

# LA IMPORTANCIA DE LA ESTADISTICA EN LA INVESTIGACION SOCIOLOGICA

## A) UTILIDAD DE LA ESTADISTICA

### I) *Introducción*

El tema de este trabajo es susceptible de un doble tratamiento, que es el que se va a intentar aquí.

Por un lado, poner de manifiesto las características y el aporte de la Estadística, disciplina que se hace imprescindible si la Sociología se afirma como ciencia rigurosa.

Por el otro, dejar claramente expresado el carácter instrumental de la Estadística, para enmarcarla dentro del cuadro de una elaboración teórica y metodológica que la trasciende y debe guiarla necesariamente, única garantía para no caer en la cuantificación a todo trance de datos que muchas veces carecerán de sentido.

En Sociología —como en toda otra ciencia— al emprender una investigación se hace necesario ir en busca de datos sobre la realidad que se pretende estudiar. Cuanto mayor sea la cantidad de datos que tengamos se tendrá una imagen más completa de la realidad estudiada. Es necesario, pues, recoger el mayor número posible de datos para no ser imprecisos ni caer en prejuicios.

Si los datos que tenemos sobre la realidad son insuficientes, nuestras propias tendencias emotivas harán sentir su peso desviando nuestros juicios, restándoles exactitud y veracidad.

Es sumamente peligroso abocarse al estudio de cuestiones hacia las cuales tenemos una reacción emocional o un juicio previo, cuestiones tales como las discriminaciones de todo tipo, los prejuicios, las ideologías.

Para garantizar la mayor objetividad de nuestro estudio y no invalidarlo de antemano es necesario tender a conseguir una sólida base de datos concretos.

Estos datos acerca de la realidad social son obtenidos por medio de mediciones y recuentos. Medir y contar las cosas se hace, por lo tanto, imprescindible para garantizar precisión y objetividad.

Si quisiéramos, por ejemplo, detectar en la población de Rosario la opción política en relación con la estratificación tendríamos que encuestar a todos los habitantes de Rosario de nacionalidad argentina y mayores de 18 años, es decir, a todos los electores.

Pero, ¿cómo entrevistar a todos los electores de la ciudad de Rosario? Si tuviéramos que aplicar un cuestionario a todos ellos nos resultaría muy costoso y nos demandaría mucho tiempo.

Necesitamos un método de investigación que nos simplifique el recuento y este método lo encontramos en la Estadística, la que con un gasto menor y en menos tiempo nos suministrará ciertos datos a partir de los cuales podremos inferir aquellos otros que buscábamos.

Apenas comenzado nuestro intento de investigación surge la necesidad de la Estadística como disciplina auxiliar, y tenemos ya una de las tareas de la Estadística que la hace irremplazable y que justifica su importancia: la simplificación de los datos.

Paul F. Lazarsfeld en la Introducción a "Say it with figures" de Zeisel nos dice que: "La vida social moderna se ha vuelto demasiado complicada para ser percibida por observación directa... tales rasgos pueden ser comprendidos por aquellos que pueden leer tablas estadísticas o pueden contratar alguien para que se las interprete". Y más adelante agrega: "La gran complejidad de los eventos sociales requiere un lenguaje de cantidad".

## II) De la recolección de datos a la prueba de la hipótesis

Ahora bien, ¿qué es exactamente la Estadística?, ¿en qué consiste?

Justamente, es la disciplina que, sin apoyarse en la totalidad de los datos, “sabe, mediante un análisis sutil sacar partido de una pequeña cantidad de datos elegidos con acierto”. (1).

El método estadístico nos ayuda a simplificar los datos agrupándolos primero por medio de clasificaciones o agrupaciones de acuerdo a ciertas características comunes, y luego nos proporciona los datos simplificados en una tabla de frecuencia; por ejemplo: 40 niños de un grado de escuela primaria se clasifican según las edades:

<i>Edades</i>	<i>Nº de niños</i>
7 años	2
8 años	18
9 años	15
11 años	5
Total:	40

En sus orígenes la Estadística fue simple recuento, y sus formas más primitivas fueron los censos. A través de creciente perfeccionamiento llegó a ser un instrumento capaz no sólo de ayudar a describir numéricamente la realidad sino también de probar las hipótesis, de hacer previsiones, de estimar “parámetros” en el universo a partir de los datos obtenidos en la muestra, etc.

## III) Objeto de la Estadística

El objeto de la Estadística es la “población total” y no el individuo. Este sólo interesa en tanto forma parte de la “población total” o “universo”.

