



Revista Electrónica de Metodología Aplicada  
1998, Vol. 3 nº 1, pp. 16-29

URL:<http://www3.uniovi.es/~Psi/REMA/v3n1/a2v3n1.wp5>

---

**LA CALIDAD DEL MUESTREO EN LAS  
INVESTIGACIONES SOCIALES.**

**Vicente Manzano Arrondo**  
**Dpto. de Psicología Experimental**  
**Universidad de Sevilla**  
**e-mail: [vmanzano@cica.es](mailto:vmanzano@cica.es)**

**ABSTRACT.**

Even though the importance of the sampling within scientific procedure for the knowledge generation has is recognized, professionals of the social and human sciences, customarily neglect the extraction phase to a sample. In the present work we concerns very especially to accomplish an attention call on the transcendancy of this problems; at the same time, we will indicate possible causes and some of the consequences derived. Moreover, it presents help system software SOTAM, with the objective to improve the samples quality.

**Key word:** Sampling.

**1.- Acerca de la estadística**

Cada vez es mayor la proporción de investigadores, en las más diversas disciplinas científicas, que realizan análisis estadísticos de datos como procedimiento formal para llegar a conclusiones o apoyar procesos de decisión sobre las hipótesis de la investigación (Vallecillos, 1996). Para muchos de ellos, utilizar las herramientas de la Estadística en la investigación, dentro de su área de conocimiento, es un quehacer poco motivador, puesto que su dominio puede alejarse bastante de los objetivos de formación, investigación y aplicación del experto en ese área. La duda que se genera tanto en el investigador como en el docente es: «Cuando se estudia estadística, no para convertirse en especialista en la materia sino para trabajar en ciencias humanas, en psicología en concreto, ¿hasta dónde habría que profundizar en aquella materia?» (Carpintero, 1989:476).

Este dilema entre motivación y necesidad, en la utilización de los recursos estadísticos tiene consecuencias importantes. «La necesidad de acudir a las técnicas estadísticas matemáticas para la resolución científica de la mayor parte de los problemas plantea una dificultad a veces insuperable. ¿Cómo puede un investigador usar dichas técnicas si carece de conocimientos sólidos de la ciencia matemática?» (Alberdi y otros, 1969:21).

A esta problemática hay que añadir la resistencia de algunos investigadores de las ciencias humanas a considerar métodos *excesivamente técnicos* que «pierden simultáneamente su interés humano» (Boulanguer, 1971:14). El mismo autor señala más adelante (pág. 93): "Hay que pasar por una etapa fastidiosa de análisis (...) Es preciso reconocer que el

investigador de las ciencias humanas cuenta todavía con demasiados pocos medios para abordar esta etapa que enfría su entusiasmo y le obliga a perderse entre los recovecos, idas y venidas, de unas técnicas con frecuencia mal comprendidas».

Los recursos de la estadística para la investigación científica pueden considerarse como un conjunto de herramientas conceptuales que poseen sus propias condiciones de aplicación y pautas de uso. Respetando ambos aspectos, las conclusiones que se realizan sobre las informaciones tratadas convenientemente con herramientas estadísticas, tendrán una base sólida. En caso contrario, las conclusiones pueden ser erróneas e, incluso, engañosas.

La investigación en Psicología, Sociología y Educación, al igual que ocurre en otras ciencias, en buena medida se basa en el manejo de recursos estadísticos como elementos indispensables para llegar a conclusiones aceptables por el resto de la comunidad científica. Dada la peculiaridad de su objeto de estudio, inabordable en la mayoría de los casos si no es a través de perspectivas complejas de relación entre variables, la atención de los investigadores en las ciencias humanas y sociales se concentra cada vez más en la llamada Estadística Multivariante. Los diseños complejos de investigación y análisis, las aportaciones más recientes de la informática para la aplicación de técnicas avanzadas de manipulación de datos y la discusión de estos aspectos desde perspectivas teóricas y aplicadas, preocupan y concentran a multitud de profesionales cuyo quehacer cotidiano es el estudio de cómo se investiga, haciendo de ello su especialidad. Paralelamente, otras especialidades dentro de estas ciencias utilizan el conocimiento ya elaborado y retransmitido, preocupadas más por los resultados y posibilidades que por las condiciones de aplicación y el fundamentos de uso, de tal forma que se ha propiciado la utilización de las técnicas estadísticas, sin considerar la adecuación de éstas a las condiciones en las que se aplican (Solanas y Sierra, 1992). A su vez, las ciencias sociales se han visto *apabulladas* en los últimos años por avances vertiginosos en informática y aplicaciones estadísticas (Manheim, 1982; Rossi y otros, 1983), y muy especialmente en la psicología (Judd y otros, 1995), lo que favorece una absorción de poca calidad por parte de los especialistas en áreas no metodológicas. Por otro lado, la adopción de procedimientos informáticos para realizar tareas metodológicas no parece ser una solución inmediata, considerando la ansiedad que generan los ordenadores, fenómeno muy generalizado (Fariña y Arce, 1993).

