

SERIE DE MANUALES DIDACTICOS

Nº 6

## ESTADISTICA DESCRIPTIVA EN SALUD ANIMAL

VOLUMEN I



organización panamericana de la salud  
oficina sanitaria panamericana, oficina regional  
de la organización mundial de la salud



ORGANIZACION PANAMERICANA DE LA SALUD

*Oficina Sanitaria Panamericana, Oficina Regional de la*

ORGANIZACION MUNDIAL DE LA SALUD

**CENTRO PANAMERICANO DE FIEBRE AFTOSA**

CAIXA POSTAL 589 - ZC/00 - RIO DE JANEIRO, BRASIL

ESTADISTICA DESCRIPTIVA EN SALUD ANIMAL

por

Vicente M. Astudillo

Hernán Málaga C.

Melba Wanderley

VOLUMEN I

1 9 7 6

VOLUMEN I

METODO ESTADISTICO . . . . .	1
RECOLECCION DE DATOS . . . . .	15
ELABORACION DE DATOS: ORDENACION . . . . .	25
ELABORACION DE DATOS: TABULACION . . . . .	47
ELABORACION DE DATOS: REPRESENTACIO GRAFICA . . . . .	57

VOLUMEN II

RESUMEN DE ATRIBUTOS . . . . .	1
TASAS DE MORBILIDAD . . . . .	9
RESUMEN DE DATOS CUANTITATIVOS: MEDIDAS DE POSICION Y DISPERSION . . . . .	13
REGRESION LINEAL . . . . .	27
CORRELACION RECTILINEA . . . . .	37
CORRELACION MULTIPLE Y PARCIAL. COEFICIENTE DE CORRELACION . . . . .	45
SERIES CRONOLOGICAS . . . . .	49
NUMEROS INDICES . . . . .	57
ESTADISTICAS DEMOGRAFICAS . . . . .	63

VOLUMEN I

C O N T E N I D O

	pág.
METODO ESTADISTICO . . . . .	1
1. Hechos de los que se ocupa la estadística . . .	1
2. Aceptaciones de estadística . . . . .	3
3. El conocimiento científico de los problemas de salud animal y el método estadístico . . . .	3
4. Planificación de un estudio . . . . .	7
5. Tipos de estudio . . . . .	12
RECOLECCION DE DATOS . . . . .	15
1. Para qué se recogen datos estadísticos . . . .	15
2. Qué datos recoger . . . . .	15
3. Cómo recolectar los datos . . . . .	16
4. Métodos de recolección de datos a partir de las fuentes primarias . . . . .	18
5. Elaboración de un cuestionario o formulario . .	21
ELABORACION DE DATOS: ORDENACION . . . . .	25
1. Generalidades . . . . .	25
2. Crítica de los datos recolectados . . . . .	26
3. Ordenamiento: clasificación y recuento . . . .	27
ELABORACION DE DATOS: TABULACION . . . . .	47
ELABORACION DE DATOS: REPRESENTACION GRAFICA . . . . .	57

\* \* \*

## METODO ESTADISTICO

### 1. HECHO DE LOS QUE SE OCUPA LA ESTADISTICA

1.1 Existen muchos fenómenos (hechos) naturales y sociales que son determinados por una multiplicidad de causas, las cuales además se interrelacionan entre sí y dependen a la vez de múltiples otras causas. Muchas de ellas son desconocidas y en conjunto contribuyen en forma significativa para producir el efecto estudiado. Esto es lo que hace que un fenómeno sea de naturaleza compleja. Un sector importante de esas causas desconocidas es de microcausas; en el sentido de que su contribución individual es pequeña, lo que las hace incontrolables desde el punto de vista práctico. De ahí que al repetir la ocurrencia del fenómeno, aun cuando sea en condiciones de la mayor uniformidad, se produzcan resultados o respuestas diferentes. Esto es lo que hace que la ocurrencia de un fenómeno natural o social sea variable. La estadística permite tratar este tipo de fenómenos de una manera más elaborada que la forma vulgar.

#### 1.2 La estadística se aplica :

Cuando se hacen observaciones repetidas en condiciones similares.

Cuando a pesar de tener las mayores precauciones en uniformar las condiciones en que se hacen las observaciones, los resultados que ellas presentan son diferentes, son variables.

Cuando el valor de la expresión cuantitativa, de cualquier observación individual, es incierto, no predictible, aún cuando se conozcan algunos factores que pueden ser determinantes de carácter en estudio.

1.3 Es un hecho tratable por la estadística el nivel de anticuerpos circulantes contra la fiebre aftosa en la población bovina. No es un problema trata -

ble por la estadística el valor de la hipotenusa de un triángulo. En el problema de los anticuerpos, aún cuando sepamos cuales la raza, sexo, edad, peso, condiciones sanitarias y el período de tiempo postvacunal, etc., no se puede predecir el valor del Índice de seroprotección, al medir el nivel de anticuerpos con este indicador. Si un cierto bovino, especificadas las condiciones ya enumeradas, tiene un valor de seroprotección de 2,3, otro bovino en las mismas condiciones puede tener un valor 0,5. En cambio en el caso del cuadrado de la hipotenusa, si se dan los valores de los dos catetos (condiciones) se puede predecir en forma cierta el valor de la hipotenusa. Además, si dos triángulos tienen los mismos valores para los catetos (las mismas condiciones) el valor de la hipotenusa será el mismo, para los dos triángulos. En este último caso se puede decir que conociendo los factores determinantes, el resultado es predecible.

1.4 El hecho que las repeticiones de un carácter manifiesten variación (diferencias en el resultado) se debe a que los hechos naturales y sociales, como ya se ha dicho están influidos por una gama de causas (multicausalidad) que se interrelacionan entre sí y dependen a su vez de múltiples otras causas. La relación de factores causales es siempre incompleta, siendo la mayor parte de las causas desconocidas, razón por la cual se dice que la variación es aleatoria o causal. Se entiende por azar (aleatorio) una mezcla entre lo desconocido, lo parcialmente conocido y lo conocido.

1.5 La variación de origen aleatoria, que hace que un valor individual cualquiera sea irregular en su comportamiento sea imposible de predecir en forma lógica, tiene un orden ou regularidad cuando es observada en un colectivo, o sea, un conjunto de repeticiones individuales. En el hecho colectivo se presenta una regularidad que disminuye la incertidumbre. Los valores que puede asumir la característica se hacen más o menos probables, se toman predecibles. Los resultados manifiestan ciertas tendencias, ya que algunos valores son más frecuentes que otros.

1.6 Los métodos estadísticos permiten estudiar conjunto de observaciones y dilucida esas tendencias generales, cuando las observaciones están in fluidas por múltiples causas de variación. Permiten captar el orden, la uniformidad que manifiestan los hechos naturales o sociales.

## 2. ACEPCIONES DE ESTADISTICA

2.1 Estadísticas. Datos cuantitativos afectados por una multitud de causas.

2.2 Estadística o Método Estadístico. Es un conjunto de procedimientos , especialmente adecuados, por una parte, para coleccionar, describir e interpretar datos cuantitativos afectados por múltiples causas de variación , por otra parte, para que basados en un conjunto pequeño de datos, se tomen decisiones sobre el comportamiento del fenómeno en el conjunto mayor de datos.

2.3 Estadística Descriptiva. Conjunto de procedimientos que permiten la ordenación, presentación y resumen de grandes volúmenes de datos y formas manejables e inteligibles que permitan caracterizarlos.

2.4 Estadística Inductiva. Conjunto de métodos que permiten hacer generalizaciones sobre propiedades de un carácter en el agregado completo, a partir de un pequeño grupo de unidades o individuos.

## 3. EL CONOCIMIENTO CIENTIFICO DE LOS PROBLEMAS DE SALUD ANIMAL Y EL METODO ESTADISTICO

3.1 Los hechos en salud animal dependen de múltiples causas y son complejos, se presentan muy VARIABLES de un individuo a otro, de un rebaño a otro rebaño, de una región a otra, de un momento a otro.

3.2 El médico veterinario que trabaja en salud animal requiere hacer estos hechos inteligibles. A través del conocimiento de las relaciones causales podrá explicar racionalmente su ocurrencia, haciéndolos manejables, pudiendo de esta manera influir sobre ellos. Es decir, conociendo los factores que inciden sobre una enfermedad puede atacarlos, para reducir o eliminar la enfermedad.

3.3 Para llegar a establecer relaciones permanentes y esenciales entre los hechos y así poder manejarlos y modificarlos debe aplicar un procedimiento sistemático y lógico que cubra el ciclo completo desde el enunciado del problema hasta la evaluación de los resultados.

3.4 Para detectar una relación entre dos hechos, como camino para llegar a descubrir una "relación causal" particular entre ellos, se sigue un ciclo lógico de procedimientos que se puede descomponer, con fines pedagógicos, en las etapas siguientes:

i) Enunciado del problema a estudiar

Primero que nada, siguiendo la metodología de solución de problemas, se plantea en forma explícita el problema. Para plantear las alternativas de solución y sus correspondientes modelos es necesario tener algún conocimiento sobre las " relaciones causales " entre los componentes. Para esto será necesario averiguar la existencia de algunas de ellas para alimentar el modelo de solución.

Por ejemplo estamos frente al problema:

-inmunización deficiente de los bovinos contra la fiebre aftosa.

ii) Planteamiento de una hipótesis

Considerando el ejemplo dado, una hipótesis o explicación provisoria podría ser: un factor responsable por la inmunización deficiente de la población bovina es la mala calidad de las vacunas.

iii) Deducción de consecuencias verificables de la hipótesis

A veces no es posible verificar directamente la veracidad de la hipótesis, razón por la cual, en forma lógica se deducen consecuencias de la



hipótesis, que se pueden verificar. Para el ejemplo dado : En los rebaños bovinos, que utilizan vacunas de mala calidad, habrá una mayor ocurrencia de fiebre aftosa que en rebaños bovinos que utilizan vacunas de buena calidad, en el supuesto que los dos conjuntos de rebaños son idénticos en todas las otras características.

iv) Verificación de la hipótesis

Se hace a través de la observación del hecho. Implica la aplicación de procedimientos estadísticos. Por ejemplo, toman rebaños uniformes y se asignan al azar al tipo de vacuna a utilizar y se recogen datos sobre ocurrencia de fiebre aftosa, estableciendo la relación entre ambos hechos. Se debe tener en cuenta que una asociación estadística es necesaria pero no suficiente para establecer una "relación causal".

v) Interpretación de resultados

Si los hechos no están de acuerdo con las consecuencias deducidas de la hipótesis, entonces se rechaza la hipótesis. El uso de vacunas de mala calidad no es el único factor. Podrían haber otros factores que pudieran enmascarar el efecto de las vacunas de mala calidad. Por ejemplo, si en los rebaños que utilizan vacunas de mala calidad, mantienen normas sanitarias rigurosas y si la no aplicación de medidas sanitarias fuera otro factor influyente, entonces el efecto real de las vacunas de mala calidad, quedaría enmascarado por el efecto contrario de la aplicación rigurosa de medidas sanitarias. Si los hechos están de acuerdo con consecuencias verificables deducidas de la hipótesis, entonces se acepta la hipótesis planteada.

Si el factor más importante en la pobre inmunización de la población bovina fuera la no aplicación de medidas sanitarias y en rebaños que utilizan vacunas de mala calidad, no se aplicara ningún tipo de medidas sanitarias, la mayor ocurrencia de fiebre aftosa en esos rebaños podría interpretarse como efecto exclusivo del uso de vacunas de mala calidad por ser ése el carácter considerado según la hipótesis, lo que no sería real.

3.5 Esta forma de adquirir conocimiento científico en salud animal es posible aplicando el método estadístico. Permite verificar la hipótesis a través de la colecta y análisis de los datos y extraer conclusiones válidas. En el campo de la salud animal no se pueden verificar hipótesis con certeza como ocurre en el campo de las ciencias exactas. Una hipótesis solo es probablemente verdadera. Dado que no es posible considerar todos los factores que influyen, ni tampoco todos los individuos, por riguroso que sea el estudio, se produce variación en los resultados. La conclusión tiene cierto grado de incertidumbre. Las relaciones que se presentan en las ciencias naturales y sociales son aproximadas, tienen por fundamento la frecuencia por lo cual se tornan más o menos probables. En cambio en las ciencias exactas las relaciones entre una causa y el efecto tienen el carácter de "necesario y suficiente". La diferencia de como es una relación en estos dos campos se puede esquematizar de la manera siguiente :

- Si existiese una relación de "causa" (cuerpo lúteo retenido) y "efecto" (metritis), se podrían presentar los resultados siguientes :

i) Modelo en las Ciencias Exactas

METRITIS	CUERPO LÚTEO RETENIDO		TOTAL
	SI	NO	
SI	200	0	200
NO	0	200	200
T O T A L	200	200	400

ii) Modelo en las Ciencias Naturales

METRITIS	CUERPO LÚTEO RETENIDO		TOTAL
	SI	NO	
SI	183	24	207
NO	17	176	193
T O T A L	200	200	400

#### 4. PLANIFICACION DE UN ESTUDIO

4.1 El método estadístico proporciona procedimientos para llevar a la práctica las etapas del método científico que requieren recolección, análisis e interpretación de datos.