Cuando decimos “población total” o “universo” no que-

---

(1) VESSEREAU, Andrés: *La Estadística*, Cap. I (Eudeba, 1962).

remos significar la totalidad de los seres humanos, sino la totalidad de individuos que tienen determinada característica que los convierte en objeto de nuestro estudio.

Por ejemplo, si estudiamos los efectos de la estratificación en relación con la socialización, y nos interesamos por la clase baja, nuestra "población total" estará compuesta por los hogares de clase baja.

Antes de comenzar una recolección de datos es imprescindible definir perfectamente los límites de ese "universo" que será el objeto de nuestro estudio. Muchas investigaciones han fracasado al obtener aparentes contradicciones en los resultados por haber trabajado con un "universo" mal delimitado.

No siempre es fácil fijar los límites de la "población total" de nuestro estudio. Es simple cuando nos interesan, por ejemplo, los niños en edad escolar de la ciudad de Rosario; pero es más complejo cuando estamos interesados, por ejemplo, en los "obreros especializados" ya que aquí no existe una característica concreta fácilmente determinable sino que es convencional.

#### IV) *Estadística paramétrica y Estadística no-paramétrica*

En la Estadística moderna podemos hacer una distinción entre dos tipos de técnicas:

1) la que a partir de los valores hallados en una muestra ("estadísticas" o "estadísticos") infiere los "parámetros" en el universo, por lo que es llamada paramétrica. Los "parámetros" son los valores reales correspondientes o "esperados" del universo.

Por medio de esta técnica se pueden inferir en el universo el Medio aritmético ( $\bar{X}$ ), la Mediana ( $Md$ ), el Modo ( $Mo$ ), el Medio Geométrico ( $G$ ), la Desviación standard ( $\gamma$ ), etc.

Esta inferencia de los "parámetros" se hace basándose en suposiciones debido a lo cual se obtienen conclusiones que contienen restricciones y deben, por lo tanto, ser enunciadas con-

dicionalmente, por ejemplo: "Si las suposiciones con respecto a la población son válidas podemos concluir que..." (a)

La técnica paramétrica se apoya sobre distribuciones de frecuencia normal, es decir, sobre distribuciones que representadas gráficamente toman la forma de una curva normal.

2) otra técnica más reciente que no se apoya sobre distribuciones normales es la llamada técnica no-paramétrica o técnica de distribución libre (distribution-free). Esta técnica infiere los valores en el universo de tal manera que las conclusiones que se obtienen no contienen restricciones y, por lo tanto, no deben enunciarse condicionalmente; por ejemplo: "Teniendo en cuenta el aspecto de la población podemos concluir que..." (b).

(a y b son ejemplos dados por Siegel)

Como la técnica paramétrica se basa sobre operaciones (suma, división y multiplicación) a partir de los valores hallados en la muestra, si se aplica a características que no son realmente numéricas (ejemplo: variables cualitativas) puede dar lugar a distorsiones. Este cuidado no hay que tenerlo cuando se usa la técnica no-paramétrica que puede trabajar con características que no son realmente numéricas.

#### V) *Elementos que la Estadística proporciona al Sociólogo*

##### 1) *Descripción numérica de la realidad:*

Hemos visto que la Estadística se vuelve una disciplina auxiliar imprescindible para la Sociología en cuanto es capaz de simplificar la tarea al reducir los datos con los cuales hay que trabajar. Pero ésta no es la única contribución que hace la Estadística. Los aportes que hace son muchísimos, desde esta simplificación y agrupamiento de datos hasta la determinación de los parámetros de posición y dispersión, desde la selección de una muestra hasta la interpretación de los datos obtenidos, desde el cálculo de probabilidades hasta la determinación de correlaciones entre variables.

Estas y muchas otras técnicas aporta la Estadística a la Investigación Sociológica.

¿En qué momento la Sociología recurre a la Estadística y cuáles son las preguntas que le hace?

Desde el momento que se tiene una determinada población con cierta característica a estudiar, la Sociología recurre a la Estadística solicitándole la confección de una tabla de frecuencia, y en ésta las medidas de posición y dispersión.

¿Qué queremos significar con esto y en qué medida es ésta una contribución de la Estadística?

