

# **Tipos de variables y sus implicaciones en el diseño de una investigación**

• Universidad Pontificia Comillas • Madrid •  
*Facultad de Ciencias Humanas y Sociales*  
©Pedro Morales (última revisión, 25, Abril, 2009)

## **Índice**

1. Introducción: importancia de las variables en la preparación de instrumentos y diseños .....	3
2. Variables en cuanto características de la situación y de los sujetos .....	3
a) <i>Condiciones experimentales</i> .....	3
b) <i>Características personales de los sujetos</i> .....	4
2.1. Variables independientes y dependientes.....	4
2.2. Variables moderadoras y variables controladas .....	5
a) Variables moderadoras.....	5
b) Variables controladas.....	6
2.3. Variables confundidas .....	6
2.4. Variables desconocidas .....	7
2.5. Tipos de variables en la investigación experimental: visión de conjunto.....	7
3. Implicaciones de los tipos de variables al planificar la investigación .....	8
3.1. Implicaciones en los instrumentos de recogida de datos.....	9
3.2. Implicaciones en la elección de la muestra .....	11
3.3. Implicaciones en los métodos de análisis estadístico.....	12
4. Variables en cuantos tipos de datos o escalas .....	13
5. Referencias bibliográficas.....	14



## 1. Introducción: importancia de las variables en la preparación de instrumentos y diseños

Es importante conocer qué se entiende por *variable* y cuáles son los *distintos tipos* de variables. El clarificar el término de variable y sus diversas acepciones y modalidades nos va a ayudar a:

- 1° Preparar mejor, de manera más completa, nuestros *instrumentos* de recogida de datos. Si no pensamos previamente en los diversos tipos de variables que nos pueden interesar, podemos descuidar el recoger información adicional o necesaria para llevar a cabo nuestra investigación; no es infrecuente el que algunos caigan en la cuenta de que les faltan datos importantes cuando ya es demasiado tarde.
- 2° Pensar y escoger el *diseño* de investigación más apropiado y el *método adecuado para analizar* los datos.

El clarificar el concepto de variable y los distintos tipos de variables es muy importante como *punto de partida* al iniciar una investigación; no se trata de una mera disquisición académica o de una complicación terminológica más o menos irrelevante. En educación y en psicología solemos centrarnos inicialmente en un *tema* de interés, en un rasgo o característica que queremos estudiar, pero hay también otras variables o características de los sujetos que acompañan necesariamente al rasgo central que queremos investigar.

Este pensar *en otras cosas* desde el principio nos va a ayudar tanto a planificar el *instrumento* de recogida de datos como a buscar la *muestra* adecuada y a sopesar *distintos métodos para analizar nuestros datos*.

## 2. Variables en cuanto características de la situación y de los sujetos

Básicamente podemos distinguir dos acepciones del término variable:

- 1° En cuanto *características* de los sujetos,
- 2° En cuanto *tipos de datos* o modos de obtenerlos; en este mismo sentido se habla de *tipos de escalas*.

Aquí nos centramos sobre todo en la primera acepción; es indudablemente la que más nos interesa ahora. Variable en cuanto característica de los sujetos es el sentido habitual del término variable. Una *variable* es lo que podemos *observar, codificar o cuantificar* en los sujetos sobre los que investigamos. El término variable viene del hecho de que los sujetos pueden ser distintos con respecto a distintas variables. Las variables en cuanto tipo de datos (*escalas nominales, ordinales, de intervalo y de razón*) lo tratamos brevemente al final (apartado nº 4).

El término variable, en el contexto de la investigación experimental, puede referirse a ámbitos muy distintos.

### a) *Condiciones experimentales*

El término variable en cuanto *características* de los sujetos aplicado a un método, terapia, pertenencia a un grupo, etc., puede parecer poco apropiado, sin embargo sí podemos decir que una característica de un sujeto es el haber seguido un determinado método, el haber pasado por una experiencia, el pertenecer a un determinado grupo, etc. De hecho estos son los términos que se utilizan.

Condiciones experimentales pueden ser situaciones, pertenencia a grupos, etc. Por ejemplo un método didáctico es una variable (los sujetos pueden seguirlo o no seguirlo), el cursar un tipo u otro de carrera puede ser otra variable. Estas variables, como veremos enseguida, suelen ser las variables denominadas *variables independientes*.

### **b) Características personales de los sujetos**

Características personales de los sujetos son las variables *demográficas*, como sexo, edad, estado civil, profesión, etc., y también son características personales los *rasgos* que solemos denominar *constructos*: inteligencia, rendimiento académico, actitudes, rasgos de personalidad, valores, etc. Estos *rasgos que medimos* suelen ser las *variables dependientes* aunque también, como veremos, pueden tratarse como variables independientes.