La planificación de un estudio permite determinar previamente que es lo que se requiere conocer y la secuencia de pasos tendientes a:

- . obtener datos pertinentes y apropiados,
- .. que permitan un análisis objetivo,
- ... que conduzca a conclusiones válidas con respecto al problema en estudio.

Dentro de la fase de planificación se pueden considerar los siguientes aspectos :

#### 4.2 Definición de los objetivos

Objetivo es lo que se pretende alcanzar o a lo que se quiere llegar (PARA QUE). Si no se tiene claro lo que se quiere hacer o lo que se desea alcanzar no se puede actuar científicamente. Si este paso falta, por creerlo subentendido, está expuesto a pérdidas de tiempo, esfuerzos y recursos. Una buena definición de los objetivos orienta la planificación de la colecta de información adecuada y de los procedimientos de análisis de los datos.

Formalmente la definición del objetivo puede corresponder a una descripción o a una explicación. En el primer caso se enuncia el problema a estudiar, precisando su alcance y sus límites.

En el segundo caso se trata de la deducción de las consecuencias prácticas de la hipótesis, si ella fuera cierta.

#### 4.3 Definición del Universo o Población

Universo, es el conjunto de elementos, individuos o casos, que manifiestan las características o hecho que nos interesa observar, medir y conocer y al cual se van a aplicar las conclusiones del estudio.

Es muy importante delimitar en forma clara el conjunto de individuos a estudiar para no gastar recursos e incluir aquellos que no interesan. A sí por ejemplo, al describir la producción de leche de las vacas de una región, será importante definir qué se entiende por vacas, por producción de leche (diaria, lactancia).

#### 4.4 Determinación de la Muestra

Puede ser impracticable o innecesario el estudio del universo completo.

Muestra es una parte del universo, seleccionada de acuerdo con ciertos principios probabilísticos o casuales, lo que permite que los resultados en ella obtenidos se puedan generalizar al universo de origen.

#### 4.5 Definición del Grupo testigo o Control

Cuando el estudio es de carácter explicativo, es decir, existe una hipótesis que verificar, entonces se hace necesario considerar grupos de individuos que tengan y no tengan la característica estudiada. Se llamará grupo de estudio a aquel que tiene la característica y grupo testigo el que no la posee.

Cuando se trata de medir el efecto de una vacuna antiaftosa en un grupo de bovinos, es indispensable que exista otro grupo de bovinos cuya única diferencia con el que se le aplica la vacuna sea la ausencia de ella.

Si se aplica un tratamiento que resulta favorable en el 90% de los casos, no debemos impresionarnos con el resultado, ya que puede ser que en un grupo sin el tratamiento, los resultados favorables pudieran ser iguales o incluso superiores.

##### Requisitos de un Grupo Testigo

i) Provenir del mismo universo que el grupo en estudio. Ambos grupos deben tener composición semejante en cuanto a otras características que pudieran influir en el resultado.

ii) Diferenciarse del grupo en estudio sólo por la ausencia de la característica en estudio.

#### 4.6 Determinación de las Unidades de Observación

Definir en qué elementos se observará (medir o contar) la característica en estudio.

Puede ser que si hay varios objetivos, cada uno requiera de diferentes unidades de observación. Puede haber aspectos que se refieran a rebaños, otros a vacas, otros a cada bovino.

<u>PROBLEMA</u>	<u>UNIVERSO</u>	<u>UNIDAD OBSERVACION</u>
Hemograma	Sangre de un individuo	Un centímetro cúbico
Tamaño camada cerdo al nacer, región A	Todos los tamaños de las camadas de cerdos al nacer en región A	Cada camada al nacer en región A.
Movimiento de animales	Archivo de fichas de una Inspectoría Veterinaria.	Cada ficha por rebaño.

#### 4.7 Determinación de la información necesaria para el estudio

Se debe hacer una cuidadosa selección de la información a recolectar. Debe ser pertinente, indispensable y suficiente para cumplir con los objetivos establecidos en el estudio.

##### Requisitos de la información

Sólo deben recolectarse datos :

i) útiles a los objetivos del estudio, que sean indispensables para ellos y no los que " podrían servir en alguna oportunidad ". Se trata de reunir los datos estrictamente necesarios.

ii) que sean fáciles de obtener, ya que existen datos importantes pero que son difíciles de obtener o sólo se obtienen en algunas unidades.

iii) en lo posible sin errores, o con un error mínimo, ya que los errores de información conducen a errores en las conclusiones.

#### 4.8 Fuentes de origen de la información

En general se presentan dos alternativas para obtener los datos necesarios : utilizar datos ya recogidos y producir los datos.

##### i) datos ya recogidos

se refieren a hechos ocurridos con anterioridad. Estos datos pueden estar dispuestos en dos formas :

- . publicaciones, de organismos responsables
- .. no elaborados ni publicados, como ser en fichas, registros, tarjetas, etc.

##### ii) datos a producir

en este caso hay que recolectar los datos a través de algunos procedimientos que más adelante serán expuestos.

##### Características de los datos a producir

Ventajas : se recogen sólo los útiles, se conoce la calidad de los datos.

Desventajas : mayor gasto de recursos, tiempo y trabajo.

#### 4.9 Fijación de Unidades de Medida y Escalas de Clasificación

Tiene importancia tanto para la recolección de los datos como para la elaboración de los mismos. La unidad de medida se refiere al padrón utilizado para conocer la intensidad de un hecho. Es fundamental cuantificar, midiendo la magnitud o enumerando las modalidades.

Dato : información que proporcione la unidad de observación

Tipos de datos :

datos cuantitativos : son susceptibles de expresión numérica ya que es posible medirlos.

datos cualitativos : no son susceptibles de expresión numérica, ya que revelan un atributo o cualidad. Sólo se enumeran.

continuos : toman cualquier valor entre las unidades. Peso, estatura.

discretos : sólo toman valores enteros. Nº de bovinos por rebaño.

-Escalas de Clasificación

Datos de enumeración ————— Atributos o variables cuali  
tativas

Datos de medición ————— Variables cuantitativas

. Variable : símbolo numérico o característica susceptible de tomar diversos valores.

. Constante : cuando el símbolo numérico sólo toma un valor.

Las variables pueden ser continuas o discretas según el tipo de datos.

-Escalas de Medición

. Escala nominal o clasificatoria

Es la forma más primitiva. En ella sólo se permite contar.

Algunos autores no la consideran una escala de medición. Se aplican símbolos o números para clasificar individuos o elementos.

Ejemplos :

- i) las razas de los animales
- ii) los números de las marcas de los animales

En estos casos la asignación de un número es arbitraria y tiene el propósito de clasificar.

.. Escala ordinal

Cuando entre varios individuos es posible ordenarlos en forma jerárquica de acuerdo con la relación entre ellos de alguna característica. Unos son más viejos que otros, unos más gordos que otros, etc.

... Escala racional

Es una escala realmente cuantitativa. La razón entre dos puntos cualesquiera de la escala es independiente de la unidad de medida. Por ejemplo, al pesar los bovinos en kg y libras, al determinar los pesos de los bovinos en ambas medidas, encontramos que la relación de los dos pesos en libras o kg es igual.

#### 4.10 Plan de elaboración de los datos

Establecer el mecanismo de clasificación y recuentos.

Determinar un programa de tabulaciones y representación gráfica.

Indicar qué valores característicos se utilizarán para resumir los da  
tos.

Cuando esas medidas de resumen se refieren al universo se denominan  
parámetros y cuando se refieren a muestras se denominan estadígrafos.

#### 4.11 Organización del estudio

a) duración del estudio

b) utilización de equipos, como ser computación

c) determinar los recursos humanos materiales y financieros

d) determinación del costo del estudio.

### 5. TIPOS DE ESTUDIO

5.1 Un mismo problema puede estudiarse de distintas maneras. El tipo de estudio que se realice dependerá entre otras cosas de : los objetivos del estudio, la existencia de hipótesis, la fuente de origen de los datos, el orden en que se recoge la información y el manejo de las variables.

A continuación se muestran algunas clasificaciones de investigaciones.

#### 5.2 Retrospectiva y Prospectiva

Con respecto a la relación entre tiempo de ocurrencia y registro de la información se define al estudio retrospectivo como aquél que averigua hechos ocurridos en el pasado; el estudio prospectivo, en cambio, va registrando la información a medida que se va produciendo.

Con respecto a la relación entre causa y efecto, en el estudio retrospectivo, las unidades de observación se clasifican según la variable " efec



to " y luego se averigua la existencia o intensidad de la variable presumi  
blemente causal en las diferentes clases. En el estudio prospectivo la pri  
mera clasificación se hace según la variable que se supone causal y luego  
se va registrando la ocurrencia o no, del efecto.

Ejemplo: Frente a la hipótesis de que el consumo de cigarrillos  
favorece la aparición de cáncer pulmonar el estudio retrospectivo clasifi-  
caría a los individuos en cancerosos y sanos y averiguaría si existe dife-  
rencia en el hábito de fumar entre los dos grupos. En un estudio prospec-  
tivo se observaría si en un grupo de fumadores aparecen más enfermos o  
muertos de cáncer pulmonar que en un grupo de no fumadores.

### 5.3 Transversal y Longitudinal

El estudio transversal estudia las diversas variables en determinado  
momento, el longitudinal estudia las variables a través del tiempo. Pueden  
compararse a una fotografía instantánea y a una película de los hechos res  
pectivamente.

### 5.4 Descriptiva y Explicativa

El estudio descriptivo tiene como objetivo mostrar una situación, el  
explicativo pretende averiguar la veracidad de una hipótesis.

Los resultados obtenidos en un estudio descriptivo pueden dar origen  
a una hipótesis y cumplir así con la primera etapa del método científico.

### 5.5 Experimental y no Experimental

El estudio experimental tiene como característica que el investiga -  
dor maneja la variable independientemente.

Las unidades experimentales son adjudicadas en forma aleatoria a las  
diferentes categorías del factor presumiblemente causal, estudiándose los  
efectos que se producen. En los estudios no experimentales sólo se puede  
clasificar el material de estudio en diferentes categorías de los factores  
causa y efecto, estudiando la relación entre ambos.

Las conclusiones de estudios experimentales bien diseñados son más vá lidas que las conclusiones que se obtengan de buenos estudios no experimentales. El experimento es el mejor camino para acercarse científicamente a la verdad en el establecimiento de relaciones causales conociendo la probabilidad del error que pueda cometerse.

Las distintas clasificaciones no son mutuamente excluyentes e incluso la pertenencia a una categoría puede determinar que un estudio debe ser necesariamente de determinado grupo en las otras clasificaciones. Es así, por ejemplo, que el experimento será prospectivo, longitudinal y explicativo, en cambio el estudio no experimental puede ser de cualquier tipo en las otras clasificaciones.

## RECOLECCION DE DATOS

### 1. Para qué se recogen datos estadísticos

1.1 La recolección de datos tiene por finalidad coleccionar información sobre las características en estudio, de manera de satisfacer los objetivos.

### 2. Qué datos recoger

2.1 Deben ser medibles, ya que un propósito de cualquier estudio científico es la medición objetiva de las características estudiadas.

2.2 Que sean pertinentes, se refieran a los objetivos del estudio.

2.3 Que sean indispensables para satisfacer los objetivos. No se deben recolectar datos que " casualmente podrían servir más adelante ", como acontece muchas veces.

2.4 Que sean factibles de obtener, ya que existen datos que a pesar de ser útiles son difíciles de coleccionar o sólo pueden ser obtenidos en algunas unidades de observación.

2.5 Que sean lo más exactos posibles, es decir, sin errores, ya que los errores de información conducen a conclusiones erróneas.

2.6 Deben ser definidos, ya que si su significado no ha sido aclarado, no hay uniformidad de criterio en quienes los recogen, resultan datos no comparables, invalidando las conclusiones.

### 3. Cómo recolectar los datos

- 3.1 En general los datos necesarios para satisfacer los objetivos de un estudio se pueden producir (fuente primaria) o pueden ser utilizados datos ya recopilados (fuente secundaria).
- 3.2 La utilización de una fuente secundaria significa hacer uso de datos recogidos en estudios realizados con anterioridad. Estas fuentes secundarias pueden ser de dos tipos : publicaciones y los datos recolectados pero no publicados.
- 3.3 Los datos publicados se obtienen generalmente a partir de publicaciones de organismos estadísticos oficiales, revistas científicas y técnicas.
- 3.4 Los datos ya recogidos pero no elaborados ni publicados se obtienen de variadas fuentes, entre las que se pueden mencionar archivos epidemiológicos, fichas clínicas, registros de cooperativas pecuarias, archivos de servicios veterinarios, registros de los establecimientos pecuarios.
- 3.5 La ventaja mayor que ofrece el uso de fuentes secundarias es el ahorro de tiempo, trabajo y recursos, ya que basta encontrar las fuentes que contienen los datos. El inconveniente principal que tiene este procedimiento es la dificultad de conocer la calidad de los datos a utilizar.
- 3.6 La utilización de fuentes primarias de datos (producirlos) ofrece la ventaja de recoger sólo aquellos datos necesarios, conociendo la calidad de ellos. La desventaja está en el gasto de recursos, tiempo y trabajo que involucra. Generalmente hay que elaborar cuestionarios especialmente adecuados.