La fusión de esta creciente complicación de las herramientas de análisis, junto con la discrepancia entre los objetivos de formación y la necesidad de uso de los recursos estadísticos, consigue finalmente que el especialista en áreas aplicadas tienda a descuidar aspectos muy básicos, previos a la aplicación de estos recursos estadísticos complejos. Por otro lado, en muchas ocasiones, la aplicación de herramientas estadísticas se deja arrastrar por *hipótesis de comodidad*, en el sentido de aplicarse para permitir la ejecución de una prueba o el ajuste de un modelo, no porque son las estrategias más adecuadas, sino porque son las más cómodas (Lebart y otros, 1985).

## 2.- Sobre el muestreo

Muchos son los apartados dentro de la estadística y áreas relacionadas, expuestos a esta problemática. Pero, sin duda, uno de los capítulos de mayor trascendencia a la hora de utilizar los conocimientos generados gracias a la mediación de la estadística y, por tanto, más sensibles a una utilización impropia, es la extracción de muestras. Como indica Azorín (1969:5) «Puede considerarse como problema crucial de la Estadística el de "inferir la población" o afirmar algo sobre ella a partir de una muestra». Cabe señalar que en las ciencias sociales y muy especialmente en Psicología y Educación, además de en las más diversas situaciones cotidianas (Sukhatme, 1953; Stuart, 1984), las acciones e investigaciones versan, la mayoría de las veces, sobre muestras y no sobre poblaciones (Jolliffe, 1986; Rodrigues, 1977), tendencia que continúa en claro aumento (Silva, 1993). Por otro lado, la extracción de muestras es fundamental para las investigaciones por encuestas y ocurre que éstas dominan el panorama de la investigación social (Yates, 1981; Rossi y otros, 1983; Presser y Traugott, 1992), de tal forma que su campo de aplicación no sólo ya es enorme, sino que continúa en expansión (Raj, 1972). Es más «debe verse ya claramente que el muestreo es parte esencial de todo procedimiento científico» (Goode y Hatt, 1977:259).

A pesar de la importancia del muestreo en la investigación, existen serias deficiencias en la utilización de las muestras por parte de los investigadores. «Mi impresión es que en muchos casos estos investigadores están enterados de que están disponibles poderosos procedimientos de muestreo, pero se creen incapaces para utilizarlos porque son demasiado difíciles o costosos. En cambio, son utilizados procedimientos *ad hoc* increíblemente descuidados, frecuentemente con resultados desastrosos» (Sudman, 1976:IX).

Existen dos supuestos generalizados que mueven a investigadores de diversas ciencias a *descuidar* la fase del muestreo en sus estudios:

1. Supuesto de uniformidad. Dado que las características poblacionales son más o menos constantes en toda su extensión, cualquier porción es representativa del total. Esta es la justificación para las muestras de sangre en los análisis de laboratorio: la composición del líquido es más o menos constante en todo el cuerpo. (Cochran, 1976).

2. Supuesto de disposición aleatoria. La población no es uniforme en cuanto a los valores de la/s variable/s que interesa/n en el estudio, pero éstos se reparten aleatoriamente en toda la extensión poblacional. Por esta razón, cualquier porción de la población será una muestra aleatoria tan válida como si se hubiera cuidado muy especialmente el componente probabilístico en la selección.

Parece, no obstante, que estos argumentos dan buenos resultados en ciencias como astronomía, física y química, si bien es poco aconsejable asumirlos en las ciencias sociales, medicina y biología (Kish, 1965). Es más, en general mantener el supuesto de homogeneidad o uniformidad, es un proceder lamentable en muestras que incluyen personas (Lininger y Warwick, 1984).

### 3.- Sobre el tamaño de la muestra

Pero quizá, entre los aspectos del muestreo que antes y más preocupa a los investigadores figura en un lugar principal la pregunta *¿Cuál es el tamaño que debe tener la muestra?* (Feigl, 1978; Teijeiro, 1990; Sudman, 1983; Kalton, 1987; Fink y Kosecoff, 1989; Henry, 1990; Czaja y Blair, 1996). Para Teijeiro (1990), es la primera decisión a tomar cuando se proyecta el muestreo. Para responder a esta pregunta, debe considerarse una multiplicidad de factores, además del hecho que la estimación del tamaño estará en función de algunas suposiciones (Davis, 1971) y teniendo en cuenta que «es evidente que la calidad y validez de los resultados de una encuesta dependen del tamaño de la muestra en cuestión» (Ghiglione y Matalon, 1989:54).