Estos rasgos los medimos y cuantificamos mediante tests, escalas, cuestionarios, observación directa, etc. Cada instrumento (o cada pregunta) *mide* una variable que representa un *constructo*, que lo hace observable y medible (medible entendido en un sentido amplio).

No siempre la *variable medida* se ajusta con propiedad a un determinado constructo previamente especificado (por eso analizamos y depuramos nuestros instrumentos, escalas, etc.), pero siempre podemos pensar en el constructo que *subyace* a una determinada variable que medimos con un instrumento determinado (e incluso podemos *redefinir* el constructo en función del instrumento utilizado).

Las variables *operacionalizan* los constructos, los definen y los hacen medibles. Pero esto se puede hacer con mayor o menor precisión. Una variable tal como la medimos (un test, una escala, etc.) puede *operacionalizar* (hacer medible...) mal un constructo o medir más de uno.

También es verdad que *los constructos (rasgos humanos) se pueden concebir en diversos niveles de abstracción*, pueden ser muy genéricos (gusto por aprender lenguas extranjeras) o muy específicos (gusto por aprender inglés); la *constancia* (ser constante) puede ser un constructo muy específico, y el ser *constante* y *además* ser *organizado*, y tener *objetivos a largo plazo* y valoran *la independencia y responsabilidad individual* (y todo esto medido con el mismo instrumento) puede responder a la definición de un constructo más complejo (*motivación de logro* en este caso). Los constructos pueden definirse por lo tanto a nivel más específico o más genérico, como de manera análoga se puede hablar de peras o de manzanas (nivel específico) o de frutas (englobando peras y manzanas; nivel genérico).

Conviene analizar bien los instrumentos de medición, clarificar en lo posible qué miden y definir o explicar bien el *constructo implícito* en nuestro instrumento.

Por lo que respecta a los instrumentos de medida que preparamos para nuestra investigación (*cuestionarios*), hay constructos que podremos medir de manera muy sencilla (simples preguntas, una lista de adjetivos en los que cada uno representa un constructo o rasgo) o de manera más elaborada (como un test o una escala de actitudes, en los que el mismo constructo se operacionaliza a través de varios indicadores o ítems). Más adelante damos sugerencias más específicas.

## **2.1. Variables independientes y dependientes**

Las *variables independientes* son las que elegimos libremente, o *manipulamos*, para verificar su efecto en, o su relación con, las *variables dependientes*.

Si queremos comprobar la eficacia de un método en un tipo de aprendizaje, el método es la variable independiente (también denominada en casos como éste *variable experimental*) y el aprendizaje (definido por el modo, test, etc. con que lo medimos) es la *variable dependiente*. El

objetivo directo de una investigación es clarificar la relación entre las variables independientes y las variables dependientes.

La variable independiente puede estar dividida en *niveles* o *subclasificaciones*: si la variable independiente es el *género*, estará dividida en dos niveles (*varones* y *mujeres*); si la variable independiente es la *carrera* que se está estudiando, los niveles serán las diversas carreras que entren en la investigación; si la variable independiente es el *método*, los niveles serán las modalidades metodológicas que entren en la investigación (o con un método particular y sin ese método).

Las variables independientes y dependientes se suelen conceptualizar también como *predictores* y *criterios*. Simplificando mucho, estos conceptos se entienden bien si pensamos en términos de *causas* y *efectos*, pero cayendo en la cuenta de que todo lo más se trata de *hipotéticas* causas, *concausas* o *condiciones*. Variables *antecedentes* y variables *consecuentes* pueden ser términos más adecuados.

En los diseños experimentales, las variables independientes suelen ser *cualitativas* (distintos métodos, experiencias, condiciones, etc.) y la variable dependiente suele ser *cuantitativa* aunque no necesariamente (*medimos* los resultados o el cambio mediante un test, una escala, un cuestionario que nos da unas puntuaciones, etc.).

## **2.2. Variables moderadoras y variables controladas**

En el efecto que una variable independiente tiene en la variable dependiente, pueden influir otras variables. Por ejemplo en la eficacia en el aprendizaje (*variable dependiente*) de un método didáctico (*variable independiente*) puede influir la hora de la clase, el género del alumno, los conocimientos previos, etc.; estas serían las *variables moderadoras*.

Estas variables pueden ser conocidas o previstas por el investigador. En la medida en que se trata de variables conocidas, que pueden influir de alguna manera en la variable dependiente, y que por lo tanto tenemos en cuenta en nuestros análisis, son también variables independientes (*variables independientes moderadoras*), aunque no sean estas las variables independientes el objeto principal de la investigación.