- 3.7 Los procedimientos de recolección de datos a partir de fuentes primarias se pueden clasificar atendiendo a los criterios tiempo y cobertura. Según el tiempo, los procedimientos de recolección de datos se clasifican en continuos, periódicos y ocasionales. Según la cobertura, los procedimientos de recolección de datos se clasifican en completos y muestrales.
- 3.8 La recolección de datos continua en el tiempo se llama también registro. En esta forma, se anotan en forma permanente y obligatoria los hechos que se producen a través del tiempo. Ej.: registro civil, registro de la notificación por fiebre aftosa, registros de las cooperativas lecheras.
- Las características principales de este procedimiento son las siguientes : permanencia en el tiempo, ya que no se interrumpe; obligatoriedad para el problema a que se refiere.
- 3.9 La recolección de datos periódica en el tiempo permite la obtención de datos en lapsos de tiempos regulares. El censo pecuario es un procedimiento de este tipo, ya que se lleva a cabo cada 10 años.
- 3.10 La recolección ocasional en el tiempo, se refiere a procedimientos que permiten obtener datos en forma esporádica, con duración limitada en el tiempo. Las encuestas por muestreo y los experimentos son ejemplos de este tipo de procedimiento.
- 3.11 Los procedimientos de recolección de cobertura total son encuestas completas. Ellos son exhaustivos ya que cubren todas las unidades de observación de un universo. Un ejemplo de esta forma de procedimiento, cuando se trata de enumerar todos los individuos de una población, es el censo de población, de vivienda, pecuario, etc.
- 3.12 Los procedimientos de recolección de datos, cuya cobertura no es completa, se llaman encuestas por muestreo. Permiten obtener información

mación sobre alguna característica a partir de un conjunto de unidades de observación, consideradas representativas de todo el universo. El tema de muestreo es tratado en forma separada.

4. Métodos de recolección de datos a partir de las fuentes primarias

- 4.1 En este caso se trata de producir datos originales que interesan de acuerdo a los objetivos planteados en el estudio. Esta producción de información se hace a través de los siguientes métodos : observación, entrevista y cuestionario por correo.
- 4.2 La observación de los hechos es la apreciación directa y objetiva de las características estudiadas. Es el método clásico en muchos estudios científicos, aceptando sin embargo, que no todas las características pueden ser estudiadas a través del método de observación.
- 4.3 La observación de los hechos, de la cual resultan los datos que posteriormente serán alabados, está expuesta a errores de muy diverso origen. Es importante preocuparse de reducir al máximo estos errores, ya que en todo el proceso estadístico posterior ellos persistirán, invalidando las conclusiones del estudio. Se deben eliminar todos aquellos factores perturbadores que tiendan a hacer poco exacta la observación.
- 4.4 Si se estudia algún aspecto de la epidemiología descriptiva de la fiebre aftosa y la contabilización de la frecuencia de bovinos enfermos es defectuosa, entonces las conclusiones serán de poco valor. Este es un error del observador. El error del observador, a veces se incrementa cuando el registro de los datos, fruto de la observación, es descuidado y deficiente. De ahí que sea importante contar con instrucciones escritas acerca del contenido y forma del registro de datos observados, además de completarla con una crítica de

ellos, que asegure la eliminación o rectificación de omisiones, vi  
cios o la presencia de datos ilegibles.

4.5 En el estudio de ciertas características puede cobrar importancia  
otra fuente de error, la relativa a los instrumentos o métodos de  
medición. Puede ocurrir con algunas pruebas diagnósticas, con método  
para medir niveles de anticuerpos circulantes, etc. Estos erro  
res son llamados errores del instrumento de observación.

4.6 Otro método de recolección a partir de una fuente primaria es el de  
la entrevista. Es muy utilizado en el campo de las ciencias socia  
les y para los médicos veterinarios que trabajan en salud animal de  
be ser un método conocido ya que en varias oportunidades han de va  
lerse de él.

Cuando se habla de entrevista el término se restringe a su signifi-  
cado en la colecta de datos, es decir, la búsqueda de información .  
Para esto el entrevistador formula un conjunto de preguntas y ano  
ta las respuestas en un cuestionario uniforme, todo de acuerdo a un  
manual de instrucciones. Este método, como el anterior también pue  
de ser fuente de errores.

4.7 En cuanto a las características que ha de tener el entrevistador no  
es fácil enumerarlas. La importancia de cualquiera de ellas depende  
del tipo de encuesta. A veces puede ser importante que tenga un a  
decuado conocimiento del tema tratado. Otras veces, no es convenien  
te que tenga opiniones o prejuicios sobre el tema de la encuesta  
ya que pueden influenciar la respuesta. Sin embargo, existen algu  
nas características personales que deben ser tomadas en cuenta para  
seleccionar entrevistadores. Entre ellas : honestidad, interés, exac  
titud, nivel de preparación compatible.

4.8 Es conveniente que el entrevistador sea entrenado para el desarrollo  
de sus tareas de recolección de información. Uno de los elementos

importantes es el conocimiento del manual de instrucciones sobre el trabajo de terreno, indicando

- a) encuentro de los entrevistadores
- b) solución a problemas de " no respuesta "
- c) como hacer las preguntas
- d) como se registran las respuestas
- e) instrucciones sobre el cuestionario

En este último documento deberían indicarse las definiciones necesarias.

Otro aspecto importante en la preparación de un entrevistador son las sesiones de discusión con el encargado del estudio, el que debe orientar al entrevistador indicándole la finalidad de la encuesta y cual se considera que es el procedimiento correcto de entrevista. Un tercer aspecto del entrenamiento del entrevistador es una aplicación previa en terreno de la entrevista, así se prueban además los cuestionarios e instructivos.

- 4.9 El entrevistador una vez que ubica al entrevistado debe explicar a éste el motivo de la encuesta y la institución mandante, indicándole al entrevistado, cuando así sea necesario, el carácter confidencial de la información que él proporcione.

En seguida se deben comenzar a formular las preguntas en el orden indicado, siguiendo las instrucciones.

En las preguntas que pudieran referirse a opiniones o conocimientos (SI SE LLEGARAN A FORMULAR) no se debe sugerir la respuesta.

Evitar insertar palabras como " realmente " antes de " le gusta " , puesto que afecta la respuesta.

- 4.10 El registro de las respuestas debe ser escrupuloso. Pueden producirse errores de registro por cansancio del entrevistador, errores de uso de códigos complejos. En algunas encuestas se agrega un espacio para que el entrevistador manifieste alguna observación sobre la entrevista, especialmente sobre la validez de las respuestas.



4.11 También se debe tener en cuenta la supervisión y el control del trabajo de terreno. La supervisión debe estar a cargo de alguien que haya participado en todas las etapas, que pueda detectar deficiencias de trabajo e indicar su corrección. Debe ser el enlace entre el nivel central y el terreno.

Para mantener la buena calidad del trabajo de terreno, es conveniente someter a constante control las actividades desarrolladas en terreno.

4.12 El método de los cuestionarios por correo se menciona aunque tiene escaso uso en el campo de salud animal. En general es un método que presenta grandes limitaciones. Entre ellos se pueden mencionar :

a) sólo es utilizable si las preguntas fueran muy simples y acompañadas de instrucciones muy claras y completas.

b) generalmente hay gran cantidad de uso de respuestas.

La ventaja que ofrece este procedimiento es el bajo costo y la rapidez.

## 5. Elaboración de un cuestionario o formulario

5.1 Cuestionario o formulario es un documento impreso, especialmente confeccionado para obtener información consonante con algún objetivo bien definido. Debe contener las preguntas necesarias y los espacios correspondientes y suficientes para ser llenados con las respuestas.

5.2 Tanto para determinar el contenido como la forma de un formulario es indispensable tener bien claro los objetivos del cuestionario :

a) El de expresar los objetivos del estudio en preguntas específicas

b) El de motivar al entrevistado a proporcionar la información requerida.

5.3 Revisaremos algunos aspectos generales que se deben tomar en cuenta en la elaboración de un cuestionario. Primero que nada veremos los

aspectos propios del contenido. Se debe tener en cuenta que cuestionarios muy extensos no se llenan muchas veces con exactitud, ni a grado, influyendo ésto en la calidad de la información proporcionada. Por tanto, en su construcción deben evitarse las preguntas que no sean estrictamente necesarias para cumplir con los objetivos del estudio.

Otros de los aspectos que deben considerarse en la construcción de un cuestionario es el lenguaje. Los términos elegidos deberán estar dentro del rango del vocabulario del respondiente, evitándose el uso de modismos. Tampoco es recomendable usar términos técnicos, excepto en aquellos estudios en que los entrevistados son técnicos.

El cuestionario debe considerar además del lenguaje, la forma de las preguntas. Estas deben tener en cuenta el marco de referencia del entrevistado. Un ejemplo que muestra el efecto del marco de referencia se puede apreciar al hacer la siguiente pregunta :

Como están los animales ? Puede responder bien, mal o regular, pero no se sabe si se refiere al estado sanitario, al manejo, a la producción, etc.

La forma de una pregunta debe estar de acuerdo con el nivel de conocimiento del entrevistado. Toda pregunta debe hacerse bajo el supuesto de que el entrevistado posee una respuesta adecuada.

Las preguntas deben formularse de modo que no sugieran la respuesta . Por ejemplo, una pregunta que pretende conocer la actitud acerca de la eliminación de vacas con brucelosis, podría formularse como : que piensa sobre la eliminación de vacas con brucelosis ? La misma pregunta en forma totalmente sesgada podría formularse como sigue : no dirá usted que está en contra de la eliminación de vacas con brucelosis, no es así ?

Las preguntas deben limitarse a una sola idea o referencia.

En preguntas como : " Que opina usted de los programas de control de la fiebre aftosa y brucelosis ", las respuestas no permitirán darse cuenta a cual idea se refiere o si es a ambas. La forma aceptable para esta pregunta dependerá de los objetivos específicos. En algunos

casos será necesario formular dos preguntas separadas, una para cada idea, en otros casos, si sólo se desea tener una idea de lo que la gente piensa en cuanto al problema en conjunto, podría ser recomendable la forma señalada anteriormente, teniéndose presente al analizar la respuesta que ella se refiere a ambas ideas y no a una en particular.

Fuera de la forma de cada pregunta se debe dar importancia al orden de las preguntas en el cuestionario. Las preguntas deben ordenarse de tal modo que aparezcan con sentido ante el respondiente. Un cuestionario bien diseñado facilita el paso de una pregunta a otra y a menudo lo lleva a anticiparse a la discusión.

La secuencia debe determinarse también por lo que se denomina un " acercamiento tipo embudo ". Esto se refiere a la formulación de las preguntas más generales primero, seguidas por aquellas más específicas.

Según la forma de la respuesta, se pueden considerar dos tipos de preguntas : abiertas y cerradas.

En las preguntas abiertas el respondiente debe contestar en sus propios términos, estructurando la respuesta a su gusto. Este tipo de preguntas permite conocer el marco de referencia del respondiente, lograr información que no puede obtenerse a través de las preguntas cerradas.

A veces se utilizan preguntas abiertas en encuestas piloto para conocer posibles respuestas a ellas y luego poder confeccionar preguntas de tipo cerrado.

En las preguntas cerradas, el entrevistado debe elegir de un grupo de respuestas dadas aquella que esté más de acuerdo con su posición. En general las preguntas cerradas se adaptan bien a situaciones en que:

- a) hay sólo un marco de referencia desde el cual el respondiente puede contestar.
- b) dentro del marco de referencia hay un rango conocido de respuestas posibles.

## PRUEBA DE CUESTIONARIO

La prueba del cuestionario es una etapa fundamental. Aunque el diseño del cuestionario se haya hecho de acuerdo a todas las reglas establecidas, puede que en la práctica no funcione y deban hacerse cambios antes de aplicarlo en forma definitiva.

La prueba proporciona los medios para detectar problemas no previstos, tales como : redacción, longitud y secuencia de las preguntas. Puede indicar también la necesidad de agregar algunas preguntas y de eliminar otras. Es recomendable que la prueba del cuestionario se realice en individuos de la misma población en la cual se va a aplicar.

5.4 En cuanto a la forma del cuestionario, se puede considerar lo siguiente :

El tamaño que depende del número de datos necesarios. Se deben evitar los cuestionarios de tamaño excesivo.

Los rubros que sirven para localizar el cuestionario deben ir en el lugar más importante. Las preguntas relativas a un mismo rubro deben estar juntas.