Si tenemos, por ejemplo, una población compuesta por 1.000 obreros de una fábrica y nos interesa saber cuál es el salario medio que reciben, ¿cómo procederemos? Si tomamos todos los distintos salarios individualmente podría ocurrir que entre los 1.000 hubiera grandes diferencias que oscilarían, supongamos, entre los \$ 3.000 y los \$ 8.000 pasando por varias cifras intermedias (El caso extremo hipotético sería que cada obrero ganara un salario distinto al de los 999 restantes). Necesitamos, para simplificar el estudio, agruparlos de alguna manera. Determinamos arbitrariamente límites y establecemos una tabla de frecuencia de datos agrupados (En otros casos, cuando la población a estudiar es menor no hace falta agrupar). Ejemplo:

<i>Salarios</i>	$f_i$	$X_i$	$f_i X_i$
3 a 4 mil	500	3.500	$500 \times 3.500$
4 a 5 mil	300	4.500	$300 \times 4.500$
5 a 6 mil	145	5.500	$145 \times 5.500$
6 a 7 mil	35	6.500	$35 \times 6.500$
7 a 8 mil	20	7.500	$20 \times 7.500$

Total: 1.000

Por medio de una fórmula especial para calcular el Medio Aritmético de datos agrupados determinamos el Medio Aritmético de los salarios de los 1.000 obreros.

$$\text{Ejemplo: } \bar{X} = \frac{\sum f_i X_i}{\sum f_i} = 4.285.$$

En esta fórmula:  $\Sigma$  = sumatoria;  $f_i$  = frecuencia de cada intervalo de clase;  $X_i$  = punto medio de cada intervalo de clase.

En este ejemplo podemos darnos cuenta de dos contribuciones de la Estadística:

- 1) agrupó los datos simplificándolos y
- 2) extrajo una medida de tendencia central: el Medio Aritmético.

Trabajar con el Medio Aritmético en lugar de trabajar con los 1.000 salarios simplifica enormemente la tarea como se percibe a simple vista.

Según el tipo de problema a que estemos abocados podemos determinar la medida de posición que más convenga ( $Md$ ,  $Mo$ ,  $G$ ,  $\bar{X}$ ) haciendo un cálculo similar al ejemplificado.

Para otros problemas se torna cómodo trabajar con lo que se llaman “medidas de dispersión” (desviación media, desviación standard, varianza, cuartiles, deciles y percentiles), que son medidas que determinan las desviaciones de las medidas individuales respecto del Medio Aritmético.

Hasta aquí hemos tratado de resumir muy brevemente las técnicas por medio de las cuales la Estadística le proporciona al Sociólogo datos observados y simplificados sobre los que se pueden hacer interpretaciones de manera más cómoda y profunda. Podríamos llamar a ésta la fase descriptiva.

## 2) *Determinación de tendencias y previsiones:*

Una contribución valiosa está constituida por la determinación de las tendencias (trends) —movimientos lentos que se

producen en una serie a largo plazo y en forma suave—, de las variaciones cíclicas y de las variaciones estacionales.

En este punto encontramos una fase descriptiva: dichas determinaciones; y otra fase predictiva.

Las predicciones que se pueden realizar por medio de extrapolaciones son de valioso aporte para la Demografía (tablas de mortalidad, natalidad, etc.) y para la Sociología en general, por ejemplo, al determinar cómo será el mercado para cierto producto de una fábrica en años sucesivos, dato éste de importancia para la Sociología Industrial.

### 3) *El aporte de las leyes de probabilidad:*

En la Investigación Sociológica también tiene importancia el poder determinar el peso de la casualidad o azar en los acontecimientos. Desde el momento que se pudieron establecer regularidades en algo que parecía, justamente, tan irregular, tan azaroso, fue más fácil establecer algún tipo de predicción.

Las leyes matemáticas de la probabilidad contribuyen prediciendo o determinando la probabilidad que tiene un acontecimiento  $X$  de producirse en un tiempo  $Y$ .

### 4) *La dependencia estadística:*

Si tenemos dos acontecimientos relacionados de manera tal que a variaciones en uno corresponden variaciones en el otro, decimos que hay dependencias entre ellos. La dependencia puede ser negativa o positiva, alta o baja.

Un ejemplo: podemos querer determinar qué relación existe entre “infancia feliz” y el “éxito matrimonial”. Por medio de un coeficiente podemos conocer el tipo de relación y el grado de la misma. Tomamos 100 sujetos que de acuerdo a sus respuestas agrupamos en el siguiente cuadro:



	<i>Matr. exitoso</i>	<i>Matr. no exitoso</i>	<i>Totales</i>
Infancia feliz	40	70	110
Infancia infeliz	60	30	90
Totales	100	100	200

Aplicando un coeficiente apropiado determinamos, por ejemplo que hay una correlación positiva de 0,35 o del 35%. Decimos que la correlación hallada es baja.