Estas variables pueden ser utilizadas directamente en la investigación y podemos determinar su relación o influjo en la variable dependiente (*variables moderadoras*) o pueden ser neutralizadas para que no interfieran en la investigación (*variables controladas*).

El investigador puede pensar que el influjo del método puede depender, en parte al menos, del sexo del alumno, o de la hora de la clase (por la mañana temprano o al final de la mañana o de la tarde). Son variables que el investigador puede utilizar y analizar o que puede controlar. Son también variables independientes aunque la variable principal objeto de estudio sea otra (como puede ser un método didáctico cuya eficacia en la variable dependiente queremos comprobar).

Con estas otras variables el investigador tiene dos opciones: comprobar su eficacia, su influjo, o prescindir de ellas, neutralizándolas.

### **a) Variables moderadoras**

Las variables *moderadoras* son las variables independientes que el investigador *utiliza* para comprobar si influyen en la relación entre la variable independiente y la variable dependiente. Son variables independientes porque el investigador las escoge libremente, las utiliza en su investigación para comprobar su relación o influjo en la variable dependiente.

Por el ejemplo, el investigador puede comprobar si un método didáctico (variable independiente principal) funciona de manera distinta en los dos sexos, y puede consiguientemente incluir el sexo en el análisis, y comprobar, por ejemplo, si el efecto del método es distinto según el sexo del alumno. Estas variables suelen denominarse *moderadoras*, porque moderan, modulan, etc. el influjo del método (en este ejemplo es la variable independiente principal) en el aprendizaje (variable dependiente).

### b) *Variables controladas*

Puede haber otras variables que el investigador prefiere no incluir en el estudio, no pretende estudiar su influjo en la variable dependiente. Para que no interfieran en la interpretación de los resultados, las *neutraliza*. Son *variables controladas*.

Hay muchas maneras de neutralizar variables en los distintos diseños. La manera más sencilla de neutralizar una variable es eliminarla, no incluirla en el estudio (todos varones, o todos con los mismos conocimientos previos, etc.) y tener en cuenta esta *no inclusión* al interpretar los resultados.

En términos generales tenemos tres vías para *controlar o neutralizar* variables<sup>1</sup>:

1. Mediante el *muestreo aleatorio o probabilístico*. Una muestra es aleatoria si todos los sujetos de la población investigada han tenido idéntica oportunidad de ser seleccionados; en este caso ninguna característica estará representada en la muestra en una proporción mayor o menor de lo que está en la población. Es importante tener una idea clara de los diversos tipos de muestras aleatorias y cómo hacer un muestreo aleatorio.
2. Mediante el *diseño apropiado*. Por ejemplo igualando sujetos en la variable (una o más de una) que queremos controlar (*matching* en inglés; tendríamos en este caso muestras relacionadas aunque se trate de sujetos físicamente distintos), o mediante un diseño de *bloques igualados*; en este caso el método de análisis sería un *análisis de varianza factorial* en el que uno de los factores es la variable que queremos controlar: en cada nivel de esa variable tenemos un grupo de sujetos igualados en la variable o variables que queremos controlar.
3. Mediante *análisis estadísticos*, como pueden ser las *correlaciones parciales* y el *análisis de covarianza* (una *condición necesaria* en el *análisis de covarianza* es que los sujetos estén *asignados aleatoriamente* a las diversas condiciones, como pueden ser grupo experimental y de control).

### 2.3. *Variables confundidas*

Algunas variables o características pueden estar confundidas (*confounded*) con la o las variables independientes.

Por ejemplo el participar *voluntariamente* en una experiencia cuya eficacia se va investigar, puede estar *confundido* con características personales de los sujetos voluntarios (como podría ser una determinada edad, una determinada motivación, etc., que tendrán muchos participantes pero no todos) que a su vez pueden influir en la eficacia de la experiencia. Con frecuencia la pertenencia a un grupo, implica la pertenencia a otros grupos (entendiendo grupo en un sentido muy genérico) o el participar de alguna característica más que otros grupos (por

---

<sup>1</sup> Sobre el control de variables puede verse en la bibliografía el documento de Morales (2009) *El control de variables: control estadístico (análisis de covarianza) y control experimental mediante diseño*

ejemplo en la variable *nivel socioeconómico* pueden estar *confundidas* otras variables como nivel educacional).

Esta *confusión* de variables puede oscurecer las conclusiones: puede ser que lo que realmente influye más no es la experiencia o actividad investigada, sino otras características que *acompañan* a los que participan en la experiencia; al menos estos influjos pueden ser importantes.

No es fácil a veces separar las variables *confundidas*, pero sí podemos obtener datos sobre variables *sospechosas* y tenerlas en cuenta, o hacer un muestreo estratificado para neutralizar algunas de estas variables, etc.