Lo cierto es que la decisión sobre el tamaño muestral que debe considerarse suele zanjarse utilizando criterios no estadísticos como la inercia o la aplicación monótona de un único contexto teórico: el muestreo aleatorio simple y con niveles típicos para las variables independientes relevantes ( $\alpha=0'05$ , sin considerar  $\beta$  ni el tamaño del efecto, distribución normal y varianza desconocida, con estimación de proporciones).

Este comportamiento monótono y que consiste en perpetuar los valores que se acostumbran a considerar, queda justificado principalmente por cuatro argumentos:

1. Comodidad: se ahorra al investigador el realizar una búsqueda exhaustiva de la literatura específica para encontrar la expresión de cálculo del tamaño más adecuada. O, peor aún, debe deducirla a partir de la estimación de las varianzas generadas por el modelo de muestreo que se pretende aplicar. En cualquier caso, aún cuando consiga la expresión, debe *enfrentarse* a ella. Fuera del muestreo aleatorio simple, lo esperable es que las expresiones sean complejas, con lo que puede llegar a aconsejarse al inexperto que acuda a niveles más sencillos para estimar el tamaño de la muestra (Salgado, 1990) e, incluso, a insistir en la sencillez de la solución (Veres, 1981).

2. Conformidad: el no utilizar la expresión más idónea puede ser considerado como un *pequeño pecado* que se asume colectivamente. Más aún, cuando algunos autores en muestreo consideran que los errores generados por la utilización de una expresión incorrecta del tamaño muestral, quedan sensiblemente empujados frente al resto de errores generados en la investigación, especialmente durante el trabajo de campo (Kish, 1965).

3. Falta de insistencia en la literatura: desde el mismo cuerpo teórico no existe un respaldo conveniente para solucionar el problema de tomar un tamaño de muestra concreto. «De hecho, a nuestro juicio, algunas zonas de la estadística en general y del muestreo en particular están afectados por una especie de "discurso estadístico oficial", un tipo de censura que silencia algunos problemas reales y los suple con formulaciones académicas de escasa o nula aplicabilidad. Tal circunstancia se presenta con especial fuerza en el dominio de los tamaños muestrales» (Silva, 1993:139). El panorama se complica aún más considerando el escaso eco que esta problemática tiene en la literatura (Chevry, 1962).

4. Hábito: circulan algunos tamaños concretos por la literatura al uso y compartidos por investigadores, de tal forma que se acude a la *voz de la experiencia*, en el sentido de que parece ser que ésta fundamenta suficientemente una decisión sobre el tamaño de la muestra que no requiere discusión o replanteamiento. Así, encontramos afirmaciones según las cuales la mayoría de los estudios utilizan muestras de unas 600 unidades (Pérez Cebrián, 1987:37), o entre 1000 y 5000 (Hedges, 1980: 61), o  $\geq 100$  en el medio industrial y  $\geq 1000$  en el medio de consumo (Harvatopoulos y otros, 1992:32), o  $\geq 2000$  (Noelle, 1970:173), o al menos, se asume que son de tamaños invariablemente grandes (Hyman, 1970:45).

Como resultado esperable, frecuentemente la muestra es demasiado pequeña (Kachelmeier y Messier, 1990), otras veces demasiado grande y cualquiera de ambos extremos son usualmente fatales (Stuart, 1984), lo que aconseja dedicar un esfuerzo especial al inicio del diseño de la investigación, para obtener tamaños de muestra óptimos. «La decisión», sobre el tamaño de la muestra, «es importante. Una muestra demasiado grande implica un desperdicio de recursos y una muestra demasiado pequeña disminuye la utilidad de los resultados» (Cochran, 1976:105). Aún así, el consejo *escoja un tamaño de muestra más bien grande* (Jolliffe, 1986) no es nada inusual, quizá provocado por la pregunta típica *¿Es la muestra lo bastante grande?*, cuando la pregunta pertinente sería *¿Tiene la muestra el tamaño adecuado?*. De hecho, un tamaño muy grande de muestra no es ninguna garantía para la validez de las conclusiones (Stoetzel y Girard, 1973). Incluso frecuentemente, se oye la expresión *¿Qué porcentaje de la población debo seleccionar?* por parte de investigadores poco sofisticados (Brandburn y Sudman, 1988) que ignoran la relación *variable* que existe entre los tamaños de muestra y población. Así, «está firmemente arraigada la idea de que se debe consultar a un porcentaje determinado de la totalidad para llegar a resultados suficientemente fidedignos» (Noelle, 1970:173). En definitiva, no es inusual la falta de comprensión acerca de los factores que justifican un tamaño de muestra concreto.