Hay varios tipos de coeficientes de dependencia: a) los coeficientes de correlación, que se aplican a características mensurables (variables cuantitativas); ejemplos: C. de Pearson, C. Tetracórico, C. Punto biserial; b) los coeficientes por rangos, que se aplican a características clasificables; ejemplos: C. de Spearman y C. de Kendall; c) los coeficientes o cuadros de contingencia, que se aplican a variables cualitativas; ejemplos: C. de Yule.

¿Cuál es el aporte de estos coeficientes a la Investigación Sociológica?

En primer lugar, el uso de coeficientes es de suma importancia en aquellos aspectos de la realidad donde no se puede recurrir a la experimentación. Por ejemplo: es imposible detectar en una situación experimental las vinculaciones que pueden darse entre estas dos variables: infancia feliz o infeliz y el éxito en el matrimonio. No lo podemos hacer experimentalmente porque no podemos someter a un grupo a una "infancia feliz" y al grupo control a una "infancia infeliz". Pero si determinamos la frecuencia con que ellas se dan asociadas en una cierta población, podremos manipular esa variable y sacar conclusiones como si hubiéramos realizado la experimentación correspondiente.

En segundo lugar, la determinación de dependencias estadísticas puede ser valiosa contribución en aquellos casos donde se torna difícil o imposible la determinación de la existencia de una variable. Supongamos que se haya determinado que

existe una alta correlación entre “frustración” y “antisemitismo”, y que en determinado grupo no podamos aplicar una escala de A.S., y que, en cambio, podamos detectar fácilmente y sin despertar resistencias el grado de “frustración”. Al hacerlo, comprobamos, por ejemplo, que el grado de “frustración” es alto, y por ende, podemos suponer que el A.S. será alto también.

Es necesario no confundir dependencia estadística con relación de causa a efecto.

#### 5) *La Muestra:*

La Muestra se extrae del universo o población total y es una pequeña porción que contiene las mismas características que aquél.

Hay varios tipos de Muestra (aleatoria simple, estratificada con fracción de muestreo, estratificada con fracción variable de muestreo, por conglomerado, por etapas múltiples) que coinciden en algo esencial: al extraerlas se lo debe hacer de tal modo que todos los individuos del universo tengan exactamente la misma oportunidad de ser incluidos. Existen varios sistemas de elección de los individuos, siendo probablemente el más estricto el que se vale de una tabla de Números al azar.

Hay otro tipo de Muestra en la cual los individuos no son elegidos al azar (Muestra por cuotas).

El tipo de Muestra y el tipo de selección de los individuos depende en cada caso del estudio que se está realizando.

Una Muestra Aleatoria Simple sirve, por ejemplo, cuando se quiere detectar la opción política de la población total; pero cuando se trata de detectar eso mismo en relación con otro factor, por ejemplo, con la edad, lo más conveniente será, sin duda, realizar una Muestra Estratificada con estratos de acuerdo a las edades.

Sobre este universo reducido de la muestra se trabaja y los datos que se obtienen en él se llaman “estadística” o “es-

tadístico" (Medio aritmético, Modo, Mediana, etc.). A partir de ellos y por medio de las técnicas paramétricas ya enunciadas anteriormente encontramos los valores correspondientes en el universo total; estos valores se llaman "parámetros".

Por ejemplo, si queremos averiguar la cantidad de analfabetos en la ciudad de Rosario (suponiendo que no hubiera un Censo que nos diera las cifras reales), y tomamos una Muestra Aleatoria Simple obtendremos en la Muestra un valor "estadístico"  $X$ , del cual inferiremos el "parámetro"  $Y$  en el universo.

#### 6) *La inferencia estadística:*

La inferencia estadística es uno de los tópicos centrales de la Estadística moderna, y está relacionada con a) la estimación de los parámetros, y b) las pruebas de las hipótesis.

a) la estimación de los parámetros: cuando se ha averiguado el "estadístico" en la Muestra (Medio Aritmético, Desviación Media, etc.) y se infiere el parámetro en el universo, este valor se llama "estimación", porque es una estimación de la característica real que sigue siendo desconocida ya que no se hace el recuento de la población total.

En el ejemplo de los analfabetos en Rosario, ese parámetro  $Y$  no es el valor real ya que no hicimos un censo de la población total sino que es una "estimación" del valor real.

Dadas las cosas de esta manera, es evidente que las técnicas estadísticas no pueden suministrar nos nunca certezas puesto que no conocemos la totalidad; pero lo que podemos esperar es una mayor o menor aproximación al valor real de acuerdo al número de observaciones. Cuanto mayor es el número de observaciones mayor será la probabilidad de acercarse al valor real. El margen de error disminuye en la medida que aumenta el tamaño de la muestra.