#### 2.4. Variables desconocidas

Muchas características individuales nos son desconocidas, al menos por dos razones:

- a) Porque no podemos obtener datos de todo,
- b) Porque ni siquiera se nos ocurre que algunas características pueden ser relevantes y pueden estar influyendo en el efecto estudiado.

Cada persona lleva consigo su propia historia (familiar, social, etc.), y las personas son tan únicas que es imposible conocer o controlar todas las variables que pueden interferir en nuestros resultados. La mejor solución para controlar estas variables desconocidas es el *muestreo aleatorio*.

Mediante el *muestreo* adecuado (aleatorio simple, estratificado, etc.) podemos controlar también variables conocidas pero *irrelevantes* en nuestro estudio y cuyo posible influjo no queremos analizar en una situación dada (experiencia previa, edad, momento en el que se recogen los datos, etc.).

El *muestro aleatorio* no es tampoco una solución definitiva en estudios experimentales (sí puede serlo en la práctica en un sondeo de opiniones en la población general) por eso diseños experimentales en sentido propio, que permitan mediante el control todas las variables extrañas, detectar *relaciones puras* entre variables independientes y dependientes, no son posibles.

La existencia e influjo de estas *variables desconocidas* es una limitación de la investigación en las ciencias humanas; por eso ahora el énfasis no se pone en *estudios experimentales perfectos*, sino en la *acumulación de resultados* de estudios semejantes (*meta-análisis*), que permiten llegar a conclusiones más generalizables y matizadas (y siempre tentativas) sobre el *estado de la cuestión* en un momento dado.

#### 2.5. Tipos de variables en la investigación experimental: visión de conjunto

Es útil tener disponer de una visión de conjunto de los distintos tipos de variables que pueden estar presentes en nuestra investigación (figura 1).

Esta visión de conjunto *puede ayudarnos a preparar nuestros instrumentos de recogida de datos*; nos va a ayudar a determinar:

- a) Cuáles son las variables independientes principales que van a centrar y dar estructura a nuestra investigación,
- b) De qué otras variables moderadoras puede ser conveniente obtener información,
- c) Qué variables nos interesa controlar, etc.

Con esta finalidad (pensar en qué incluimos en nuestro instrumento, cómo elegimos a la muestra, qué tipos de análisis podemos considerar, etc.) presentamos en la figura 1 una síntesis de los diversos tipos de variables, para tener una visión de conjunto que facilite nuestra reflexión.