Las instrucciones, si son cortas, deben ir cerca del rubro a que se aplican, puestas en renglones cortos y no a lo ancho del cuestionario. Si sirven para todo el cuestionario, debajo del encabezamiento. Si son muy largas, al reverso o en hoja anexa.

Para las respuestas : " si ", " no ", puede hacerse casilleros en que debe especificarse si debe tarjarse lo que corresponde o lo negativo.

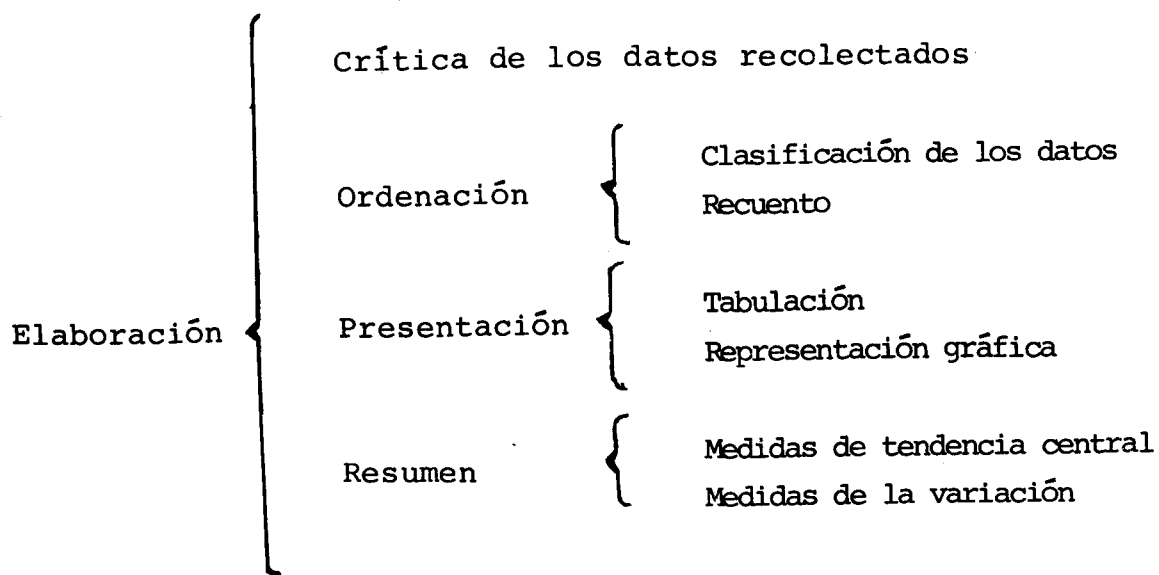
ELABORACION DE DATOS : ORDENACION

1. GENERALIDADES

1.1 La elaboración de datos es un proceso estadístico que permite la ordenación, presentación y resumen de los datos de manera de hacerlos inteligibles y manejables.

1.2 La elaboración de los datos recolectados permite hacer una descripción de naturaleza estadística de los hechos observados. La masa de datos obtenida no permite interpretarlos. Entonces se ordenan sistemáticamente, clasificándolos y haciendo un recuento del número de unidades de observación que pertenecen a cada clase propuesta. En seguida se presenta esta situación en tablas o gráficos, que permiten destacar los hechos fundamentales, con lo cual ya es posible la interpretación y manejo de los datos. Sin embargo, este conocimiento se puede hacer más preciso al resumir toda la masa de datos a un par de valores característicos, fácilmente manejables y comprensibles, que reflejan las propiedades principales del hecho observacional en estudio.

1.3 Las etapas de la elaboración de datos estadísticos se puede sintetizar en un simple esquema.



## 2. CRITICA DE LOS DATOS RECOLECTADOS

Consiste en someter a control los cuestionarios con el objeto de criticar los datos que contienen. Se someten a juicio las respuestas defectuosas, se verifican omisiones, se constatan imperfecciones como datos poco legibles. En muchos de estos casos se puede intentar reparar la deficiencia volviendo a la fuente. Sin embargo, no siempre esta solución es posible, sea por falta de tiempo, por el elevado costo que significaría, por la dificultad de encontrar al entrevistado, o simplemente porque el hecho ya ocurrió y no es posible volver a observarlo en las condiciones establecidas. En estos casos se produce la pérdida de la información correspondiente.

La crítica de los datos debe ser la primera operación a ejecutar como parte del proceso de elaboración de los datos. También puede ser considerada como una etapa complementaria de la recolección de los datos. Donde quiera que a ella se le ubique, lo que solo tiene importancia didáctica, ella debe ser un paso inexcusable a continuación de la recolección o previo a la clasificación y recuento.

Mediante la crítica es posible llegar a realizar la clasificación y recuento en forma depurada, evitándose así imperfecciones en los datos, que posteriormente pueden influir en forma negativa sobre las conclusiones que de ellos se extraen.

### 3. ORDENAMIENTO: CLASIFICACION Y RECUESTO

3.1 Una vez recogidos y criticados, los datos deben ser clasificados. Se confeccionan grupos homogéneos, llamados clases que permiten el agrupamiento de las unidades de observación en ellos mediante la operación de recuento.

El resultado de la clasificación y recuento, sea de uno (simple) o varios (combinada) caracteres, produce las llamadas series estadísticas o simplemente distribuciones.

Conforme a la naturaleza de las clases se les da una denominación especial. Cuando se observa un carácter a través del tiempo, o la respuesta de un hecho fuera una fecha, se habla de serie cronológica (histórica o temporal). Ej: número de casos de fiebre aftosa, en la población bovina de una ciudad, por mes. En las series cronológicas permanece fijo el carácter y el espacio. Otro producto de la clasificación y recuento que recibe el nombre de serie es la llamada serie geográfica (territorial o de espacio). Ellas son resultantes de la observación de un carácter a través de diferentes lugares. Ej: número de bovinos por municipio, en un momento de tiempo. En las series geográficas permanece fijo el carácter y el tiempo.

Las series ya nombradas se referían a la observación de un carácter o a través del tiempo o a través del espacio. En muchas ocasiones se fija el tiempo y también el espa

cio, importando sólo las diferentes modalidades que presenta un carácter de expresión cualitativa (atributos) o las diferentes magnitudes que presenta un carácter de expresión cuantitativa (variables). Los atributos se refieren a cualidades de los individuos como ser : raza, sexo, aptitud zootécnica, especie. Las modalidades de ocurrencia de cada uno de estos atributos no son posibles de medir (en el sentido cuantitativo). Se puede considerar la presencia o ausencia de cada modalidad, contabilizando la cantidad de veces que ocurren, considerando un conjunto de n unidades de observación. Al resultado de la clasificación y recuento de atributos se llama Serie de Atributos. Algunos estadísticos consideran las series geográficas como atributos.

Las variables se refieren a magnitudes que pueden variar al ser observadas en varios individuos. En este caso el carácter tiene expresión cuantitativa, como ser : edad, peso, estatura. Al observar la ocurrencia de un carácter en varios individuos, se aprecia que se presentan diversas magnitudes del carácter (clases) y que los individuos se pueden agrupar teniendo en cuenta el valor de esa magnitud. El resultado de la clasificación y recuento de variables se llama Serie de Variables (distribuciones de frecuencias). En este último caso se puede tratar de variables discretas o continuas según admitan expresiones enteras o fraccionarias. También de acuerdo con la cantidad de variables consideradas se habla univariada, bivariada o multivariada.

En general, se puede decir que las series o las distribuciones en la realidad aparecen muchas veces conjugadas, formando situaciones combinadas.

Las series cronológicas las estudiaremos en forma separada.



### 3.2 Clasificación

La clasificación de los datos es una operación estadística tendiente a permitir el agrupamiento de las unidades de observación en clases (categorías, intervalos numéricos, modalidades). Estas clases deben ser mutuamente excluyentes, es decir, una unidad de observación no puede ser clasificada simultáneamente en dos clases, con la misma escala. El conjunto de clases debe ser exhaustivo, es decir, no puede quedar ninguna unidad de observación sin tener una clase donde ser clasificada.

Según el número de caracteres considerados en la clasificación, esta puede ser simple o combinada. Se llama clasificación simple, cuando las clases son determinadas por la variación de un sólo carácter. Se llama clasificación combinada cuando las clases son determinadas por la variación de dos o más caracteres simultáneamente. Ejemplo de clasificación simple: raza de los cerdos de un municipio. Ejemplo de clasificación combinada : (dos entradas) : raza y sexo de los cerdos de un municipio.

La clasificación también depende del tipo de dato, según se trate de atributos, cuantitativos discretos o cuantitativos continuos. Si los datos son atributos, puede ser que cada modalidad de ocurrencia constituya una clase. Sin embargo, cuando el número de modalidades posibles, es alto, entonces conviene agrupar estas de acuerdo a su analogía, sin vulnerar la condición de que dentro de una clase deben quedar individuos de similares características con respecto al carácter estudiado. Ej: a partir del registro de una clínica veterinaria de perros, se quiere clasificar los enfermos por causa. Resultaría una cantidad excesiva de clases (enfermedades). Se puede reunir en una sola clase aquellos cuadros clínicos que presentan cierta analogía. Se podría considerar como clases grandes grupos de

enfermedades como ser : microbianas, metabólicas, traumáticas, etc.

Si los datos son discretos, generalmente, cada clase está determinada por la unidad. Ej: Número de hijos por familia. En algunos casos, por ser muy extensa la diversidad de clase, también puede ser conveniente agrupar varias unidades en una sola clase, teniendo en cuenta los objetivos del estudio . Ej: La clasificación de los rebaños según su población. Una clasificación, en que cada unidad fuese una clase, sería muy extensa, de ahí la necesidad de agrupar en una clase varias unidades.

Si los datos son continuos, los valores que presentan las unidades de observación pueden ser de cualquier magnitud entre ciertos límites. Por ejemplo, la estatura, la edad, el peso corporal, etc. Por muy grande que fuese la coincidencia de valores, ocurre una gran cantidad de valores diferentes. Imaginemos que en un grupo de 300 novillos tuviéramos 240 valores diferentes de peso corporal. Si cada valor diferente se considera como una clase, tendríamos 240 clases. Aún cuando estas clases fueran ordenadas de menor a mayor, tendríamos una tabla con 240 clases, lo que no permitiría hacer comprensible este conjunto de valores, por muy hábil que una persona fuera para interpretar datos.

De ahí que se haga necesario " agrupar " valores próximos en intervalos de clases. La primera pregunta que surge es, qué cantidad de clases o intervalos de clases tendremos ? Algunos autores recomiendan un número de clases de 10 a 20; otros de 8 a 16 y los hay que aconsejan entre 6 y 14. La decisión depende de varios aspectos, entre ellos si se van a calcular medidas de posición y variación a partir de los datos agrupados, del número de unidades de observación del estudio, de la regularidad

de distribución entre los límites, de los objetivos de la presentación y de la experiencia que se tenga con dichos datos. A mayor número de intervalos de clases, mayor trabajo y mayor exactitud en el cálculo de las medidas. Si el número de clases fuese muy grande, debemos considerar que se puede perder la finalidad de la agrupación de los datos. A menor número de intervalos de clases, menor trabajo y también menor exactitud en cálculos posteriores.

En este curso indicaremos con la letra  $m$  el número de intervalos de clases. Su determinación es en cierta forma arbitraria, ya que es difícil dar una regla invariable. Podríamos repetir la afirmación de Croxton y Cowden (Estadística general aplicada. Fondo de Cultura Económica. México. 1948) en el sentido que " raras veces deberán usarse menos de 6 u 8 clases, y que más de 16 serán útiles solamente para trabajar con datos muy abundantes ".

Teniendo en cuenta los comentarios hechos se puede agregar que la diferencia entre el valor mayor y el valor menor en la muestra o universo en estudio, puede ser dividido en un número  $m$  de clases, de manera que la amplitud y los valores límites de cada intervalo de clase sean cómodos para trabajar.

Si a los resultados de la observación de un hecho les llamamos  $X$ , la diferencia entre el valor mayor ( $X_{\max} = l_2$ ) y el menor ( $X_{\min} = l_1$ ) se llama recorrido, amplitud total, campo de variación o rango (1).

$$\begin{aligned} (1) &= l_2 - l_1 \\ &= X_{\max} - X_{\min} \end{aligned}$$

Supongamos que se estudia el peso del cuerpo de 50 cerdos mestizos de alrededor de 6 meses de edad. El peso mayor observado ha sido 104 kg y el menor 56 kg. Entonces

$$l_2 = 104$$

$$l_1 = 56$$

$$(1) = 104 - 56 = 48$$

Decidimos trabajar con 6 intervalos de clase, por tanto

$$m = 6$$

La amplitud del intervalo de clase  $i$ -ésimo se designa por  $c_i$ . Es preferible que todos los intervalos de clases tengan la misma amplitud ( $c$ ) ya que así se facilitan los cálculos posteriores. Sin embargo, en algunos casos es válido el uso de intervalos de clase con diferentes amplitudes, debido a la naturaleza del hecho estudiado.