Podemos decir que el grado de certeza es probabilístico. Se fija de antemano el grado de certeza que se desea obtener. Si la estimación se hace con un solo valor: se llama "estima-

ción puntual". Si se hace con un intervalo entre dos valores: se dice que es una estimación hecha con un "intervalo de confianza". El porcentaje de veces que este intervalo contiene al parámetro es el "nivel de confianza".

b) la prueba de las hipótesis: si en una investigación encontramos que una serie de datos confirman la Hipótesis de trabajo, podemos interesarnos en cerciorarnos de que tales datos responden a la realidad o si se han debido al azar o a errores en la aplicación de la técnica.

Hay un procedimiento que nos proporciona la misma Estadística para comprobar la validez de los datos obtenidos. Este procedimiento consta de varias etapas:

1) Se establece una  $H_0$  (hipótesis nula). Se llama así porque existe la probabilidad de que sea rechazada y se acepte otra (hipótesis alternativa). Esta hipótesis nula es la opuesta a la hipótesis comprobada o es la que enuncia que los valores hallados se deben al azar.

2) se elige un test estadístico para probar esa  $H_0$ :

—si lo que se desea averiguar es la probabilidad que tiene el valor estadístico hallado de provenir del universo, se elige un "test de significación" del valor estadístico correspondiente ( $Md$ ,  $Mo$ , etc.).

— si lo que se desea averiguar es si la diferencia hallada entre dos valores similares de dos muestras del mismo universo (dos medios aritméticos, dos varianzas, etc.) es significativa o no, es decir, si hay un error o si sólo se debe al azar se elige un test de significación de las diferencias entre dos medios aritméticos, o dos varianzas, etc.

3) Se determina el tamaño de la Muestra y se elige un nivel de significación, por ejemplo, un nivel de significación del 5%. El nivel de significación del 5% tomado como ejemplo, implica que hay una probabilidad del 95% de que el valor "estadístico" o la diferencia entre valores "estadísticos" no se deba al azar.

4) Trabajando con la Tabla de Areas y Ordenadas de la curva normal se determina la zona de rechazo para ese nivel de significación.

5) Si el valor que se obtiene cae fuera de la zona de rechazo no se podrá rechazar la  $H_0$  con ese nivel de significación que se había elegido. ¿Qué significa esto? ¿Qué aceptamos la hipótesis contraria a la original? No, simplemente podemos decir que la  $H_0$  no está contradicha por los hechos.

#### B) LIMITACIONES Y ERRORES EN EL USO DE LA ESTADÍSTICA

Es necesario —para no caer en graves errores— conocer las limitaciones y errores que pueden aparecer al manipular las técnicas estadísticas. Apreciar la utilidad de la Estadística no significa no darnos cuenta de las limitaciones que pueden surgir.

Estas limitaciones pueden ser de dos tipos:

- 1) limitaciones y errores por falta de adecuación de la Técnica Estadística al objeto de estudio; y
- 2) limitaciones y errores por falta de un conocimiento adecuado de esa Técnica.

#### I) *Limitaciones y errores por falta de adecuación de la Técnica Estadística al objeto de estudio:*

a) Los datos sociológicos no se pueden medir como los de otras ciencias. Pasteur pudo aislar materia esterilizada y demostrar que no existe la generación espontánea de los virus. En cambio, en Sociología no se puede hacer otro tanto; no se puede tampoco aislar el fenómeno social como se puede aislar un cuerpo en el vacío para estudiar las leyes de la caída de los cuerpos. En un fenómeno social aun cuando se estudie en un experimento de laboratorio no se pueden eliminar todos los factores espúreos o intervinientes.

b) Los métodos estadísticos sólo se pueden aplicar a da-

tos que pueden expresarse en alguna clase de medida. Hay ciertos puntos de interés de la Sociología que no son cuantificables. En esos casos la Estadística se torna inaplicable (En algunos casos queda el recurso de hacer cuantificable el fenómeno que nos interesa; pero este recurso no siempre es factible).

c) El análisis estadístico no nos puede resolver los problemas en su esencia. La Estadística colabora en la investigación sociológica, pero es el sociólogo el que debe buscar en el marco de la Sociología la explicación de los hechos observados.

d) En algunos casos los datos encontrados pueden ser significativos estadísticamente, pero igualmente pueden carecer de importancia para la investigación sociológica. "Cuando una predicción de una actitud, por ejemplo, muestra que dos grupos difieren por sólo uno o dos puntos de porcentaje, es fácil que para algún propósito de predicción la diferencia sea insignificante, no importa cuán significativa sea estadísticamente la respuesta"... (2).