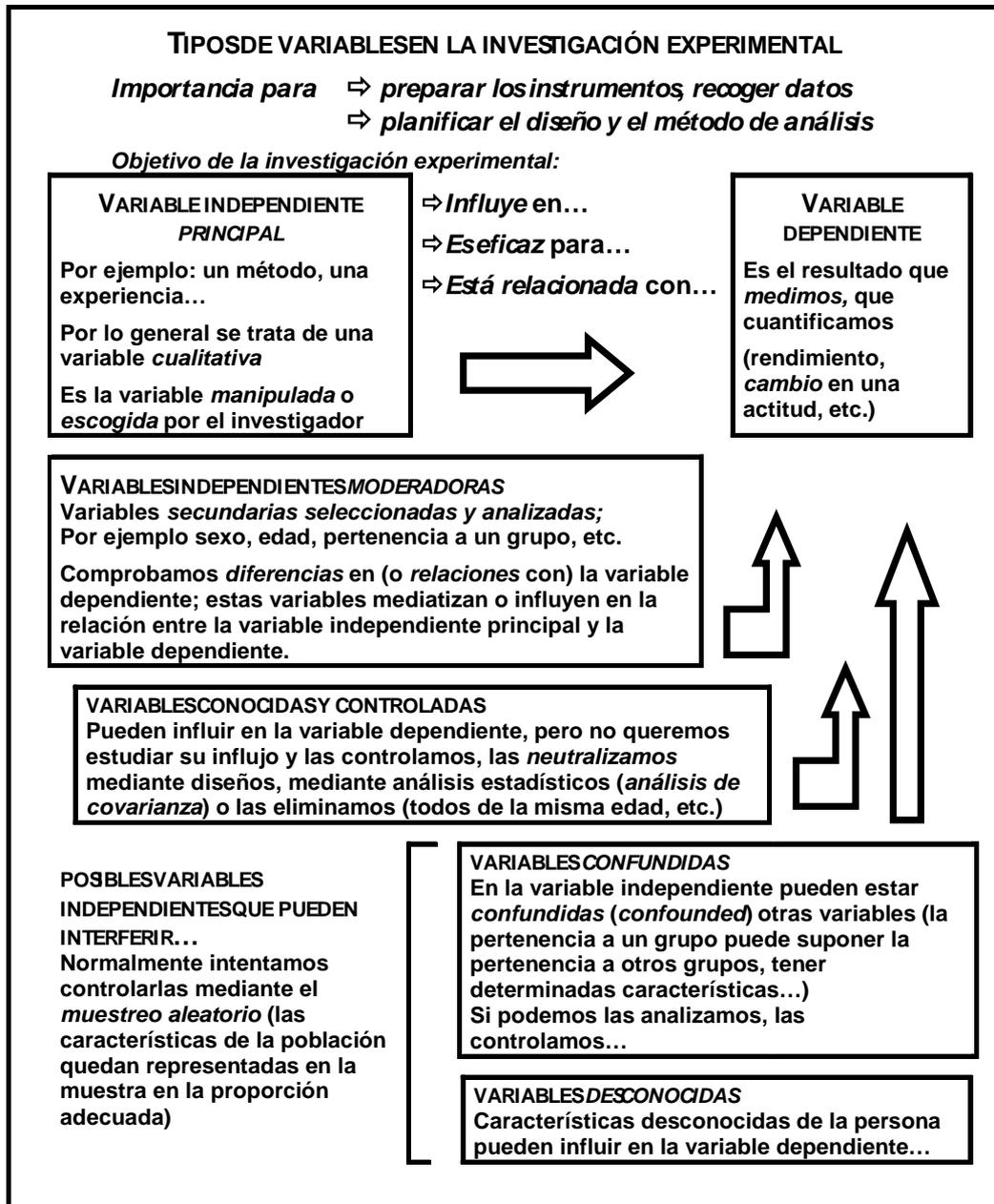


Figura 1

### 3. Implicaciones de los tipos variables al planificar la investigación

Normalmente cuando se inicia una investigación, el interés suele centrarse en pocas variables; normalmente en una variable independiente principal y en alguna o algunas variables dependientes.

Por ejemplo, se desea investigar la satisfacción de los alumnos de una universidad en función del tipo de carrera que están estudiando. La variable independiente en este caso sería la carrera que se está estudiando (una variable independiente, carrera, dividida en varios *niveles* que

serían las diversas carreras), y la variable dependiente será la satisfacción, medida con una escala o con cuestionario apropiado (o incluso con una sola pregunta).

El investigador debe preguntarse en primer lugar ¿Qué otras variables pueden estar influyendo en la satisfacción además de la carrera que se está estudiando? Puede tratarse de hipótesis plausibles, o de simples conjeturas, pero pronto saldría (y conviene hacer) un listado de *posibles variables moderadoras* que se podrían incluir en la investigación, incluso por simple curiosidad.

A título de ejemplo podemos mencionar estas variables que pueden estar relacionadas con la satisfacción con la carrera:

- \*El género del alumno (en algunas carreras las chicas pueden sentirse mejor que en otras),
- \*El tipo de motivación para estudiar esa carrera (y se puede pensar en una breve tipología de motivaciones),
- \*La personalidad del alumno,
- \*La profesión de sus padres,
- \*La satisfacción no ya con la carrera en general sino con aspectos específicos (profesorado, medios, incluso horario de mañana o de tarde, etc.).
- \*Determinados valores, actitudes... (nivel de aspiraciones, autoeficacia, motivación de logro, etc.)

*Uno de los primeros pasos en la preparación de una investigación debe ser por lo tanto hacer una lista con las variables de interés potencial para la investigación.*

Ya se han ido señalando algunas implicaciones de los diversos tipos de variables en varios aspectos o fases de la investigación; merece la pena resumir y subrayar algunas de estas implicaciones en la práctica de la investigación; son precisamente estas *implicaciones prácticas* las que hacen necesario un estudio inicialmente más teórico y conceptual de las variables.

### **3.1. Implicaciones en los instrumentos de recogida de datos**

Muchas investigaciones en el ámbito de las ciencias sociales parten de un cuestionario que va a responder una serie de sujetos. Estos cuestionarios son de la máxima importancia, porque todo va a salir de ese cuestionario. *Lo que no se pregunta no existe*, y a veces se cae en la cuenta *demasiado tarde* de lo que se debería haber preguntado y no se preguntó por no haberlo pensado a tiempo.

Esta lista *inicial* de variables se puede después someter a examen: las habrá más y menos importantes, las habrá centrales y las habrá secundarias. No se trata de medir cada variable con largos tests. Muchas se podrán medir con simples preguntas (sexo, pertenencia a grupos, etc.), o con unos pocos indicadores, sobre todo posibles *variables independientes moderadoras*.