El conocimiento insuficiente sobre las consecuencias del tamaño de muestra en el proceso y resultados de la inferencia estadística, no es monopolio de alumnos y legos en la materia. Se encuentran importantes lagunas y creencias incorrectas incluso en profesionales de amplia experiencia (Vallecillos, 1996). La misma autora señala que los conceptos más problemáticos asociados con el muestreo son la *representatividad* de las muestras y la *variabilidad* de los resultados muestrales (distribución muestral). Al respecto, existen las falacias de que un tamaño elevado de muestra por sí mismo hace a ésta más representativa, (Stoetzel y Girard, 1973) y se encuentra muy extendido el desconocimiento de la relación inversa entre el tamaño de la muestra y la variabilidad (Well y otros, 1990): «La mayoría de la gente no parece darse cuenta de que un resultado extremo de una muestra estadística es más probable en una muestra pequeña» (Timothy, 1989:52).

#### 4.- Acerca de las soluciones

La búsqueda de los diseños de muestreo óptimos o los tamaños de muestra más adecuados, sólo es factible desde la perspectiva de un investigador que goza de las suficientes dosis de conocimiento sobre muestreo. No obstante, la mayoría de los investigadores que recurren a la extracción de muestras, sólo cuentan con un conocimiento útil mínimo, a partir

del que realizan las selecciones. Como resulta obvio, es imposible dominar cada una de las facetas que intervienen en un proceso completo de investigación. El conocimiento más que suficiente en todos los aspectos que se requieren en un estudio, sólo está al alcance de los grupos de trabajo amplios, constituídos por profesionales con una formación específica.

Así, parece que sólo existen dos alternativas para la realización de investigaciones:

1. Un investigador individual sólo puede aspirar a resolver problemas mediante procedimientos que no requieran conocimientos muy amplios.
2. Para utilizar procedimientos que exigen mayor especialización, o bien recurre a otros profesionales, o bien utiliza estrategias que le posibilitan llegar a los fines propuestos, con un claro menoscabo en la credibilidad de los resultados finales.

Existe, aún así, una tercera alternativa: utilizar sistemas de ayuda informatizados. Al respecto, se encuentra bastante generalizada la utilización de programas de análisis de datos que no sólo ahorran esfuerzo, sino que posibilitan la realización de determinadas tareas que de otro modo no serían abordables con ciertas dosis de seguridad en los resultados. No obstante, no sólo en los programas de análisis estadístico se encuentra el apoyo a las investigaciones mediante muestras, en general, y mediante encuestas, en particular.

Utilidades como DEFINE (Moreno, Martínez y Chacón, 1997) ayudan al usuario a clarificar sus objetivos; otras, como N (Glaxo, 1994) permiten realizar un cálculo para el tamaño de la muestra, si bien en contextos muy limitados; los hay que, como DYANE (Santesmases, 1997), ayudan al usuario a culminar diferentes etapas desde la construcción del cuestionario hasta los análisis de datos; en el apartado de software comercial, los profesionales pueden encontrar multitud de productos, generalmente orientados a instituciones especializadas y grupos con amplio presupuesto (ver, por ejemplo, <http://www.surveytracker.com/>).

Una de las utilidades, sujetas a la libre distribución, merece nuestra atención en el presente trabajo: SOTAM (Manzano, 1997). Se trata de un sistema de ayuda a la decisión para las investigaciones mediante encuestas, orientado específicamente hacia la decisión del tamaño de la muestra, pero que acompaña múltiples utilidades satélite.

## **5.- SOTAM**

SOTAM es un sistema informático compuesto por un paquete de programas de ordenador que se gestionan desde un módulo principal. Su objetivo es ayudar a un investigador a encontrar el tamaño apropiado de una muestra, especialmente en el entorno de las investigaciones por encuestas.

Con tal objetivo, SOTAM dispone de un conjunto amplio de herramientas de decisión, ayudas y varias utilidades, gracias a las cuales el investigador no sólo puede tomar una decisión muy fundamentada sobre cuántas unidades seleccionar de la población de interés, sino además qué modelo de selección de muestras seguir, qué procedimiento de encuesta utilizar, qué unidades concretas seleccionar, etc. Todo ello con la correspondiente documentación en disco

generada específicamente para cada decisión.

Es recomendable que el usuario del sistema cuente con algunos conocimientos mínimos sobre muestreo y sobre investigación por encuestas. Si no es así, también puede utilizar el paquete, pero no obtendrá el mismo rendimiento, ni estará preparado para aprovechar todas las herramientas de que dispone SOTAM.

Las herramientas que componen el sistema pueden ser agrupadas en tres categorías: definición del entorno de trabajo, herramientas fundamentales y utilidades. Vamos a describirlas muy brevemente.