## II) *Limitaciones y errores por falta de conocimiento adecuado de la Técnica Estadística:*

Como hemos visto, la Estadística en la Investigación Sociológica es una herramienta de trabajo imprescindible. Pero si no se la sabe usar puede tornarse un arma de doble filo. Si no sabemos manipular la Técnica Estadística corremos el riesgo de elegir mal y usar un instrumento que no se adecúe al estudio.

Según sea el problema que interese hay una muestra más conveniente, un test de prueba más adecuado, un coeficiente de correlación más exacto.

Si estamos interesados, por ejemplo, en una encuesta para averiguar si hay correlación entre la opción política y otros factores: edad, sexo y ocupación, y elegimos una Muestra

---

(2) COHEN y NAGEL: *An Introduction to Logic and Scientific Method.*

Aleatoria Simple, no será ésta la más conveniente y sí lo será seguramente una Muestra Estratificada de acuerdo a esos factores. Luego dentro de la Muestra, el hecho de determinar las viviendas a entrevistar por algún sistema de elección al azar, no basta para garantizar el que todos los individuos electores tengan la misma probabilidad de ser elegidos. Es muy importante fijar un sistema correcto de elección de la persona a entrevistar en cada vivienda. Si decidimos entrevistar al azar al primero que abre la puerta, lo más probable es que sean en su mayoría amas de casa, y descartamos a todas las personas que por su ocupación están fuera de su hogar.

Esto y muchos otros detalles menudos pueden anular la eficacia de una técnica en sí misma muy efectiva.

Otro problema frecuente aparece cuando no se conocen de antemano todos los factores intervinientes que podrían aparecer. La situación óptima de conocimiento de todos los posibles factores es sólo ideal. La Estadística puede ayudarnos una vez más a detectar que tal correlación positiva que hemos obtenido entre dos factores, por ejemplo, no es tal sino debida a un tercer factor interviniente. Si no conocemos la Técnica Estadística con rigor nos podemos quedar conformes con esa correlación positiva primera sin ponerla en duda.

Es posible que se den también pequeños errores que invalidan los resultados, al confeccionar la distribución de frecuencias. Por ejemplo, cuando se hace la tabulación de alguna variable cuantitativa si el intervalo de clase no se elige cuidadosa y razonablemente estamos viciando los datos.

La fase más difícil y en la cual se varan muchas investigaciones es la interpretación de los datos. La Estadística nos auxilia hasta ese momento ordenándonos los datos y simplificándolos. Pero esta fase descriptiva no es la única en la cual interviene —como ya lo hemos visto—. Falta la Interpretación, el paso de los números a conceptos. Y este paso no puede hacerse exitosamente si no se conocen las técnicas apropiadas.

Son bien conocidos los casos, y con seguridad todos hemos visto de cerca alguno — en que se inician investigaciones muy promisorias y ambiciosas pero que no llegan a etapas posteriores a la recolección de datos.

Cuando se trabaja con correlaciones se corre el riesgo de caer en un error muy frecuente entre los que son poco conocedores de la Estadística cuando confunden correlación con relación de causa a efecto.

La Estadística trabaja con medidas, coeficientes, tests, etc., cada uno de los cuales es adecuado para ciertos propósitos pero inadecuado para otros. En consecuencia, es de suma importancia saber elegir el instrumento estadístico que responde más a las características del trabajo que se está haciendo o al tipo de datos que se manipula. En un caso, por ejemplo en el que la distribución se aleja mucho de una distribución normal es incorrecto usar el Medio Aritmético; lo que corresponde es usar la Mediana, ya que ésta no es afectada por los valores extremos como lo es el Medio Aritmético.

Un error considerable se puede cometer al manipular coeficientes de correlación cuando hay un conocimiento poco profundo de la técnica estadística y de la Metodología en general. Si se consideran dos variables hay que verificar previamente si las mediciones hechas de las dos variables son independientes entre sí. Puede ocurrir que se esté trabajando con lo mismo pero con rótulos distintos y lógicamente obtendríamos una alta correlación.

Hyman llamó a esto “contaminación” de una variable con la otra. Analizando el trabajo sobre Personalidad Autoritaria halló que en una correlación encontrada por Adorno entre “orientación político-económica” y “rasgos de personalidad” las variables estaban contaminadas entre sí.

Errores más gruesos pueden cometerse por desconocer que determinado coeficiente se aplica a variables cualitativas y no a variables cuantitativas, o viceversa. Por ejemplo, el coeficiente tetracórico ( $r_t$ ) que se aplica a variables cuanti-



tativas normalmente distribuidas y ambas dicotomizadas pueden por error aplicarse a variables cualitativas.