Es útil tener a la vista (tanto para medir las variables independientes como las posibles variables dependientes) una *tipología de tipos de preguntas o de cuestionarios*. Con frecuencia la variable dependiente principal requerirá un instrumento más elaborado, en buena medida para asegurar mejor la fiabilidad, para poder analizar el mismo constructo (por ejemplo con un análisis factorial) o dividirlo en subconstructos, pero también caben instrumentos y preguntas adicionales muy simples para obtener datos sobre variables de interés.

Una tipología sencilla de tipos de cuestionarios o de simples preguntas puede ser ésta<sup>2</sup>:

- Escalas y tests completos (ya hechos o de construcción propia),
- *Subescalas* de tests (para medir aspectos de especial interés),
- Unos pocos *indicadores* o ítems en torno a la misma variable, sin que se trate de una escala o test en sentido propio (pocas preguntas cuyas respuestas tenga sentido sumar en un total),
- Simples preguntas (como es normal en cuestionarios sociológicos),
- Listas de adjetivos (para medir diversos rasgos de personalidad o de temperamento),
- Listas de *motivaciones, problemas, valores*, etc. (una serie de elementos del mismo ámbito cuya importancia se puede valorar con respuestas idénticas, por ejemplo desde *nada importante* hasta *muy importante*).

Es muy útil tener a la vista modelos de cuestionarios ya utilizados por otros, publicados en libros y revistas, etc. Además de pensar qué variables nos puede interesar medir, tenemos que escoger el procedimiento posible y adecuado de medirlas.

En la figura 2 recogemos las ideas más importantes para pensar en la conveniencia de buscar otros datos sobre variables que pueden tener que ver con el objetivo principal de nuestra investigación.

**Qué INFORMACIÓN ADICIONAL debemos recoger**

**Siempre tenemos una o más variables dependientes objeto principal de nuestro estudio**  
(*determinados conocimientos, actitudes, valores, etc.*)

**PERO ADEMÁS nos puede interesar disponer de *más datos* sobre otras variables que también podemos *medir*:**

- Que pueden estar relacionadas con la variable dependiente (*resultados más altos o más bajos en función del sexo, edad, profesión, determinadas actitudes o valores, etc.*)
- Que pueden constituir *explicaciones rivales* a nuestra propia hipótesis; para poder controlarlas o analizarlas
- Que pueden afectar a determinados segmentos de nuestra muestra.
- Que pueden enriquecer nuestros análisis con un coste muy bajo

**Estrategia**

1. Hacer un *listado inicial* de variables relevantes
2. Valorarlas, jerarquizarlas

Nos ayudará a concretar nuestros instrumentos según su relevancia:

*una pregunta,  
una breve escala,  
un instrumento más cuidado, etc.*

**Cuando construimos o elegimos un instrumento para medir un rasgo, no debemos pensar sólo en ese rasgo, sino en toda una teoría relacional centrada en ese rasgo.**

Figura 2

<sup>2</sup> Ejemplos de diversos tipos de preguntas pueden verse en el documento de Morales (2009) *Guía para construir escalas de actitudes*

Sin entrar aquí en cómo hacer un cuestionario, modalidades de preguntas, etc., es importante caer en la cuenta en la necesidad de *detectar diferencias* en las variables que medimos, tal como se indica en la figura 3.

**Necesidad de *detectar diferencias*...**

- *Medir es diferenciar...*
- *Sin diferencias en los sujetos no podemos detectar relaciones entre variables...*
- *Sin dar a los sujetos la oportunidad de matizar sus respuestas es más difícil comprobar cambios o diferencias entre grupos...*
- *La heterogeneidad de la muestra no es la real y objetiva, sino la que somos capaces de detectar con nuestro instrumento...*
- *El modo de recoger los datos (pocas respuestas, sí o no) puede hacer muy semejantes a los que realmente no lo son tanto...*

**Estrategias**

1. No dicotomizar la información que por naturaleza no es dicotómica...
2. Disponer de *varios indicadores* (que incluso se pueden *sumar*) en vez de una única pregunta...
3. Aumentar el número de respuestas...

Figura 3

### 3.2. Implicaciones en la elección de la muestra

El listado de variables de interés también nos puede sugerir cómo elegir y organizar la muestra que va a ser objeto de nuestro estudio.

No entramos aquí en el tema de los *tipos de muestras*, cómo seleccionarla, etc., pero conviene pensar en algunas posibilidades<sup>3</sup>.

- La utilidad del *muestreo estratificado*; si hay varios *tipos de sujetos* (denominados *estratos*) en la población (*sexos, profesiones, edades, etc.*), estos deben estar representados en la muestra proporcionalmente a su número en la población.

Si en cada estrato o clasificación los sujetos son escogidos *aleatoriamente* es cuando se denomina con propiedad *muestreo estratificado*; si de cada clasificación se escoge la proporción adecuada de sujetos pero no aleatoriamente se denomina *muestreo por cuotas*, que suele ser una buena alternativa (más cómoda que la preferible, si es posible, elección aleatoria).