Las utilidades se exponen mediante la sucesión de opciones que permiten llegar a ellas. Una frase única indica que la opción se encuentra disponible directamente en el menú principal. Las barras inclinadas (/) separan una opción del menú de donde proceden. Por ejemplo, «6. Utilidades/2. Probabilidades y distancias» representa la utilidad de traducción de probabilidades a distancias estandarizadas y viceversa, que se encuentra disponible abriendo una ventana de menú, escogiendo la opción «6. Utilidades» del menú principal, y tomando la opción «2. Probabilidades y distancias» de la nueva ventana de menú.

### *Herramientas fundamentales.*

#### 1. Pantalla tanteo para tamaños

Es, sin duda, la utilidad principal del sistema, permite obtener un tamaño de muestra adecuado al modelo de muestra que se va a aplicar (aleatorio, estratificado,...), la medida objetivo (proporción, diferencia entre medias, ...), las medidas procedimentales (varianza poblacional, ...), el modelo de probabilidad que sigue la distribución muestral (normal o desconocida), el nivel de significación, etc.

Existen multitud de opciones por defecto, pero todas son modificables por el usuario. A su vez, esta herramienta permite acceder a utilidades tales como la creación de archivos de disco con matrices de tamaños de muestra en función de una variable de filas y una de columnas, que selecciona el usuario.

#### 2. Decisión procedimiento de encuesta

Aplicando recursos de la teoría de la decisión, esta herramienta permite decidir qué procedimiento se va a aplicar en la investigación por encuestas: cara a cara, por teléfono o por correo. Para ello, necesita del usuario una medida de importancia por cada objetivo de entre una lista de nueve de ellos. Con esta información y algunos otros recursos internos (que el usuario puede modificar), el sistema toma una decisión e informa de ésta.

#### 3. Decisión modelo de muestreo

Mediante un conjunto de preguntas cuyo contenido y orden depende de las respuestas del usuario, esta herramienta toma una decisión sobre el modelo de muestreo que mejor se

adecua a las circunstancias. Éstas quedan definidas por las características de la población y la información con que cuenta el investigador para realizar la encuesta.

#### 4. Pantalla tanteo para costes

Consta de un conjunto de partidas presupuestarias en pantalla. La introducción de valores modifica, de forma inmediata y automática, el monto o coste total. Determinadas opciones permiten acceder a algunas ayudas para definir mejor algunas de las partidas que entran en juego.

#### 5. Múltiples variables

En la mayoría de las investigaciones mediante encuestas, no existe una única variable reina, sino que más bien pueden resaltarse algunas características principales y otras secundarias. En definitiva, se cuenta con un conjunto de variables, hipótesis o procesos de inferencia. Cada uno de estos elementos puede aconsejar un tamaño de muestra diferente.

Esta herramienta ayuda a decidir un tamaño de muestra único, que se adecue a la situación.

#### 6. Utilidades/1. Recursos de optimización

La decisión sobre el tamaño de la muestra (y los aspectos tangenciales que ellos conlleva) puede cerrarse con la consulta al módulo de tanteo para tamaños de muestra o, por el contrario, culminar tras un proceso más largo, pero también más asentado.

Pongamos por caso que el investigador ha tanteado los tamaños que corresponden a determinados modelos de muestreo y que, igualmente, cuenta con sendas alternativas para el cálculo de los costes. En tal caso, mediante la herramienta de los recursos de optimización, puede establecer comparativas entre los diferentes ensayos y decidirse finalmente por un modelo de muestreo, con su coste y tamaño asociados.

Otra de las utilidades disponibles mediante el menú de recursos de optimización permite definir un coste total en función del cambio que se provoca en otras variables. Por ejemplo, se puede poner en marcha un procedimiento que calcule en nivel de significación asociado a un determinado coste. Acto seguido se provocan cambios en el coste y se observan las consecuencias en el nivel de significación. El proceso de ensayo culmina cuando los cambios en el coste no van seguidos de modificaciones significativas en el nivel de significación.

#### 6. Utilidades/6. Generación de informes

El menú de generación de informes permite obtener información, en archivos de disco escritos en código ASCII, que justifica las decisiones tomadas con respecto al tamaño de la muestra, el procedimiento de encuesta o el modelo de muestreo.

De esta forma, el investigador puede utilizar un módulo fundamental y, acto seguido, activar la generación de un informe que justifique el proceso y resultado (decisión) que han

tenido lugar durante y tras la interacción con la herramienta.

### *Utilidades*

#### 6. Utilidades/2. Probabilidades y distancias

A través de una pantalla del tipo formulario, el usuario puede traducir probabilidades a distancias estandarizadas y viceversa. Para ello, se contemplan tres modelos de probabilidad diferentes: normal, Student y desconocida (acotación de Chebyshev); así como cuatro esquemas de traducción alternativos: rechazo una cola, rechazo dos colas, intervalo centrado o distribución acumulada. Las interacciones pueden grabarse en disco.