La amplitud del intervalo puede ser determinada, dividiendo el recorrido por el número de clases deseado.

$$c = l/m$$

$$c = 48:6$$

$$= 8$$

En resumen, hasta ahora se tiene

a) el recorrido  $(1) = 48$

b) cantidad de clase  $m = 6$

c) la amplitud del intervalo de clase  $c = 8$

Cada intervalo de clase queda delimitado entre un límite inferior y un límite superior. En consecuencia el límite superior del segundo intervalo será al mismo tiempo el límite inferior del tercer intervalo. Esta situación se repite en  $m-1$  intervalos. De manera que al hacer el recuento vamos a encontrar unidades de observación cuyos valores coinciden con un límite, y que por tanto pueden ser clasificados en dos intervalos simultáneamente, lo que contradice la propiedad que las clases deben ser mutuamente excluyentes.

Para evitar situaciones como éstas se adopta el uso de signos como

$| - \circ - |$

Si se escribe  $2 - | 4$  indica que se excluye el valor límite 2 y se incluye el valor límite 4.

En nuestro ejemplo del peso de los cerdos los dos primeros intervalos, determinados por el procedimiento ya indicado, serían

56 - 64

64 - 72

La forma como se han anotado no señala un criterio para clasificar un cerdo cuyo peso sea igual a 64 kg. Se podría clasificar tanto en el primer como en el segundo intervalo.

De acuerdo a la proposición hecha, la forma de anotar los intervalos podría ser

a)  $56 - | 64$

$64 - | 72$

b)  $56 | - 64$

$64 | - 72$

En el caso a el cerdo se clasificaría en el primer intervalo, ya que  $- |$  indica que se incluye el valor 64. En cam

bio, en el caso b, el cerdo sería clasificado en el segundo intervalo, ya que 64  $\left| \text{---} \right.$  indica que en ese intervalo se incluye el valor 64.

La forma de encontrar los valores límites de los intervalos de clases es tener el valor  $l_1 = 56$  el cual se convierte en el límite inferior del primer intervalo de clase. Se le suma el valor  $c$  y encontramos el límite superior del primer intervalo y a la vez el límite inferior del segundo intervalo.

$$56 - 64$$

Si a 64 kg lo consideramos límite inferior del segundo intervalo y le sumamos  $c$  obtenemos el límite superior del segundo intervalo.

$$64 - 72$$

A su vez 72 kg es también el límite inferior del tercer intervalo. Si a él le sumamos  $c$  obtenemos el límite superior del tercer intervalo e inferior del cuarto intervalo.

$$72 - 80$$

El procedimiento continúa en forma sucesiva hasta obtener el límite superior del último intervalo ( $i=1,2 \dots m$ ; donde  $m=6$ ). En nuestro ejemplo este valor será 104 kg, valor que coincide con  $l_2$ .

Para evitar la ambigüedad de clasificación de una unidad de observación cuyo valor coincide con un límite, adoptaremos, en este caso, la inclusión en el intervalo de clase en el cual el valor es límite superior.

Entonces,

INTERVALO DE CLASE

56	—	64
64	—	72
72	—	80
80	—	88
88	—	96
96	—	104

En el primer intervalo obviamente también se debe incluir el límite inferior ya que forma parte del campo de variación de este estudio.

Otra solución a la agrupación de los pesos de los cerdos podría ser, utilizar cinco intervalos de clases ( $m=5$ ), aumentando ligeramente el recorrido de modo que comience en 55 kg y termine en 105 kg. Con este procedimiento se tienen amplitudes de clase de valor 10 lo que es muy cómodo para trabajar. Lo mismo que los promedios de los límites de cada intervalo de clase serían múltiplos de 10, hecho que, como veremos más adelante, también facilita el trabajo de cálculos.

INTERVALO DE CLASE

55	—	65
65	—	75
75	—	85
85	—	95
95	—	105

Recuento

El recuento es una operación mediante la cual se determina la cantidad de unidades de observación (frecuen-

cia absoluta) que corresponde a cada clase (frecuencia de clase).

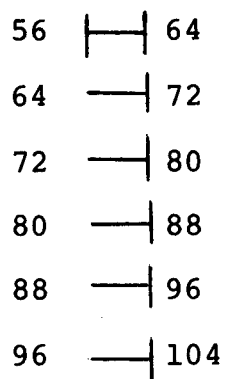
La forma de llevar a cabo el recuento va a depender, entre otros aspectos, del volumen de datos. Existen algunas formas de realizarlo.

Recuento manual

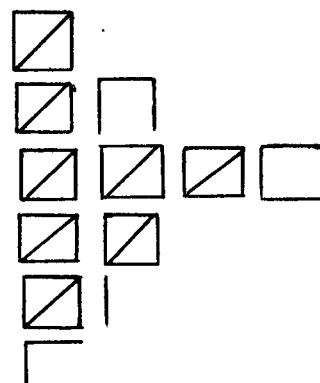
Consiste en trazar rayas frente a las clases indicando con cada raya una unidad de observación. Se utiliza esta forma cuando el número de unidades de observación no es muy grande.

Ejemplo :

INTERVALOS DE CLASE



FRECUENCIA



Este procedimiento presenta inconvenientes cuando hay que hacer el recuento de clasificaciones cruzadas.

Tarjeta simple

Cada unidad de observación está representada por una tarjeta. Su tamaño es como una de visita o 10 x 15 cm. Los datos se anotan en lugares predeterminados en una tarjeta modelo. Su tamaño limita el número de caracteres a considerar, siendo el máximo 10 ó 12. Permite hacer el recuento de clasificaciones cruzadas. La información contenida en los cuestionarios se



traspasa a ella, tratando de conservar el orden que ofrecen. Se debe comenzar por un borde, supongamos el superior izquierdo y continuar hacia abajo y a la derecha o viceversa si se comienza del borde superior derecho. Se deben utilizar, en cada rubro, códigos simples y fáciles de memorizar.

Estas tarjetas facilitan mucho el recuento porque son fáciles de manejar, rápidas de llenar y permiten la combinación de clasificaciones.

### Tarjeta de perforación marginal

Para establecer un modelo de tarjeta hay que tener cierta experiencia y una preparación adecuada y hay que trabajar sin apremios. Aún así, puede ser conveniente hacer varios ensayos, repetidos en condiciones difíciles, antes de empezar a utilizar la tarjeta.

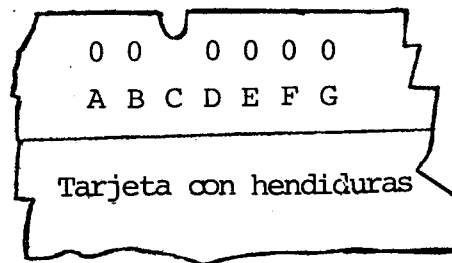
Es importante que el modelo de tarjeta sea sencillo y fácil de emplear pues de nada sirve ponerse a buscar datos si no es posible obtenerlos con la exactitud, el detalle y la prontitud indispensables. La cartulina o el papel de las tarjetas deben ser de buena calidad y su impresión debe hacerse con una disposición tipográfica clara y que no se preste a equívocos.

La tarjeta debe ir acompañada de instrucciones detalladas para su empleo. Las instrucciones escritas son indispensables como elemento de consulta cuando el usuario tropieza con una duda y permiten obtener la necesaria uniformidad en las anotaciones.

Cuando el número de unidades sobre las que es necesario reunir información pasa de unos centenares, las tarjetas de uso más común son las de perforación marginal, de

nominadas indistintamente tarjetas Paramount, Cope Chat o McBee, cuya característica principal es una serie de perforaciones prácticas a lo largo de uno o varios de sus bordes. Cada perforación es un elemento de información que se identifica de manera inequívoca por un número o por una letra de la clave utilizada. El ángulo superior derecho de la tarjeta está cortado en sesgo para facilitar el archivado, pues tan pronto como una tarjeta está mal colocada aparece en el rimerio una esquina que sobresale de las demás.

Una vez anotados en las tarjetas los datos necesarios, se abren con un sacabocados los orificios correspondientes, para transformarlos en muescas de forma de U, según se indica en la ilustración siguiente :



Para extraer las tarjetas correspondientes a un grupo de clasificación determinado, se forman lotes de mil tarjetas de manera que coincidan los ángulos cortados. Por el orificio correspondiente a la característica del grupo se introduce cuidadosamente una aguja larga, como las utilizadas para tejer, moviéndola suavemente de lado a lado. Todas las tarjetas que tienen en lugar del orificio una muesca quedarán en el rimerio al levantar la aguja y podrán contarse sin dificultad. La selección es rápida y exacta; en un minuto pueden clasificarse por el procedimiento indicado mil tarjetas.

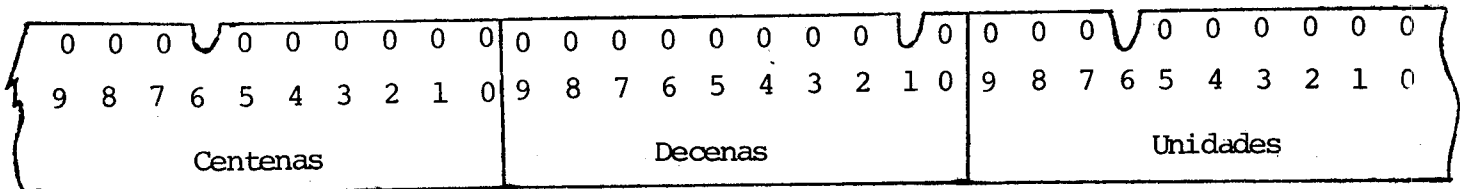
Las tarjetas del tipo indicado no pueden llevar un número ilimitado de perforaciones y su uso no resulta práctico cuando haya que manejar menos de varios centenares o

más de pocos millares. Cabe, sin embargo, modificar el modelo para adoptarlo a la consignación de los datos básicos que han de ser analizados por los servicios de zona encargada de la ejecución del programa. Con los datos tomados de las tarjetas pueden prepararse tabulaciones para comunicar la información más importante a los escalones superiores.

El sistema de cifrado que haya de utilizarse para transcribir y registrar los datos debe elegirse cuando se establezca el modelo de tarjeta. La finalidad fundamental de un sistema de cifrado es distinguir todas las categorías en que pueden entrar los datos, de manera que el encargado de reunir la información pueda anotar en cada caso la categoría apropiada. La cifra se utiliza luego para registrar esa información mediante perforaciones, con arreglo a una clave que comprenda todas las categorías necesarias, representadas por sendos orificios numerados unívocamente. Es necesario prever el registro de categorías especiales, como "desconocido", "no disponible" y "varios".

Existen numerosas claves para trabajos especiales. Los esquemas siguientes representan algunas de las claves numéricas de uso común.

Clave decimal (0-9)



Esta es la más sencilla de las claves numéricas y se llama "de cifrado directo". En la figura se ven las perforaciones correspondientes a la categoría 616. Hay una serie de 10 orificios para las unidades, otra para las decenas, otra para las centenas y así sucesivamente. Cada



Es difícil concebir hoy día que exista una persona que no conozca las tarjetas perforadas. Las cuentas de gastos, la lotería esportiva, etc. son ejemplos de los muchos usos que hoy tiene.

Esta tarjeta tiene 80 columnas y 10 filas impresas con números.

Las columnas son numeradas horizontalmente, de izquierda para la derecha, de uno a ochenta. En cada columna, una letra, un dígito o un carácter especial pueden ser guardados, completando un total de ochenta caracteres por tarjeta. Por ejemplo, supongamos que se desee perforar en una tarjeta el número de empleado 40875. Como cada columna solamente puede guardar un carácter, cinco columnas serían necesarias para contener el número 40875. De hecho, cualquiera de cinco columnas consecutivas servirían. Si decidimos usar las cinco primeras columnas para aquel fin, entonces el dígito 4 perforado en la columna uno, el dígito 0 en la columna dos, el dígito 8 en la columna 3, el dígito 7 en la columna cuatro y el dígito 5 en la columna cinco.

Además de las ochenta columnas verticales, la tarjeta IBM está también dividida en " dos posiciones de perforaciones " : (1) posición de zona y (2) posición de dígito.

La posición de perforación de zona consiste de tres columnas, dos de las cuales son usadas para perforación de zona, solamente. Ellas representan el espacio en la parte superior de la tarjeta. La primera altura, a partir de arriba, es denominada " altura doce ". La siguiente, que es la altura once, es comúnmente llamada de " altura X ". Una perforación en la altura X (referida como perforación X), en una columna seleccionada (de uno a ochenta), es frecuentemen-

te usada para distinguir un dato tarjeta (o tarjetas) de los demás de una masa. La tercera altura es conocida como " altura cero " y puede ser usada como una posición de perforación de zona o de dígito.

La posición de perforación de dígito consiste de diez alturas, numeradas verticalmente, para guardar los dígitos 0-9, correspondientes a los números impresos en la tarjeta. La altura cero guarda el dígito 0, la altura uno el dígito 1; la altura dos el dígito 2, y así por delante.