- Podemos *no incluir* en la muestra un determinado tipo de sujetos para eliminar una variable (*sólo niñas*), o estudiarlos por separado o en otra ocasión (el investigador se marca sus propios límites).

<sup>3</sup> Conviene tener una idea clara sobre los diversos *tipos de muestras*; esta información se encuentra con facilidad en textos sobre investigación y también en fuentes de Internet. Una dirección cómoda es StatPac Inc (2003) *Questionnaires & Survey Design* (ver referencias bibliográficas); en *Sampling Methods* (al final de la columna central, también se puede ir directamente a <http://www.statpac.com/surveys/sampling.htm>) hay una clara y breve revisión de los distintos *tipos de muestras* (útil para explicar lo que se ha hecho); una exposición más elaborada en Trochim (2006), *Research Methods Knowledge Base* (dirección electrónica en las referencias bibliográficas) en el menú de la izquierda *Sampling* y en el menú de *samplig* los tipos de muestras en *Probability Sampling* y *Nonprobability Sampling*).

- Podemos buscar *intencionadamente* determinados tipos de personas que responden a características (variables) que nos interesa incluir en la investigación; podemos circunscribir la investigación a un determinado tipo de población.
- La utilidad del *muestreo aleatorio* para neutralizar el influjo de *variables confundidas y desconocidas*; es éste un requisito de los diseños experimentales *en sentido propio*, que con frecuencia no se cumple aunque no siempre es tan complicado en la práctica, sobre todo cuando se investiga sobre poblaciones de tamaño limitado (como los alumnos de una facultad).

El tipo de muestreo (aleatorio simple, aleatorio estratificado, por cuotas, etc.) no tiene que ver solamente con las variables, sino con la posibilidad de *extrapolar los resultados a la población* representada en la muestra o muestras de nuestro estudio.

### 3.3. Implicaciones en los métodos de análisis estadístico

Este apartado puede coincidir de hecho con lo dicho en el anterior, pero desde otra perspectiva que puede ayudar a planificar toda la investigación.

Hay modalidades en los métodos de análisis que permiten controlar variables y en los que con frecuencia no se piensa. En realidad investigar es *responder a preguntas*, y los métodos de análisis estadístico (junto con los *diseños*) no son otra cosa que un *repertorio de respuestas*; hay preguntas, y posibilidades, que no nos hacemos (ni se nos ocurren) porque a veces tenemos un repertorio de respuestas muy limitado (como calcular una correlación, comparar dos grupos) y solamente nos preguntamos lo que sabemos responder.

Sin entrar a fondo en este tema (métodos de análisis estadístico), podemos pensar por ejemplo que:

Con las *correlaciones parciales* (muy sencillas) podemos controlar una o más variables (mediante correlaciones parciales podemos comprobar la relación entre dos variables igualando a los sujetos en una tercera variable)<sup>4</sup>.

Mediante *correlaciones múltiples (ecuaciones de regresión)* podemos verificar el impacto relativo de una serie de variables independientes en un determinado efecto o variable dependiente.

Hay modalidades de *análisis de varianza* especialmente diseñadas para controlar variables, o para verificar *tendencias*, etc.

Al comenzar una investigación de carácter *experimental* (o simplemente *empírico, cuantitativo*) es de suma importancia tener a la vista (o repasar) una serie de métodos de análisis estadísticos. Este repaso puede influir en las hipótesis que nos planteamos, en el mismo cuestionario para recoger datos, etc. Por ejemplo, si sabemos que podemos verificar *tendencias* (a crecer o decrecer) podemos obtener los datos apropiados (curso, edad, etapa, etc.).

En cualquier caso el disponer de una serie de métodos de análisis (*a qué preguntas responden, cómo se interpretan*) puede por supuesto enriquecer una investigación que de otra manera puede resultar excesivamente simple.

Es frecuente acudir a un *consultor experto* para que nos diga qué análisis podemos hacer o para que haga él mismo los análisis que crea oportunos; en estos casos el experto puede decidir qué análisis se pueden hacer *con estos datos y con esta muestra*, y como hay análisis que

---

<sup>4</sup> Sobre la correlación parcial puede verse Morales (2008), cap. 5, p. 168

responden a determinadas hipótesis, el planteamiento de la investigación puede cambiar bastante o quedar muy enriquecido añadiendo algunos análisis de interés en los que no se había pensado previamente.

#### 4. Variables en cuanto tipos de datos o escalas

El término variable es de hecho equívoco pues se utiliza con significados distintos, aunque el contexto suele dejar claro en qué sentido se emplea el término variable. Aun así y tratando de las variables en el ámbito de la investigación experimental, conviene decir algo brevemente sobre otro significado del término variable en este mismo contexto de la investigación. Hablamos también de tipos de variables (o de *tipos de escalas*) en otro sentido.