#### 6. Utilidades/3. Generación aleatoria/1. Matriz de números aleatorios

Construye archivos de disco con matrices de números aleatorios. El usuario controla la cantidad de números (filas y columnas), el formato de éstos, la existencia o no de una cabecera informativa, el modelo de probabilidad (normal, binomial, exponencial,...), el generador básico (Marsaglia, ...), etc. Como en el resto de herramientas, se pueden controlar específicamente todos los parámetros de la generación o dejar las opciones por defecto.

#### 6. Utilidades/3. Generación aleatoria/2. Números de teléfono

La utilidad permite generar un listado de números de teléfono según las indicaciones del usuario: prefijos, estructura en dígitos, etc. Consigue, con ello, un listado de números a utilizar en una investigación con encuestas por teléfono.

#### 6. Utilidades/3. Generación aleatoria/3. Tablas para encuestadores

Especialmente en las investigaciones por encuesta donde se ha seleccionado un modelo de conglomerados con submuestreo, es de amplia aplicación la selección de unidades mediante el procedimiento denominado *rutas aleatorias*. El encuestador construye sobre la marcha un itinerario y toma decisiones sobre qué calle transitar, qué viviendas de qué pisos de qué escaleras de qué edificios encuestar.

Con el fin de que el encuestador tome decisiones objetivas, es necesario que éstas sean aleatorias, es decir, se realicen sin depender de las subjetividades de cada encuestador. Para ello, se recurre a tablas de números aleatorios. Así, si el encuestador debe seleccionar una vivienda dentro de un edificio de 8 plantas, consulta su tabla aleatoria, la fila del número de orden de la encuesta a realizar y la columna 8 que corresponde a las situaciones en las que debe seleccionar un elemento de un conjunto de 8. La casilla de la tabla le indica la planta seleccionada.

Tales decisiones, por tanto, requieren la existencia de tablas aleatorias. La utilidad que estamos presentando aquí, genera tablas en archivos de disco, donde el investigador controla el número de tablas, fila y columnas el criterio de clasificación (por ejemplo, la columna 8 no puede tener número superiores a 8), el formato estético de la tabla, etc.

#### 6. Utilidades/4. Selección de unidades

En función del modelo de muestreo que se vaya a aplicar, esta herramienta permite seleccionar unidades. Así, por ejemplo, si el modelo es aleatorio, la población tiene un tamaño de 500 mil elementos y la muestra es de tamaño 600, la herramienta escribe un archivo de disco con 600 números aleatorios comprendidos entre el 1 y el 500000. Si el muestreo es estratificado, además decide el tamaño de cada estrato y, dentro de cada uno, establece qué unidades deben ser seleccionadas. De forma análoga, se adecua a los modelos por conglomerados.

#### 6. Utilidades/5. Tamaños para dominios

Si la investigación se realiza sobre una población compuesta por subpoblaciones y se requiere hacer inferencia para cada subpoblación, éstas reciben generalmente la denominación de dominios.

Una de las salidas es decidir un tamaño para toda la población y después consultar esta herramienta. Mediante ello, el sistema comunica al investigador cuántas unidades debe seleccionar de cada dominio para que el nivel de seguridad en las estimaciones sea el mismo en todos ellos.

#### 6. Utilidades/7. Utilización de software comercial

La práctica totalidad del software estadístico comercializado, únicamente considera un modelo de selección aleatorio simple. Si la investigación mediante encuestas se ha realizado utilizando otro modelo, si los análisis de inferencia se realizan con el software estadístico al uso, se están manejando estimaciones incorrectas para los errores típicos.

Esta herramienta aconseja al investigador qué acciones debe realizar en su programa habitual para que éste realice inferencias contemplando el modelo de muestreo real y no sólo el aleatorio simple para poblaciones infinitas.

### CONTROL DEL ENTORNO DE TRABAJO

#### 6. Utilidades/8. Visualización de archivos

Esta herramienta es un lector de archivos ASCII. El usuario puede desplazarse por la pantalla hacia delante y atrás, pero no editarla, es decir, se pueden leer los archivos, pero no modificarlos.

#### 7. Entorno/1. Colores del entorno

El sistema, en su interacción con el usuario, diferencia con colores los diferentes elementos que intervienen: color de fondo principal, del fondo de ventana, del texto, de las respuestas del usuario y de las teclas de opción. Esta herramienta permite variar tales colores.

### 7. Entorno/2. Control de inicio

Al volver a iniciar una sesión con SOTAM puede comenzarse con las opciones por defecto o con las opciones que se encontraban seleccionadas al finalizar la última sesión. Esta opción controla cuál es el comportamiento del sistema al respecto.