No se nos escapa que muchos o la totalidad de estos errores no invalidan la eficacia de la Estadística en sí misma. Lo que se quiere es reiterar que una técnica tan eficaz e imprescindible para la investigación puede transformarse en una fuente de errores y tergiversaciones si se tiene un conocimiento insuficiente de la misma.

### III) *La Estadística no debe ser un fin sino un medio.*

El reconocer la utilidad de la Estadística como disciplina auxiliar no implica de ningún modo convertirla en el fin de la investigación.

Relacionado con lo que decíamos anteriormente, por buena que haya sido la planificación de la recolección de los datos y por exitosa que haya sido su realización, si no se va más allá no se dará respuesta a la hipótesis. Quedarse en esa etapa implicaría hacer de la recolección de datos un fin en sí mismo, cuando en realidad los datos deberán ser tan sólo un medio para responder a las hipótesis originales.

Son muchas las investigaciones que han fracasado por perder de vista el planteo hipotético que se había formulado previamente, olvidándolo y haciendo de la recolección de datos el fin último.

### IV) *Limitaciones y errores por falta de un planteo hipotético adecuado.*

Problema similar surge cuando no se enuncia previamente en forma adecuada el planteo hipotético.

Nos dicen Goode y Hatt que "las tabulaciones de cuestiones pobremente formuladas pueden tener poca significación".

... "El estudiante puede decidir llevar a cabo varias operaciones estadísticas en este caso, pero él debe entender que

estas no cambian el diseño subyacente del estudio .Ningún refinamiento remediará tales debilidades" (3).

Tener un planteo hipotético previo bien definido significa que todos los factores relevantes son conocidos y, por ende, que no surgirán problemas inusitados en el análisis.

En caso contrario puede suceder algo similar a lo que acontece cuando el cuerpo hipotético es deficiente. Puede, una vez más, una investigación quedar varada a la altura de la recolección de datos, datos que quedarán a la espera de una reformulación del planteo previo e interpretación adecuada. O bien puede suceder que la interpretación de los datos estadísticos conduzca a cualquier conclusión viciada de nulidad.

Un ejemplo bien conocido de esta tendencia a hacer de la recolección de datos un fin en sí mismo, olvidando que la importancia y eficacia de la Estadística radica, precisamente, en ser un medio auxiliar para probar las hipótesis, es la investigación "The American Soldiers" de Stouffer.

Fue éste un caso típico de lo enunciado en los puntos III y IV. Se inició una investigación sin un planteo previo bien delimitado y se terminó haciendo de la Técnica Estadística de recolección de los datos un fin en sí misma. La recolección misma fue muy exitosa y posteriores trabajos aprovecharon la gran cantidad de datos obtenidos, pero el análisis ex-post facto no es el ideal de investigación.

### C) CRITICAS AL USO DE LA ESTADISTICA

#### I) 2 posiciones extremas:

No todas las corrientes de la Sociología actual asignan a la Estadística la misma importancia en la Investigación Sociológica.

Por un lado están las corrientes que no sólo le reconocen

---

(3) GOODE y HATT: *Methods in Social Research* (McGraw-Hill Book, 1952).

importancia sino que hacen de la fase estadística la más elaborada de la investigación, descuidando un tanto o completamente, la fase posterior de interpretación de los datos. En casos extremos de esta corriente, se parte a la búsqueda de datos sin un cuerpo hipotético previo. Se va en busca de los datos por los datos mismos. Se determina una población que interesa por sus características y se comienza a aplicar una serie de pruebas. Indudablemente, los datos que se recogen no son desdeñables, pero careciendo de un planteo previo que dan a la espera de interpretaciones ex-post facto que no siempre llegan. Se pierde de esta manera esfuerzo, tiempo y dinero.

Mucha de la sociología norteamericana contemporánea puede ubicarse dentro de esta tendencia. Existen innumerables ejemplos de tales tipos de trabajos, siendo quizá uno de los más elocuentes el ya citado "The American Soldiers" realizado por S. Stouffer, investigación que se limitó a detectar las opiniones, actitudes y comportamientos de dos grupos, uno que estaba en un frente de batalla y el otro no.

El material fue reestructurado por Merton en un análisis secundario. Ha servido para innumerables trabajos de análisis ex-post facto.

Esta tendencia no es compartida por todos los sociólogos contemporáneos. Entre los que disienten podemos distinguir dos matices: los que disienten simplemente y que realizan una crítica al estilo de C. W. Mills, y los que son detractores y hacen una crítica destructiva como la de P. Sorokin.

