule el índice tomando 1960 como base fija.

3. Para el índice de producción industrial se tienen los siguientes datos:

Años	Indice de prod. indust. base 1955	Indice de prod. indust. base 1960
1955	100	-
1956	106	-
1957	114	-
1958	120	-
1960	130	100
1961	-	112
1962	-	120
1963	-	130
1964	-	145

Se le pide que empalme ambos índices y señale claramente las limitaciones del método.

4. Una economía ha sido dividida en tres sectores y sus valores agregados se presentan a continuación (en millones de unidades monetarias corrientes).

	1956	1957	1958	1959	1962	1965
Agricultura	2 000	2 600	3 100	4 000	6 000	8 000
Servicios y otros	4 000	5 000	5 800	7 900	9 000	10 000
Industria	1 600	2 000	2 800	3 800	5 500	6 000
Producto	7 600	9 600	11 700	15 700	20 500	24 000

Por otra parte, se dispone de los siguientes índices para los mismos años.

Indice costo vida	80	100	130	150	190	210
Indice de precios al por mayor	100	120	140	170	210	220
Indice de producción industrial	100	110	125	140	170	200
Indice de precios de la construcción	90	100	120	160	200	210
Indice de sueldos y salarios	100	125	140	160	180	200

Se le pide que:

a) Calcule el producto bruto real para los seis años en unidades monetarias de 1956 y luego de 1962. Compare los resultados.

b) Calcule el índice deflactor implícito del producto bruto y explique su significado.

5. Una familia incrementa sus ahorros nominales en 360 por ciento entre 1955 y 1962. Si en el año 1962 dicha familia desea invertir sus ahorros en materiales de construcción, sólo podrá adquirir una cantidad 20 por ciento mayor que en 1955; en cambio, si destina sus ahorros a la compra de productos

agrícolas, no alcanzará a obtener la cantidad que habría comprado en 1955. Pero al mismo tiempo, si invirtiera sus ahorros para comprar ambos tipos de bienes, de tal modo que gastase igual suma en cada uno, podrá adquirir una cantidad 5 por ciento superior al año 1955. Se le pide calcule el índice de precios de productos agrícolas para 1955 con base 1962. ¿Qué limitaciones tendría el método empleado?

6. Se disponen de antecedentes, sobre la base de una muestra de cuatro productos, de las exportaciones de cierto país.

Años	Muestra de productos de exportación A		B		C		D		Valor total de las exportaciones
	p	q	p	q	p	q	p	q	
1958	2	10	20	6	4	12	4	2	250
1959	3	10	20	8	4	14	4	2	320
1960	3	10	25	6	4	16	6	2	420
1961	4	10	30	8	5	20	6	3	450
1962	4	12	30	8	5	20	6	3	480
1963	4	15	25	8	6	25	8	2	500

Se le pide que estime

a) Un índice de cantidad de exportaciones para el total de éstas con base 1958.
b) Un índice de Paasche para los precios de exportación con base 1959.
c) Si el índice de precios de importaciones tiene el siguiente comportamiento:

	1958	1959	1960	1961	1962	1963
IPM	100	120	160	240	310	430

Se le pide calcule el índice de relación de términos de intercambio con base 1959.

7. Si los precios de los cigarrillos se incrementan en 70 por ciento y, como consecuencia, el índice de costo de vida sube en 1,8 por ciento, ¿qué ponderación dentro del costo de vida tiene este bien?

SOLUCION DE EJERCICIOS

1. a) $IPL = \dfrac{\Sigma\, p_n q_o}{\Sigma\, p_o q_o}$

Años	A $p_n q_o$	B $p_n q_o$	C $p_n q_o$	D $p_n q_o$	Total $\Sigma\, p_n q_o$	Indice base 1959=100
1959	120	60	10	300	490	100.0
1960	144	60	10	300	514	104.9
1961	180	75	20	500	775	158.2
1962	240	75	20	500	835	170.4
1963	360	90	20	600	1 070	218.4

b) $IQP = \frac{\Sigma q_n p_n}{\Sigma q_o p_n}$

Año	A $p_n q_n$	B $p_n q_n$	C $p_n q_n$	D $p_n q_n$	Total $\Sigma p_n q_n$	Indice
1959	120	60	10	300	490	100.0
1960	144	60	15	450	669	130.2
1961	300	50	40	1 000	1 390	179.4
1962	400	50	60	1 000	1 510	180.8
1963	900	90	100	1 200	2 290	214.0

c) El índice de valor puede obtenerse multiplicando los dos índices antes encontrados; y en virtud de la prueba de reversión de factores:

Años	IPL	IQP	IV
1959	100.0	100.0	100.0
1960	104.9	130.2	136.6
1961	158.2	179.4	283.8
1962	170.4	180.8	308.1
1963	218.4	214.0	467.4

2. Es necesario "desencadenar" el índice previamente:

		Indice base fija 1957=100
1957	100	100.0
1958	100 . 104	104.6
1959	100 . 104 . 102	106.1
1960	100 . 104 . 102 . 108	114.6
1961	100 . 104 . 102 . 108 . 120	137.6
1962	100 . 104 . 102 . 108 . 120 . 150	206.2
1963	100 . 104 . 102 . 108 . 120 . 150 . 130	268.1

Luego será necesario cambiar de base, dividiendo por el valor de 1960 (114.6).

Años	Indice base fija 1960=100
1957	87.3
1958	90.8
1959	92.6
1960	100.0
1961	120.0
1962	179.9
1963	233.9

3. Para empalmar el índice con base 1955, bastará aplicar la siguiente regla de tres:

$$\begin{array}{cc} 130 & 100 \\ I_{61.55} & I_{61.60} \end{array}$$

y de igual forma para todos los años. Para empalmar hacia atrás el otro índice, el procedimiento es similar. Los índices empalmados se muestran a continuación:

	Indice base 1955	Indice base 1960
1955	100.0	76.9 $(100 \cdot \frac{1.00}{1.30})$
1956	106.0	81.5 $(100 \cdot \frac{1.06}{1.30})$
1957	114.0	87.7 $(100 \cdot \frac{1.14}{1.30})$
1958	120.0	92.3 $(100 \cdot \frac{1.20}{1.30})$
1960	130.0	100.0
1961	145.6 (130 . 1.12)	112.0
1962	156.0 (130 . 1.20)	120.0
1963	169.0 (130 . 1.30)	130.0
1964	188.5 (130 . 1.45)	145.0

La limitación estriba en el hecho que a lo largo de cualquiera de los dos índices empalmados aparecen implícitamente dos bases; cuanto más diferentes sean los supuestos de una y otra base, tanto más defectuoso será el empalme.

4. a) El primer problema por resolver es la elección de los deflactores adecuados, es decir que exista relación entre el concepto que se desea deflactar y el índice elegido como deflactor. El segundo problema consiste en expresar los deflactores elegidos, con la misma base para disponer de unidades homogéneas. Admítase que después de un estudio se ha decidido deflactar el valor agregado de la agricultura por el índice de costo de vida, el de servicios por un índice de sueldos y salarios y en industria se haría la proyección a través del índice de cantidad respectivo.
Los índices con base 1956 serán los siguientes:

	1956	1957	1958	1959	1962	1965
Deflactor agricultura	100	125	163	188	238	263
Deflactor servicios	100	125	140	160	180	200
Indice de produc. ind.	100	110	125	140	170	200

Los valores agregados reales, en u.m. de 1956, serán:

	1956	1957	1958	1959	1962	1965
Agricultura	2 000	2 080	1 902	2 128	2 521	3 042
Servicios	4 000	4 000	4 143	4 938	5 000	5 000
Industria	1 600	1 760	2 000	2 240	2 720	3 200
Producto real	7 600	7 840	8 045	9 306	10 241	11 242

Para obtener los valores agregados reales, en u. m. de 1962 será necesario expresar los deflactores con base 1962.

	1956	1957	1958	1959	1962	1965
Deflactor agricultura	42	53	68	79	100	111
Deflactor servicios	56	69	78	89	100	111
Indice de prod. ind.	59	65	74	82	100	118

Los valores agregados reales en u. m. de 1962, serán:

	1956	1957	1958	1959	1962	1965
Agricultura	4 762	4 906	4 559	5 063	6 000	7 207
Servicios	7 143	7 246	7 436	8 876	9 000	9 009
Industria	3 245	3 575	4 070	4 510	5 500	6 490
Producto real	15 150	15 727	16 065	18 449	20 500	22 706

Los resultados evidentemente serán distintos, ya que se expresan en unidades distintas; incluso, la tasa de crecimiento del producto real, por las razones citadas, será diferente.

b) Para calcular el deflactor implícito del producto, se aplicará la relación.

$$DI = \frac{\text{Producto nominal}}{\text{Producto real}}$$

En cuanto al producto real, puede tomarse cualquiera de los ya calculados, según la base del deflactor implícito que se desee. De esta manera, si se pretende que la base sea el año 1956, se tendrá:

	1956	1957	1958	1959	1962	1965
Producto nominal	7 600	9 600	11 700	15 700	20 500	24 000
Producto real (u.m.de 1956)	7 600	7 840	8 045	9 306	10 241	11 242
Deflactor implícito	100.0	122.4	145.4	168.7	200.2	213.5

5. Con los datos presentados pueden construirse los siguientes índices:

	Indice de valor ahorro nominal	Indice de cantidad construcción	Indice de cantidad mixto
1955	100	100	100
1962	460	120	105

En virtud de la prueba de reversión de factores pueden obtenerse los siguientes índices de precios:

	Indices de precios construcción	Indices de precios mixto
1955	100	100
1962	383	438

El índice de precios mixto, será una combinación lineal de los índices de precios parciales, ponderados por el gasto:

$$IP\ (\text{mixto}) = IP\ (\text{const})\ W_c + IP\ (\text{agric})\ W_a$$

pero $W_c = W_a = \frac{1}{2}$

$$438 = \frac{383 + IP\ (\text{agric})}{2}$$

$$IP\ (\text{agric}) = 876 - 383 = 493$$

IP (agric) base 1955=100	IP (agric) base 1962=100
100.0	20.3
493.0	100.0

Los supuestos admitidos fueron los siguientes:

1. Que los índices cumplen con la prueba de reversión de factores;
2. Que el índice de precios mixto se obtiene a partir de un promedio ponderado de los índices parciales;
3. Que la estructura del índice de precios agrícolas es la misma entre los años 1955 y 1962.

6. a) Se calculará un índice de cantidad de Laspeyres

$$IQL = \frac{\Sigma\, q_n p_o}{\Sigma\, q_o p_o}$$

Años	A $q_n p_o$	B $q_n p_o$	C $q_n p_o$	D $q_n p_o$	Total $\Sigma\, q_n p_o$	Repre-sent. %	$q_n p_o$ Ajust. 100%	Indice base 1958=100
1958	20	120	48	8	196	78.4	250	100
1959	20	160	56	8	244	79.8	306	122
1960	20	120	64	8	212	60.9	348	139
1961	20	160	80	12	272	88.4	308	123
1962	24	160	80	12	276	84.6	326	131
1963	30	160	100	8	298	85.2	350	140

b) $IPP = \frac{\Sigma\, p_n q_n}{\Sigma\, p_o q_n}$ donde 0: 1959

Años	A P_oq_n	B P_oq_n	C P_oq_n	D P_oq_n	ΣP_oq_n	ΣP_nq_n	Indice
1958	30	120	48	8	206	196	95.1
1959	30	160	56	8	254	254	100.0
1960	30	120	64	8	222	256	115.3
1961	30	160	80	12	282	398	141.1
1962	36	160	80	12	288	406	141.0
1963	45	160	100	8	313	426	136.1

c) El índice de la relación de intercambio está dado por

$$IRI = \frac{IPX}{IPM}$$

Para disponer de los dos índices de igual base, se efectuará un cambio de base aritmético en el IPM.

	1958	1959	1960	1961	1962	1963
IPX	95.1	100.0	115.4	141.1	141.0	136.1
IPM	85.0	100.0	133.3	200.0	258.3	358.3
IRI	111.9	100.0	86.6	70.6	54.6	38.0

7. Se considerarán dos grupos : uno compuesto por los cigarrillos (C) y otro compuesto por todo el resto de los bienes (R)

$$IC\ (W_C) + IR\ (W_R) = 1.018$$

$$1.7\ W_C + 1.0\ W_R = 1.018$$

$$\text{pero } W_C + W_R = 1$$

$$1.7\ W_C + 1.0\ (1 - W_C) = 1.018$$

$$0.7 W_C = 0.018$$

$$W_C = \frac{0.018}{0.7} = 0.026$$

Los cigarrillos tienen un 2.6% de ponderación dentro del índice de costo de vida.

BIBLIOGRAFIA

Cansado, Enrique, Curso de Estadística General, Comisión de Educación Estadística del Instituto Interamericano de Estadística, Biblioteca Interamericana de Estadística Teórica y Aplicada, Rosario, Argentina, 1958.

Cochran, William G., Sampling Techniques, Nueva York, John Wiley and Sons, 5ª ed., 1966.

Croxton, Frederic y Cowden, Dubley, Estadística General Aplicada, trad. de Teodoro Ortiz y Manuel Bravo, México, Fondo de Cultura Económica, 1962.