### 7. Entorno/3. Tipo usuario

Algunas opciones del sistema sólo están disponibles si el entorno se encuentra configurado en modo usuario experto. Hay dos estados posibles: usuario experto o inexperto. La distinción es importante por cuanto que un mayor control sobre el proceso también implica mayores exigencias en cuanto a información que debe suministrar el usuario, mayor nivel de conocimientos y experiencia de éste y más trascendencia en las decisiones.

En un apartado posterior, entraremos de nuevo en la distinción entre modos de actuación según las características del usuario.

### 7. Entorno/4. Opciones por defecto

Inicialmente, el sistema viene de fábrica con unas opciones por defecto. El usuario puede modificar éstas mediante la utilidad de que trata este punto. Siempre queda la oportunidad de recuperar los valores originales.

### 7. Entorno/5. Señal audible

Cuando el usuario pulsa una tecla equivocada o selecciona una opción desactivada, puede emitirse un sonido de aviso, agudo y breve, desde el ordenador. Esta opción permite activar o desactivar esa señal audible o aviso por acción inadmisibles.

### 7. Entorno/6. Otro archivo de salida

Tal y como se explicita de nuevo en un apartado venidero (primeros pasos) las acciones que se realizan con el sistema tienen eco en un archivo de disco denominado *archivo de registro* o *archivo de salida*. Está escrito en código ASCII, por lo que puede ser consultado con la inmensa mayoría de los procesadores y editores de texto, incluyendo la utilidad «6. Utilidades/8. Visualización de archivos».

### 8. Ejecutar programa del DOS

Esta opción permite ejecutar un programa como si el usuario se encontrara trabajando directamente en la línea de comandos del sistema operativo. Una vez finalizada la ejecución del programa, se vuelve a SOTAM previa pulsación de una tecla (según el mensaje que aparece en pantalla).

## 6.- Referencias

- Alberdi, J.; Lorente, S. y Moreno, E. (1969). *Metodología de investigación por muestreo*. Madrid: Euramérica.
- Azorín, F. (1969). *Curso de muestreo y aplicaciones*. Madrid: Aguilar.
- Boulanguer, G. (1971). *La investigación en ciencias humanas*. Madrid: MAROVA.
- Bradburn, N. M; Sudman, S. (1988). *Polls and surveys. Understanding what they tell us*. The Jossey-Bass Social and Behavioral Science & Public Administration Series. Nueva York: John Wiley & Sons.
- Carpintero, H. (1989). Comentarios a la enseñanza de la estadística en psicología. El aprendizaje de la estadística. *Estadística Española*, 31, 122, 476-479.
- Cochran, W. G. (1976). *Técnicas de muestreo*. México: Compañía Editorial Continental S.A..
- Czaja, R. y Blair, J. (1996). *Designing surveys. A guide to decisions and procedures*. Thousand Oaks, California: Pine Forge Press.
- Chevry, G. R. (1962). *Practique des enquêtes statistiques*. Paris: Presses Universitaires de France.
- Davis, J. A. (1971). *Elementary Survey Analysis*. Nueva Yersey: Prentice-Hall.
- Fariña, F. y Arce, R. (1993). *Ansiedad ante los ordenadores*. Recursos Humanos. Serie Psicología. Madrid: Eudema.
- Feigl, P. (1978). A graphical aid for determining sample size when comparing two independent proportions. *Biometrics*, 34, 111-122.
- Fink, A. y Kosecoff, J. (1989). *How to conduct surveys: a step-by-step guide*. Beverly Hills: Sage.
- Ghiglione, R. y Matalon, B. (1989). *Las encuestas sociológicas. Teoría y práctica*. México: Trillas.
- Glaxo (1994). *N. Programa para el cálculo del tamaño muestral en diseños de experimentación*. Unidad Estadística. Departamento Médico. Madrid: Glaxo.
- Goode, W. J. y Hatt, P. K. (1977). *Métodos de investigación social*. México: Trillas.
- Harvatopoulos, Y.; Livan, Y. F. y Sarnin, Ph. (1992). *El arte de la encuesta. Principios básicos para no especialistas*. Bilbao: Ediciones Deusto.
- Hedges, B. (1980). Sampling. En G. Hoinville y R. Jowell (Eds.). *Survey research practice* (pp. 55-89). London: Heinemann Educational Books.
- Henry, G.T. (1990). *Practical sampling*. Applied Social Research Methods Series, vol 21. Newbury Park, California: Sage Publications.
- Hyman, H. (1970). *Diseño y análisis de las encuestas sociales*. Buenos Aires: Amorrortu Editores.
- Jolliffe, F.R. (1986). *Survey Design and Analysis*. Mathematics and it's applications. Chichester: Ellis Horwood Limited.
- Judd, Ch. M., McClelland, G. H. y Culhane, S. E. (1995). DATA ANALYSIS: Continuing Issues in the Everyday Analysis of Psychological Data. *Annual Review of Psychology*, 46, 433-465.
- Kachelmeier, S. J. y Messier, W. F. Jr. (1990). An investigation of the influence of a nonstatistical decision aid on auditor sample size decisions. *Accounting review*, 65, 1, 209-226.
- Kalton, G. (1987). *Introduction to survey sampling*. SAGE University Paper. Beverly Hills: SAGE Publications Inc.