## II) *Una crítica: C. W. Mills:*

C. W. Mills en "La imaginación sociológica" hace una crítica a la sociología norteamericana actual y refiriéndose al desarrollo de las técnicas metodológicas en general, aclara que su concepción se opone a lo que él llamó Empirismo abstracto, al que describió "como conjunto de técnicas burocráticas que impiden la investigación social con sus pretensiones metodoló-

gicas que congestionan el trabajo con conceptos oscurantistas o que lo trivializan interesándose en pequeños problemas públicamente importantes”.

Y más adelante agrega:

“El estudio de los hechos contemporáneos fácilmente puede convertirse en una serie de datos de ambiente sin relación entre sí y con frecuencia insignificantes”...

*Una “caricatura” de crítica: P. Sorokin*

Sorokin en su libro “Achaques y manías de la Sociología Moderna y Ciencias afines” hace una crítica a lo que él llama “Quantofrenia”. Insiste en la ficción de la traslación de los métodos de la matemática hacia las Ciencias Sociales puesto que se pretende aplicar fórmulas matemáticas a fenómenos que no son susceptibles de cuantificación.

Considera que se hace con esto una “deformación” y “contrabando” de las fórmulas matemáticas.

Toda su crítica está hecha en un lenguaje cáustico, completamente destructivo, casi agravante al referirse continuamente a la investigación empírica con abundancia de términos tales como “obsesión quantofrénica”, “ficción”, “deformación”, “metrofrenia”. No deja nada en pie ni aprovechable; nada de lo que hacen los investigadores tiene algún valor.

III) *Una “tercera posición”: Merton.*

Merton en su trabajo “Teoría Sociológica e investigación empírica” hace un balance entre las dos tendencias extremas; entre aquella que “huye de la “trivialidad” de la observación detallada y en pequeña escala y busca la grandeza de las síntesis globales” y la que “no se preocupa demasiado por descubrir las inferencias de sus investigaciones, pero permanece confiada y segura de que la información que suministra es verdadera”.

Postula que “no hay base lógica para el hecho de que se enfrenten con hostilidad”.

## D) PERSPECTIVAS Y CONCLUSIONES

### I) *Necesidad de encontrar el "justo medio"*:

Entre la posición que considera a la Estadística y a la ubicación e importancia de la Estadística es necesario encontrar la posición exacta.

Entre la posición que considera a la Estadística y a su técnica de Investigación una cosa superflua y que oscurece la realidad social que se pretende conocer y la opuesta que hace de la Estadística un fin en sí mismo, olvidando que sólo es un medio para conocer la realidad, debemos tratar de hallar un "justo medio", es decir, debemos darnos cuenta perfectamente de qué y cuánto es lo que podemos esperar de esta técnica auxiliar de la investigación sociológica.

Si la descartamos haremos una sociología que no será tal, sino posiblemente "teoría" sin base empírica, más cercana a la filosofía social que al conocimiento científico de la realidad social; si le asignamos más importancia de la que debe tener y la convertimos en fin dejaremos la investigación a medio camino y sin posibilidades de generalización y de inserción en un cuerpo más amplio de teoría.

### II) *Necesidad de que el Sociólogo conozca Estadística*:

Como consecuencia de lo visto concluimos que es necesario y casi diríamos imprescindible que el Sociólogo conozca las Técnicas Estadísticas. Muchos investigadores, especialmente los que recién se inician, desdennan a veces las técnicas y consideran un poco denigrante el que se les exija un conocimiento de las mismas.

Piensen en estos casos que la presencia y asesoramiento de un Estadístico es suficiente y correcto para llevar a cabo de manera exitosa la parte técnica de la recolección de datos. Pero olvidan que el aporte o la contribución del perito en Estadística no puede ir más allá, y que debe ser el Sociólogo

—tal como lo son en sus especialidades el biólogo, el economista, el psicólogo social, etc.— el que realice la interpretación final de los datos obtenidos.

No se puede y se debe evitar por sumamente peligroso que sea el técnico en Estadística el que realice la interpretación final de los datos recogidos.

Con todas estas consideraciones creemos haber destacado la función que cumple la Estadística en el conocimiento riguroso de la realidad social, alejando el peligro de la teorización vacía con escasa o ninguna base empírica; y al mismo tiempo hemos querido ubicar el tema en el contexto más general de la metodología sociológica y su relación con la teoría, señalando las limitaciones de un conocimiento estadístico librado a sus propias fuerzas, que al desarrollarse sin una guía teórico-metodológica más amplia puede llevar la Sociología a extremos que a toda costa se deben evitar.

**HILDA HABICHAYN**

Entre Ríos 750, Rosario