Hoel, Paul G., Introducción a la Estadística Matemática, trad. de Enrique E. Dieulefait, Biblioteca Interamericana de Estadística Teórica y Aplicada, Rosario, Argentina, 2ª ed., 1955.

Mills, F. Cecil, Métodos aplicados a la economía y los negocios, trad. de Juan Ruiz Magan y José Ruiz Rubio, Madrid, Ed. Aguilar, 1960.

Ostle, Bernard, Estadística Aplicada, trad. de Dagoberto De La Serna Valdivia, México, Editorial Limusa-Wiley S. A., 1965.

Ríos, Sixto, Métodos Estadísticos, Madrid, Ed. del Castillo S. A., 5ª ed., 1967.

Vusković, Pedro, Apuntes de Estadística, Santiago, Escuela de Economía Universidad de Chile, Edición mimeografiada, 1959.

Yamane, Taro, Elementary Sampling Theory, New Jersey, Prentice-Hall Inc., 1967.

Yamane, Taro, Statistics, An Introductory Analysis, Nueva York, Harper and Row Publishers, 1964.

Yule, G. U. y Kendall, M. G., Introducción a la Estadística Matemática, trad. de José Ros Jimeno, Madrid, Ed. Aguilar, 1954.

EL INSTITUTO

El Instituto Latinoamericano de Planificación Económica y Social (ILPES) es un organismo autónomo creado bajo la égida de la Comisión Económica para América Latina (CEPAL) y establecido el 1º de julio de 1962 en Santiago de Chile como proyecto del Programa de las Naciones Unidas para el Desarrollo (Fondo Especial) con amplio apoyo del Banco Interamericano de Desarrollo (BID). Cuenta además con aportaciones directas de los gobiernos latinoamericanos y de otros organismos internacionales y privados. El objeto principal del Instituto es proporcionar, a solicitud de los gobiernos, servicios de capacitación y asesoramiento en América Latina y realizar investigaciones en diversos campos económicos y sociales. Desde su fundación, el Instituto ha venido ampliando y profundizando la acción iniciada por la CEPAL en materia de planificación merced al esfuerzo conjunto de un grupo de economistas y sociólogos dedicado por completo al estudio y búsqueda de soluciones de los problemas que preocupan en la actualidad a los países de esta parte del mundo.

ESTOS CUADERNOS

Con el nombre común de Cuadernos del Instituto Latinoamericano de Planificación Económica y Social se inician diversas publicaciones, que abrigan en su conjunto un mismo propósito. Por el momento los cuadernos se compondrán de tres series distintas que declaran en su título la naturaleza de su contenido: apuntes de clase; anticipos de investigación, y manuales operativos.

Con la publicación de sus cuadernos el Instituto persigue informar a un público más amplio de algunas de sus tareas de investigación y de enseñanza que no pueden menos de modificarse continuamente, ya sea por nuevas orientaciones de la ciencia o por la aparición de problemas antes desconocidos. Esa información quiere hacerse de tal modo que constituya invitación a un diálogo en el que se apoye realmente una auténtica cooperación intelectual. Por ello, es indudable que la mejor manera de alcanzar esas metas es hacer comunicables algunas de las tareas del Instituto en sus etapas de formación. Se trata, pues, de trabajos o fragmentos de trabajos que no pretenden en modo alguno la plena madurez de forma o contenido y que, por consiguiente, en uno u otro plano han de ser modificados en su día de acuerdo en lo posible -y ese sería el ideal que pretenden alcanzar los cuadernos- con el consenso científico suscitado por el diálogo y la discusión.

Los apuntes de clase dicen por sí mismos lo que la serie significa: lecciones o fragmentos de lecciones que pueden ser útiles no sólo al becario de los cursos de capacitación del Instituto y al estudiante de otros centros de enseñanza, sino al interesado en determinadas cuestiones no obstante las insuficiencias que necesariamente lleva consigo la expresión académica. Los anticipos de investigación tratan de hacer viable el estado de esfuerzos de conocimiento en sus etapas iniciales y que, sin embargo, contienen ya en ciernes el horizonte de la investigación perseguida. Los manuales operativos se conciben como instrumentos de trabajo que faciliten la acción de los organismos gubernamentales, y en general de los especialistas en ese campo, en tareas prácticas de la planificación muchas veces de carácter urgente.

En consecuencia, se presenta estos cuadernos al público con una conciencia crítica de todas sus limitaciones por ver precisamente en ella el mejor estímulo para la tarea que el Instituto tiene por delante.

IMPRESO EN LA SECCION DE REPRODUCCION DE DOCUMENTOS DE LA CEPAL