- Kish, L. (1965). *Survey Sampling*. New York:: John Wiley & Sons.
- Lebart, L.; Morineau, A.; y Fénelon, J. P. (1985). *Tratamiento estadístico de datos. Métodos y programas*. Barcelona: Marcombo.
- Lininger, Ch. A. y Warwick, D. P. (1984). *La encuesta por muestreo: teoría y práctica*. México: Compañía Editorial Continental.
- Manheim, H. L. (1982). *Investigación sociológica. Filosofía y Métodos*. Barcelona: CEAC.
- Manzano, V. (1997). SOTAM: Sistema de Optimización para el Tamaño de la Muestra. V Congreso de Metodología de las CC. Humanas y Sociales. Sevilla, Septiembre de 1997.
- Moreno, R.; Martínez, R; y Chacón, S. (1997). DEFINE: Un programa tutor en HTML para la delimitación de conceptos. V Congreso de Metodología de las CC. Humanas y Sociales. Sevilla, Septiembre de 1997.
- Noelle, E. (1970). *Encuestas en la sociedad de masas. Introducción a los métodos de la demoscopia*. Madrid: Alianza Editorial.
- Pérez Cebrián, F. (1987). *La planificación de la encuesta social*. Zaragoza: Secretariado de prensas universitarias de la Universidad de Zaragoza.
- Presser, S. y Traugott, M. (1992). Correlated response errors in a panel study of voting. *Public Opinion Quarterly*, 56, 1, 77-86.
- Raj, D. (1972). *The design of sample survey*. Nueva York: McGraw-Hill.
- Rodrigues, A. (1977). *Investigación experimental en psicología y educación*. México: Trillas.
- Rossi, P.H.; Wright, J.P. y Anderson, A.B. (1983). Sample survey: history, current practice, and future prospect. En P.H. Rossi, J.P. Wright y A.B. Anderson (Eds.). *Handbook of survey research* (pp. 1-20). Quantitative Studies in Social Relations. Orlando, Florida: Academic Press.
- Salgado, J.A. (1990). La práctica del muestreo. En E. Ortega (Eds.). *Manual de investigación comercial* (pp. 344-377). Madrid: Pirámide.
- Santesmases, M. (1997). DYANE: Diseño y análisis de encuestas en investigación social y de mercados. Madrid: Pirámide.
- Silva, L.C. (1993). *Muestreo para la investigación en ciencias de la salud*. Madrid: Diaz de Santos.
- Solanas, A. y Sierra, V. (1992). Bootstrap: fundamentos e introducción a sus aplicaciones. *Anuario de Psicología*, 55, 143-154.
- Stoetzel, J. y Girard, A. (1973). *Las encuestas de opinión pública*. Madrid: Instituto de la Opinión Pública.
- Stuart, A. (1984). *The ideas of sampling*. High Wycombe, England: Charles Griffin.
- Sudman, S. (1976). *Applied sampling*. Quantitative Studies in Social Relations. Nueva York: Academic Press.
- Sudman, S. (1983). Applied sampling. En P.H. Rossi, J.P. Wright y A.B. Anderson (Eds.). *Handbook of survey research* (pp. 145-197). Quantitative Studies in Social Relations. Orlando, Florida: Academic Press.
- Sukhatme, P.V. (1953). *Sampling theory with applications*. Iowa: Iowa State College Press.
- Teijeiro, F. (1990). Técnicas de muestreo. En E. Ortega (Eds.). *Manual de investigación comercial* (pp. 312-343). Madrid: Pirámide.
- Timothy, R. (1989). Variations on a seminal demonstration of people's insensitivity to sample size. *Organizational behavior and human decision processes*, 43, 1, 52-57.
- Veres, E. (1981). Tamaño muestral óptimo de maximización del valor esperado de un

- experimento. *Estadística española*, 92,47-65.
- Vallecillos, A. (1996). *Inferencia estadística y enseñanza: una análisis didáctico del contraste de hipótesis estadísticas*. Colección Mathema. Granada: Comares.
- Well, A.; Pollatsek, A. y Boyce, S.J. (1990). Understanding the effects of sample size on the variability of the mean. *Organizational behavior and human processes*, 47, 2, 289-312.
- Yates, F. (1981). *Sampling methods for censures and surveys*. High Wycombe, England: Charles Griffin.