El código de máquina usado para guardar informaciones en una tarjeta IBM es el código HOLLERITH. El Dr. Herman Hollerith concibió la técnica de guardar determinados datos en una tarjeta de dimensiones padronizadas, siguiendo un código predefinido, a fin de producir los orificios en localizaciones específicas de aquél.

#### Registro numérico y alfabético de informaciones

Los números son registrados en una tarjeta perforándose solamente " un orificio " en una columna, para cada dígito. Si necesitáramos guardar, por ejemplo, un número de serie de nueve dígitos en una tarjeta serían necesarios nueve orificios practicados en nueve columnas consecutivas. Las perforaciones serían hechas en la posición de perforación de dígito.

Si " datos alfabéticos " deben ser registrados, son necesarios " dos orificios " en cada columna, para " cada letra " del alfabeto. Por ejemplo, la palabra EQUINO perforada en una tarjeta requiere seis columnas consecutivas con " dos orificios " en cada una. Uno de los orificios debe estar en la posición de perforación de " zona " y

el otro en la posición de perforación de dígito. En otras palabras, es necesaria " una perforación " para registrar " un dígito " en una columna, en cuanto que " dos orificios " son precisos para guardar " una letra ". Esa técnica es padrón en todas las tarjetas IBM y se basa en la teoría de HOLLERITH para código de máquina.

Tabla de codificación

Para describir el lenguaje de máquina más claramente, tablas separadas son presentadas a continuación; una titulada CODIFICACION NUMERICA y la otra CODIFICACION ALFABETICA.

Tabla de Codificación Numérica IBM

Dígito	Posición de Perforación de dígito	Posición de Perforación da zona	Altura
0	si	no	0
1	si	no	1
2	si	no	2
3	si	no	3
4	si	no	4
5	si	no	5
6	si	no	6
7	si	no	7
8	si	no	8
9	si	no	9

Codificación numérica IBM

Si fuésemos a perforar el número 14036 en las columnas trece y diecisiete, procederíamos como sigue :

La columna 13 debería tener una perforación en la altura 1 para el dígito 1  
La columna 14 debería tener una perforación en la altura 4 para el dígito 4  
La columna 15 debería tener una perforación en la altura 0 para el dígito 0  
La columna 16 debería tener una perforación en la altura 3 para el dígito 3  
La columna 17 debería tener una perforación en la altura 6 para el dígito 6

### Codificación alfabética IBM

Como existen tres alturas de zona para la perforación de caracteres alfabéticos en la tarjeta, a saber, alturas doce, once y cero, el alfabeto es dividido en tres partes, cada parte conteniendo un número de letras igual al número de alturas de dígito. Esto es practicado porque " dos perforaciones son necesarias en una tal columna para presentar una letra ". Dependiendo de la letra en cuestión, una de las perforaciones debe estar en las alturas doce, once o cero y la otra deberá situarse en las alturas uno, dos, tres, cuatro, cinco, seis, siete, ocho e nueve. Cualquier conjunto de columnas puede ser usado para cualquiera de los tres grupos alfabéticos.

A - I (primeras nueve letras)

J - R (las nueve letras siguientes)

S - Z (las ocho restantes)

Las letras de A hasta I son, cada una, codificadas, perforándose 1 altura doce y, en la misma columna, las alturas de uno a nueve, respectivamente. Por ejemplo, la letra A sería codificada por una perforación en la altura doce y otra en la altura uno, de la misma columna. La letra B sería codificada por una perforación en la altura doce y una en la altura dos de la misma columna, y así en adelante para las demás letras.



Letra	Posición de Perforación de zona en la Altura	Posición de Perforación de dígito en la Altura
A	12	1
B	12	2
C	12	3
D	12	4
E	12	5
F	12	6
G	12	7
H	12	8
I	12	9

En este punto, ya está claro que la letra I ocupa la última altura utilizable en la tarjeta. Para codificar el segundo grupo de letras, de J hasta R, precisamos partir de la altura uno, otra vez, usando la altura once para la perforación de zona, en vez de la altura doce, y las alturas de uno a nueve para la perforación de dígito.

La letra R ahora ocupa la última altura de la tarjeta. Para codificar las letras restantes, de S a Z, precisamos por tanto, comenzar otra vez usando la altura cero para la perforación de zona y las alturas dos hasta nueve para la perforación de dígito. Como hay solamente ocho letras en el tercer grupo alfabético, la IBM decidió usar las alturas dos a nueve, en vez de las alturas uno a ocho, para la perforación de dígito. Eso fue hecho posiblemente para evitar perforaciones muy próximas unas de las otras, como sería el caso si las alturas cero y uno fuesen usadas, porque en las máquinas antiguas la precisión era algunas veces sacrificada cuando las perforaciones eran practicadas muy llegadas una a otra. Por esta razón, se resolvió guardar un intervalo y se hizo parte del procedimiento padrón, a pesar que las modernas perforadoras hayan sido perfeccionadas a tal punto que la proximidad de las perforaciones no constituye más cualquier problema.

Deberíamos resaltar que cualquier grupo de columnas puede ser utilizado para registrar cualquier palabra, conforme muestra la siguiente discusión. Para ilustrar, véase como la palabra ACID es guardada en las columnas diecisiete hasta veinte, la palabra MONK en las columnas tres hasta seis y la palabra STY en las columnas setenta hasta setenta y dos. En ACID, el A es guardado perforándose las alturas doce (zona) y uno (dígito), directamente abajo, en la columna diecisiete. Cualquier otra columna podría haber sido usada. El código predeterminado, arriba mencionado, ya viene inserto en las perforadoras modernas, de manera que todo lo que el operador tiene que hacer es apretar la tecla correspondiente a la letra deseada en el teclado de la máquina y las perforaciones aparecen en la tarjeta conforme descrito. Experimente, usando tablas, codificar la palabra INVOICE, para ver si Ud. entendió.

Letra	Posición de Perforación de Zona en la Altura	Posición de Perforación de dígito en la Altura
J	11	1
K	11	2
L	11	3
M	11	4
N	11	5
O	11	6
P	11	7
Q	11	8
R	11	9
S	0	2
T	0	3
U	0	4
V	0	5
W	0	6
X	0	7
Y	0	8
Z	0	9

ELABORACION DE DATOS : TABULACION

La tabulación es una operación mediante la cual se presentan los resultados de la clasificación y recuento organizados en tablas. Estas tablas presentan una serie o distribución de frecuencias, simple o combinada, mostrando la forma como se reparten las unidades de observación en las diferentes clases.

1. Tabla es un cuadro mediante el cual se presentan series de datos ordenados en columnas y filas, destacando los aspectos principales de la distribución, facilitando así la descripción e interpretación de los hechos estudiados.

2. Desde el punto de vista de la utilización existen dos tipos de tablas. Las tablas de referencia o muestras que se usan como fuentes de información rápida. En ellas se vacían los datos recolectados, para poder posteriormente construir tablas más pequeñas. En ellas sólo se encuentran valores absolutos y nunca porcentajes. El otro tipo de tabla es la tabla específica o de resumen; son de tamaño más reducidos que las anteriores, contienen una menor cantidad de información, da énfasis a aspectos particulares de los datos. Ellas responden a un objetivo bien preciso generalmente.

3. Obviamente en este curso nos preocuparemos de las tablas de tipo específico. Frecuentemente ellas permiten responder una pregunta. El primer aspecto que se ha de tener en cuenta al elaborar una tabla es que ella tenga un propósito. No se trata de confeccionar una serie de tablas, tantas como sean posibles o tantas como nuestra imaginación des

cubra. La manera más sencilla y directa de definir el propósito de una tabla es a través de la formulación de una o más preguntas que se intenta responder mediante la tabla. Ejemplos :

¿Cuál es la distribución geográfica de la morbilidad en bovinos por fiebre aftosa ?

¿Qué grupos etarios de la población bovina son más afectados por la fiebre aftosa?

4. Una vez establecido el propósito de una tabla, hay que parar a considerar la estructura de la tabla. Tal como lo indica la Resolución nº 886 de la Junta Ejecutiva Central del Consejo Nacional de Estadística de Brasil (1966), la estructura de una tabla tiene elementos esenciales y elementos complementarios.

5. Los elementos esenciales de la estructura de una tabla son los siguientes :

a) Título : Es una declaración que se ubica sobre la tabla y que debe indicar el hecho observado, el lugar y la época en que fueron colectados los datos correspondientes. Es decir, el título debe dar respuesta a las preguntas ¿qué? ¿dónde? y ¿cuándo?

b) Cuerpo : El cuerpo de una tabla está formado por un conjunto de COLUMNAS y FILAS que contienen los datos en orden vertical y horizontal respectivamente. Las CASILLAS o CASAS son cruzamientos de una columna y una fila. Las casillas no deberían quedar nunca vacías, sin llenar, sea con números o con señales convencionales.

c) Cabecal : Es la parte superior de la tabla y debe especificar el contenido de las columnas.

d) Columna indicadora : Es la parte de la tabla que señala el contenido de las filas. En ciertos casos puede ocurrir que una tabla tenga más de una columna indicadora.

e) Trazos : Excluido el título, la tabla debe ser delimitada, tanto en la parte superior como en la parte inferior de la misma. En la parte superior de la tabla se pueden emplear dos trazos horizontales paralelos gruesos que delimiten por encima y por abajo el cabezal y el título de la columna indicadora.

En la parte inferior de la tabla también se puede hacer lo mismo, especialmente cuando es necesario delimitar el total.

En general no se delimitan las tablas con trazos verticales ni en el extremo izquierdo ni en el extremo derecho de la tabla.

Título                      Distribución de los bovinos por edad y sexo  
Municipio XX. Mayo de 1970

The diagram shows a table with a header section and a body section. The header section is divided into two columns: 'Edad' and 'Sexo'. The 'Sexo' column is further divided into 'Machos' and 'Hembras'. The body section has three rows: 'Menores de 1 año', 'Entre 1 y 2 años', and 'Mayores de 2 años'. The table is labeled with 'Cabezal' at the top, 'Cuerpo' on the right, and 'CASILLA' at the bottom. Arrows point from the labels to the corresponding parts of the table.

Edad	Sexo	
	Machos	Hembras
Menores de 1 año		
Entre 1 y 2 años		
Mayores de 2 años		

6. Los elementos complementarios de la estructura de una tabla, convencionalmente se anotan al pie de la tabla. Ellos son los siguientes :

a) Fuente : Señala la (o las) entidad responsable por la obtención o por la elaboración de los datos utilizados . Este elemento generalmente se omite cuando los datos son originales.

b) Notas : Son informaciones generales tendientes a aclarar el contenido de la tabla o a señalar el método utilizado en la recolección o elaboración de los datos.

c) Llamadas : Son informaciones específicas sobre partes determinadas de la tabla destinadas a aclarar el contenido. Muchas veces pueden cobrar la forma de llamadas las citas a varias fuentes de origen de los datos utilizados. Las llamadas se indican en el cuerpo de la tabla con números arábigos, entre paréntesis, a la izquierda de las casillas y a la derecha en la columna indicadora.

7. Enseguida se debe proceder a la colocación de los datos numéricos, es decir, se procede a llenar las casillas de la tabla con los datos correspondientes. En el caso de distribuciones de frecuencia de variables continuas, los intervalos de clases deberán ser presentados conforme ya se ha indicado en la parte relativa a clasificación. En cuanto a la forma de escribir los números se deben tener en cuenta algunas recomendaciones. La parte entera deberá separarse de la parte decimal con una coma (,). La parte entera de los números será separada en grupos de tres algarismos a través de puntos (.) de derecha a izquierda. Se exceptúan a esta recomendación los números que señalan un año calendario (1972).

Se debe evitar en lo posible usar números romanos. El "redondeamiento" de números se puede efectuar sobre las siguientes consideraciones. Si el algarismo a ser eliminado es 0, 1, 2, 3, 4, permanece inalterado el último algarismo a permanecer. En cambio, cuando el algarismo a eliminar es 6, 7, 8, 9, entonces se aumenta en una unidad el algarismo a permanecer.

Ej.: 15,13 ————— 15,1  
25,38 ————— 25,4

Cuando el algarismo a eliminar es 5 y el algarismo siguiente es mayor que 0, entonces se aumenta en una unidad el algarismo a permanecer. Cuando el algarismo a eliminar es 5 y el algarismo siguiente es 0, se aumenta en 1 unidad el algarismo a permanecer si este fuera un número impar; si se tratase de un número par éste permanecerá inalterado.

Ej.: 15,151 ————— 15,2  
15,350 ————— 15,4  
15,450 ————— 15,4

Es conveniente evitar redondeamientos sucesivos, sobre todo cuando se están realizando operaciones.

Cuando se trate de cifras muy extensas, como ser, millares, millones- se pueden simplificar los números, situación que puede ser señalada en el título de la tabla o en una nota al pie.