Las variables, en cuanto rasgos o en cuanto características podemos analizarlas porque sobre las mismas recogemos datos con determinados *instrumentos*: cuestionarios, tests, guías para codificar la observación, etc. Estos instrumentos (o *escalas* en un sentido genérico) nos dan datos que pueden ser de varios tipos:

**Variables nominales** o categóricas: los datos son *categorías de clasificación*, como el sexo, lugar de procedencia, etc.; no hay números que en sentido propio indiquen *orden* o *cantidad*.

Es normal codificar con números las respuestas a estas preguntas, pero estos números no lo son en sentido propio cuando no tiene sentido hablar de *más* o *menos*.

Cuando las categorías de clasificación son *dos mutuamente excluyentes* (como responder *sí* o *no* a una pregunta) se denominan **dicotómicas** (y suelen codificarse con 1 ó 0). Pueden tener un significado más *cuantitativo*, o numérico en un sentido más propio, en la medida en que signifiquen *presencia* o *carencia* de *algo* (por ejemplo los *unos* y *ceros* en las preguntas de los tests objetivos).

**Variables continuas** son aquellas que pueden adquirir valores numéricos consecutivos de menos a más (como la edad, rendimiento académico, grado de acuerdo con una afirmación, etc.) y pueden ser de tres tipos:

**1. Ordinales:** sólo conocemos, o sólo utilizamos, el *número de orden*, prescindiendo de la *distancia*. Por ejemplo si tres sujetos puntúan en un test 38, 27 y 12, podemos *ponerlos por orden* y los números anteriores se convierten en 1, 2 y 3; tenemos en este caso una escala o variable *ordinal*. También puede suceder que el *número de orden* (que denominamos *rango*) sea el único dato disponible.

**2. De intervalo** cuando podemos asumir que existe una *unidad* en sentido propio y podemos *calcular distancias* (como cuando medimos longitudes). Esto es más problemático que suceda con los datos que obtenemos con los instrumentos de medición psicológica y educacional, aunque de hecho los utilizamos como si se tratara de escalas de intervalo, y tenemos buenas razones para hacerlo así.

**3. De razón**, que son escalas de intervalo en las que además hay un origen o un *punto cero real*; estas escalas en sentido propio no son propias de nuestras ciencias sociales; por ejemplo el puntuar cero en un test de inteligencia no supone tener *cero inteligencia*, el punto cero es arbitrario.

Las variables continuas podemos también convertirlas en variables **dicotomizadas** (como *por encima de la mediana* = 1 y *por debajo de la mediana* = 0). Al dicotomizar una variable perdemos información, aun así los datos pueden ser *más seguros* que sin dicotomizar, o los únicos disponibles.

Estos tipos de escalas tienen implicaciones importantes en la teoría de la medición, porque el tipo de dato o escala condiciona los tipos de análisis matemáticos que pueden hacerse. En rigor las operaciones aritméticas que normalmente hacemos (como calcular medias) sólo son posibles con datos de *intervalo* o de *razón*. Hay sin embargo razones teóricas y estudios experimentales que justifican suficientemente el tratar muchos de los datos de las ciencias sociales (psicología, educación) como si fueran realmente datos o escalas de *intervalo*, que es lo que de hecho se hace habitualmente (Una exposición y discusión más amplia sobre los *tipos de escalas y su relación con las operaciones permisibles* puede verse en el capítulo I de Morales, 2006).

## 5. Referencias bibliográficas

MORALES VALLEJO, PEDRO (2006). *Medición de actitudes en psicología y educación*, 3ª edición. Madrid: Universidad Pontificia Comillas.

MORALES VALLEJO, PEDRO (2008). *Estadística aplicada a las Ciencias Sociales*. Madrid: Universidad Pontificia Comillas.

MORALES VALLEJO, PEDRO (2009). *El control de variables: control estadístico (análisis de covarianza) y control experimental mediante diseño*  
<http://www.upcomillas.es/personal/peter/investigacion/Controldevariables.pdf> (última revisión, 2 de Abril de 2009)

MORALES VALLEJO, PEDRO (2009). *Guía para construir escalas de actitudes*  
<http://www.upcomillas.es/personal/peter/otrosdocumentos/Guiaparaconstruirescalasdeactitudes.pdf>

STATPAC INC (2003) *Questionnaires & Survey Design*  
<http://www.statpac.com/surveys/index.htm#toc>;

TROCHIM, WILLAM M. (Cornell University) (2006), *Research Methods Knowledge Base*  
<http://www.socialresearchmethods.net/kb/index.php>