Otro aspecto dice relación con el uso de porcentajes. Ellos podrían ser utilizados de preferencia con dos

finalidades. Una, cuando se quiere destacar la frecuencia de un hecho en un total. Es más sencillo decir que el 16% de los bovinos de un establecimiento enfermaron de fiebre aftosa, es decir que 60 de los 375 bovinos del establecimiento pecuario enfermaron de fiebre aftosa. Otra finalidad, puede ser la necesidad de comparar la ocurrencia de un hecho en dos o más grupos que tienen diferente número de unidades de observación. Supongamos que en un rebaño A con 83 bovinos, enfermaron 9 de ellos de fiebre aftosa. En un rebaño B de 128 bovinos, 24 enfermaron de fiebre aftosa. Es difícil saber cual de las dos situaciones es mejor. Al computar los porcentajes vemos que en el rebaño A un 13% de los bovinos enfermó y que en el rebaño B enfermó un 19% de los bovinos. No deben ser utilizados porcentajes cuando el número de unidades de observación sea muy pequeño. Por ejemplo se vacunan con una vacuna antiaftosa de tipo experimental 4 bovinos, entonces cada uno representa un 25% del total. Si 3 de ellos se protegen contra la enfermedad se habla de un éxito de 75%, lo que da una falsa impresión de estabilidad de nuestra información, no reflejando la realidad de una escasa experiencia.

Algunos autores recomiendan no calcular porcentajes con números de unidades de observación inferiores a 50. Otros fijan en 20 el límite entre un número reducido en que no debe calcularse porcentaje y un número grande en que se puede calcular. En general es aconsejable expresar los porcentajes sin decimales.

La suma de los datos numéricos de una columna o de una fila se indicará en forma destacada con la palabra total, excepto cuando se trata de una serie geográfica en cuyo caso el total se reemplaza por el nombre del conjunto de la misma. El total puede preceder o seguir a las casillas. Si existieren totales parciales se debe indicar la suma de ellos, por la expresión total general.



La unidades de medida deben ser expresadas de acuerdo con los signos convencionales aprobados oficialmente .  
Ej.:

t	-	tonelada	cm	-	centímetro
kg	-	kilogramo	s	-	segundo
g	-	gramo	min	-	minuto
m	-	metro			

Estos símbolos no deben ser seguidos ni de punto ni de letra "s" que indique plural. El símbolo de las unidades de medida debe ser escrito al final de una expresión numérica mismo que ésta sea fraccionaria.

En las tablas se utilizan símbolos convencionales para indicar la existencia de alguna situación en relación con los datos numéricos. Entre los más comunes están los siguientes :

— Cantidad cero  
... Dato no disponible  
0 Valor numérico menor que 0,5

8. A continuación veremos algunos ejemplos de tablas según las escalas de clasificación de los caracteres considerados. Veremos tanto series simples (distribuciones de frecuencia relativas a un carácter) como series combinadas (distribuciones de frecuencia para dos caracteres simultáneamente). Se considerarán datos cualitativos y cuantitativos.

## a) Serie simple. Escala cualitativa

Título. Distribución de los rebaños bovinos según el tipo de explotación. Lugar NN, 1968.

TIPO DE EXPLOTACION	Nº	%
Carne	62	60
Mixto	26	25
Leche	16	15
<b>T O T A L</b>	<b>104</b>	<b>100</b>

## b) Serie simple. Escala cuantitativa continua

Título. Distribución de cerdos mestizos de 6 meses de edad según clases de peso corporal (kg). Zona YY, 1965.

CLASES DE PESO	Nº	%
55 ———  65	5	10
65 ———  75	12	24
75 ———  85	20	40
85 ———  95	10	20
95 ———  105	3	6
<b>T O T A L</b>	<b>50</b>	<b>100</b>

c) Serie simple. Escala cuantitativa discreta

Título. Distribución de 21 familias según el número de hijos. Lugar NN. 1967.

NUMERO DE HIJOS	Nº	%
0	2	10
1	4	19
2	7	33
3	6	28
4	2	10
<b>T O T A L</b>	<b>21</b>	<b>100</b>



d) Serie combinada. Dos escalas cualitativas.

Título. Distribución de rebaños bovinos afectados por fiebre aftosa según tipo de explotación y tipo de virus diagnosticado. Estado NN. 1967

TIPO DE EXPLOTACION	TIPO DE VIRUS						TOTAL	
	O		A		C			
	Nº	%	Nº	%	Nº	%	Nº	%
Carne	687	71	275	28	13	1	975	100
Leche	333	67	156	31	11	2	500	100
Mixta	330	75	105	24	4	1	439	100
<b>T O T A L</b>	<b>1350</b>	<b>71</b>	<b>536</b>	<b>28</b>	<b>28</b>	<b>1</b>	<b>1914</b>	<b>100</b>

e) Serie combinada. Una escala cualitativa y otra cuantitativa

Título. Distribución de 165 bovinos de 18 meses, mestizos, vacunados contra la fiebre aftosa, según valor de índice seroprotección y respuesta a la descarga de virus C. Lugar. HH. 1968.

INDICE SEROPROTECCION					T O T A L	
	PROTEGIDOS		GENERALIZADOS		Nº	%
	Nº	%	Nº	%		
 2	9	7	29	97	38	23
 2	126	93	1	3	127	77
T O T A L	135	100	30	100	165	100

f) Serie combinada. Las dos escalas cuantitativas

(+) Título. Distribución de 40 hombres sanos según contenido porcentual de oxígeno arterial y venoso. Lugar XY. 1970.

% OXIGENO ARTERIA	% OXIGENO VENOSO						T O T A L
	10	11	12	13	14	15	
17	2	1	-	-	-	-	3
18	-	2	3	1	-	-	6
19	-	1	2	4	1	-	8
20	-	-	2	3	4	1	10
21	-	-	1	2	5	1	9
22	-	-	-	-	1	3	4
T O T A L	2	4	8	10	11	5	40

(+) Tomado de nociones de Estadística de S.Vergar e I.López. Unidad de Bioestadística y Biomatemáticas. Departamento Salud Pública y Medicina Social. Facultad de Medicina. Universidad de Chile.

ELABORACION DE DATOS : REPRESENTACION GRAFICA

1. Gráfico - Es un dibujo que permite, a través de la impresión visual, presentar datos estadísticos ya ordenados, facilitando la rápida comprensión de los hechos.

2. Hemos visto que la tabla daba una imagen de los aspectos destacados de un hecho estudiado. La gráfica permite una apreciación más rápida y sencilla sobre lo que los hechos muestran. Hay cierto tipo de lectores que prestan más atención a los gráficos.

3. Las ventajas de los gráficos se podrían resumir en :

- a) permiten llegar a todo tipo de público en forma sencilla.
- b) facilitan la comparación por la rápida apreciación de similitudes y diferencias.
- c) facilitan la fijación de ciertos conceptos por ser atractivos en su presentación.

4. Sin embargo, los gráficos también tienen limitaciones. Entre ellas las siguientes :

- a) no pueden presentar tantos grupos de datos como una tabla, a riesgo de tornarse complejos.

- b) la gráfica presenta por lo general valores aproximados.
- c) su elaboración consume más tiempo y trabajo que una tabla.

5. Entre los requisitos que debe llevar un buen gráfico se ha de considerar :

- a) debe ser sencillo y auto-explicativo
- b) debe presentar fielmente los hechos en estudio. Es conveniente evitar escalas exageradas.
- c) no contener demasiada información. Se tornan complejos y pierden su efectividad.
- d) deben ser de trazado atractivo. Dibujo limpio, de trazos marcados, leyendas bien ubicadas, etc.
- e) ser adecuados al tipo de escala en que están clasificados los datos.

6. Para elaborar un gráfico, al igual que lo ocurrido con la elaboración de una tabla, se han de tener en cuenta algunas etapas.

- a) definición del objetivo
- b) elección del tipo de gráfico
- c) elaboración del gráfico propiamente tal.

7. Por supuesto que no se trata de confeccionar tantos gráficos como se nos ocurre. Cada gráfico debe ser producto de la necesidad de dar respuesta a una pregunta. Es decir, cada gráfico debe tener un objetivo. PARA QUE se quiere construir un gráfico, que se desea mostrar ?

8. Existe una variada gama de formas de re presentación gráfica, de donde elegir aquel tipo de gráfico que es adecuado. La elección del tipo de gráfico debe estar orientada fundamentalmente por los objetivos propuestos y por la esca la de clasificación en que están expresados los datos.

En términos generales, de acuerdo al ti po de datos y más particularmente al tipo de serie que resulta de la ordenación, se pueden utilizar las siguientes formas de re presentación gráfica estadística :

<u>Serie</u>	<u>Gráficos Estadísticos</u>
Atributos	Barras simples Barras proporcionales Barras agrupadas Diagrama sectorial
Variable discreta	Barras simples
Variable continua unidimensional	Histograma Polígono de frecuencia
Variable continua bidimensional	Diagrama de dispersión
Cronológicas	Barras simples Gráficos lineales aritméticos Gráficos lineales logarítmicos Gráfico polar Pictogramas Estereogramas
Geográficas	Barras simples Barras proporcionales Cartogramas

En este esquema no se han considerado aquellos gráficos no estadísticos y más propios de la administración como gráficos de organización (organigramas) y de control (Gantt).

Excluyendo los pictogramas que son representaciones gráficas de figuras y los cartogramas que son representaciones sobre mapas, los gráficos estadísticos se clasifican en :

- a) Diagramas : gráficos geométricos de dos dimensiones (plano)
- b) Estereogramas : gráficos geométricos de tres dimensiones (volumen).

9. Elegido el tipo de gráfico, es necesario construirlo o elaborarlo. Para elaborar diagramas hay que observar ciertas normas y recomendaciones.

9.1 Título - Cada gráfico debe tener un título conciso y claro que indique los siguientes aspectos :

- a) hecho del que se trata : QUE
- b) lugar donde fue observado : DONDE
- c) época en que fue observado : CUANDO

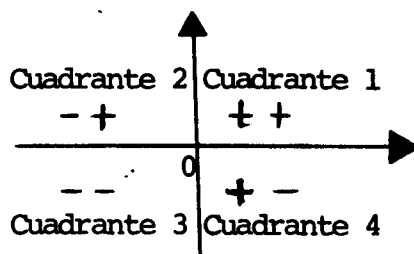
9.2 Tamaño del gráfico - En lo posible el formato del gráfico debe ser rectangular. Debe tratarse de mantener una proporción de 1:1,5 o también  $1:\sqrt{2}$  entre la altura y la anchura del gráfico. Es decir, si el eje horizontal, que debe ser mayor, mide 15 cm entonces el eje vertical debería medir 10 cm.

9.3 Coordenadas - Muchos gráficos estadísticos son contruidos utilizando un sistema de coordenadas cartesianas rectangulares. Está constituido por dos ejes perpendiculares entre sí. Estos ejes son rectas orientadas. Cada



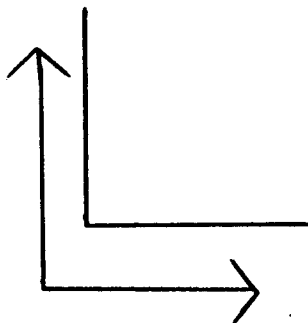
recta orientada tiene una flecha, que indica el sentido, y un punto de origen a partir del cual se aplica la unidad de medida considerada. El punto de origen divide el eje en dos semi-ejes. El semi-eje positivo tiene el mismo sentido de la flecha y el semi-eje negativo tiene sentido contrario.

En el sistema cartesiano estos ejes se cortan formando un ángulo recto, originando cuatro cuadrantes.



El punto 0 es el origen. En general en estadística se utiliza el cuadrante 1 es decir aquel cuadrante limitado por los semi-ejes positivos.

El diagrama generalmente progresa de izquierda a derecha y de abajo hacia arriba.



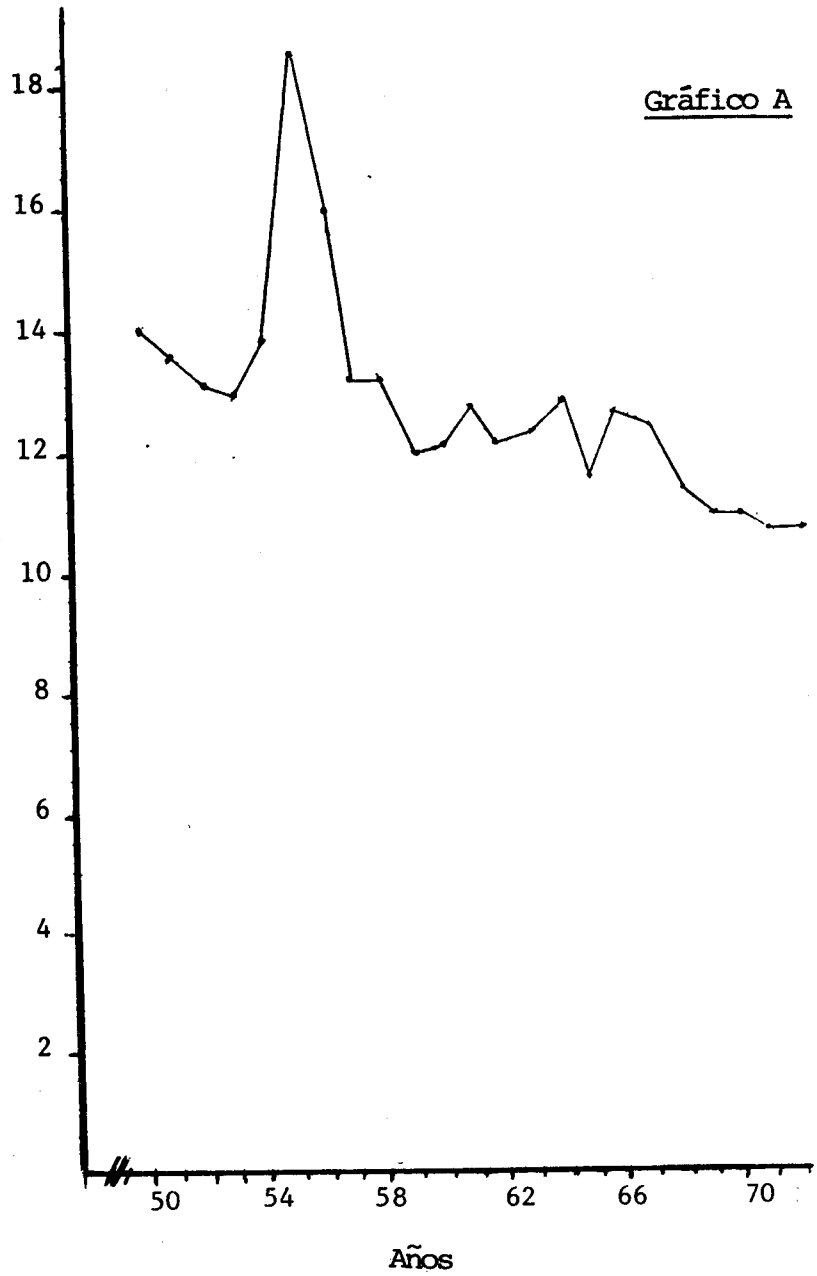
Al eje horizontal se le llama eje de ABSCISAS y al eje vertical se le llama eje de ORDENADAS.

9.4

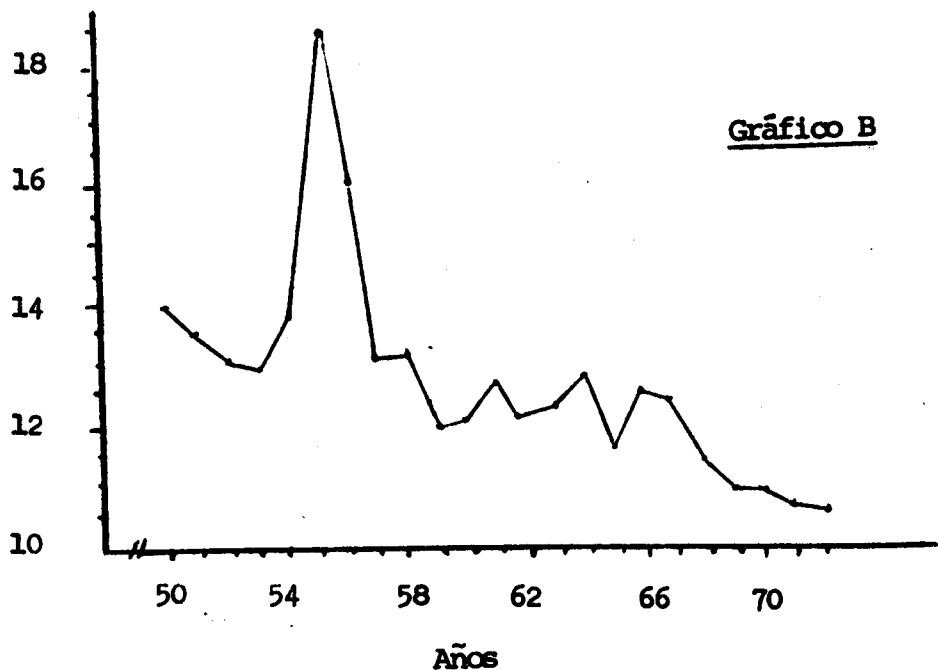
#### El punto cero en la escala vertical

Esta es una regla importante de observar y que a veces se descuida, dando como resultado confusión en la interpretación del gráfico y una visión incorrecta del hecho presentado.

Tasa de morbilidad %

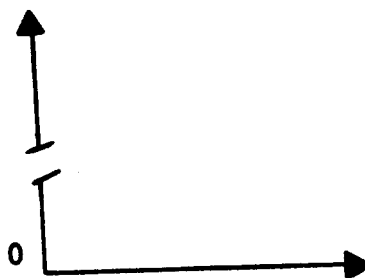


Tasa de morbilidad %



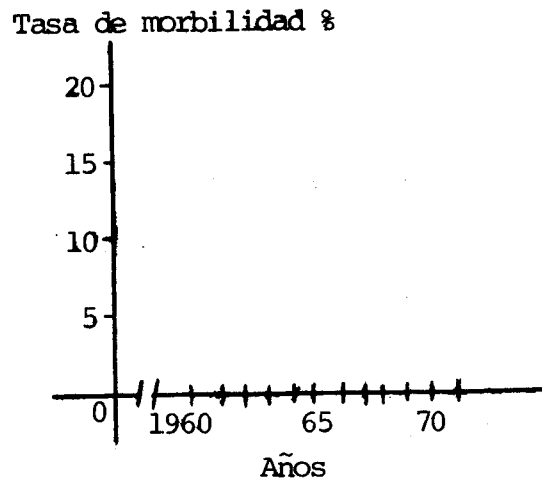
En los gráficos A y B aparece la misma serie de datos. En el gráfico B comienza en 10 la escala vertical. Esta gráfica da al lector una impresión visual no concordante con los hechos. El valor correspondiente al año 56 parece ser el 100% correspondiente al año 53, mientras que en el gráfico A se muestra que es sólo 33% mayor. Esto es consecuencia de que el gráfico B no considera el punto 0 sino que parte del valor 10.

Si existiere falta de espacio se debe insertar un quiebre en el eje vertical.



### 9.5 Leyendas

Cada eje debe llevar una lectura que identifique las unidades utilizadas. En cada eje no precisan estar todas las unidades escritas, a veces no es práctico.



El título de la escala correspondiente al eje de abscisas debe estar abajo de dicho eje, centrado, con letra clara y ser conciso.

El título de la escala del eje de coordenadas debe estar localizado en la parte superior de dicho eje. En ocasiones se escribe presentándolo en el mismo sentido que el eje, esto obliga a modificar la posición del texto para leer.

Estas leyendas deben indicar la naturaleza de la escala y las unidades empleadas.

Cuando se representa más de un hecho en el propio gráfico se debe indicar la clave correspondiente.

#### 9.6 Trazado

Las líneas deben resaltar claramente sobre el fondo de la gráfica. Si se trata de una curva ella debe dibujarse con trazos más gruesos que las coordenadas. Si se traza más de una línea se pueden adoptar trazos diferentes : línea continua, segmentada, de puntos, etc. Deben distinguirse perfectamente una de otra.

#### 9.7 Fuente

Así como en las tablas, los gráficos deben contener al pie una referencia acerca del origen de los datos presentados, cuando ellos no son originales.

### 10. Tipos de gráficos

La mayoría de los gráficos utilizados para presentar datos estadísticos son diagramas. A continuación se presentan algunos tipos de gráficos estadísticos.

#### 11. Barras Simples

Cada clase se presenta por una barra cuyo largo indica la presencia o número de unidades de observación que pertenecen a esa clase. La base de las barras es constante e igual que los espacios entre las barras. En el caso de las series de atributos, cuando no hay un orden implícito de presentación, el orden puede estar dado por la longitud de las barras, cuando no hubiera otro criterio

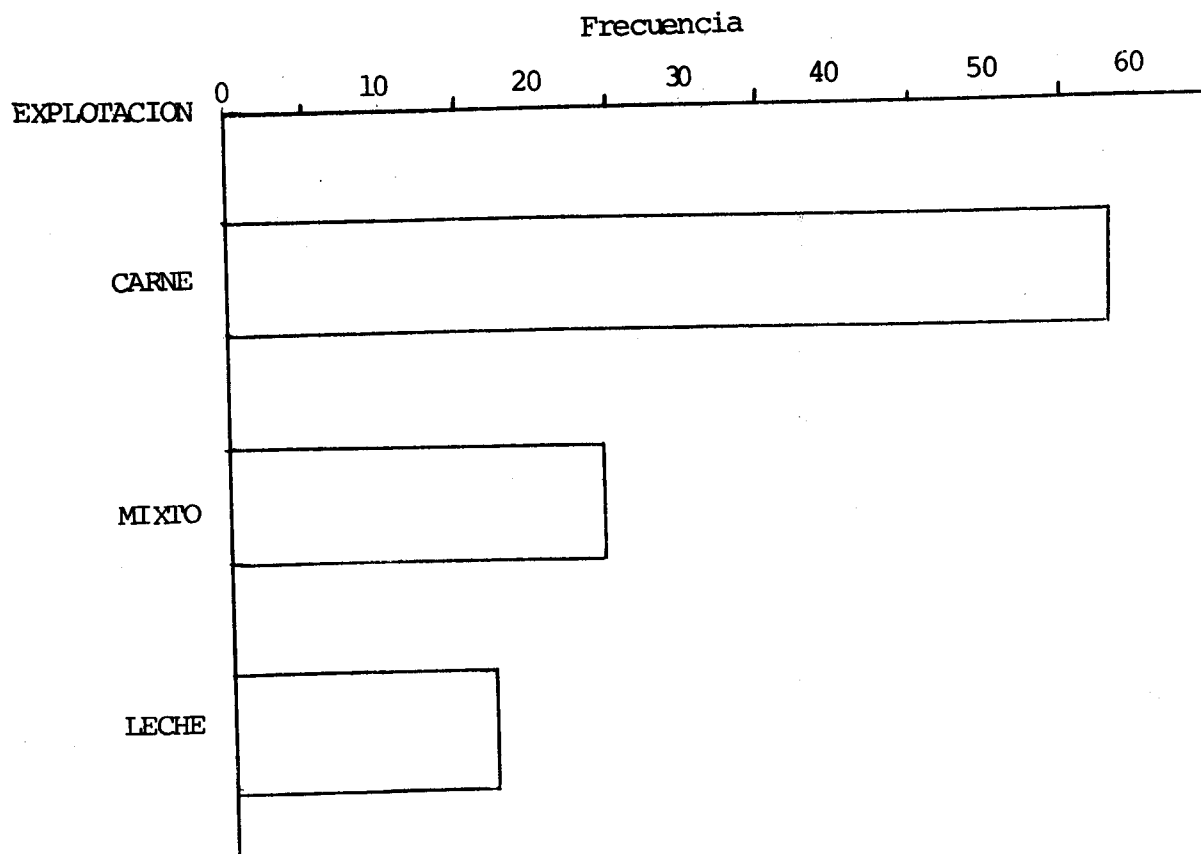
de orden a respetar. En las series de variables discretas (barras lineales), en las cronológicas y en las geográficas, generalmente el orden aparece dado por la propia naturaleza de las clases.

En las series de atributos, especialmente cuando el nombre de cada clase es largo o por facilidad de lectura, conviene construir estas barras simples en sentido transversal.

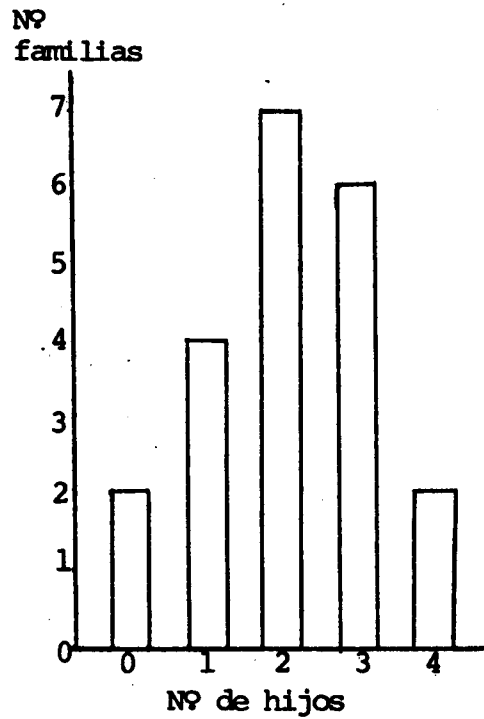
Ejemplos :

a) Distribución de 104 rebaños bovinos según tipo de explotación.

Lugar NN. 1968.



Distribución de 21 familias según el número de hijos. Lugar NN. 1967.

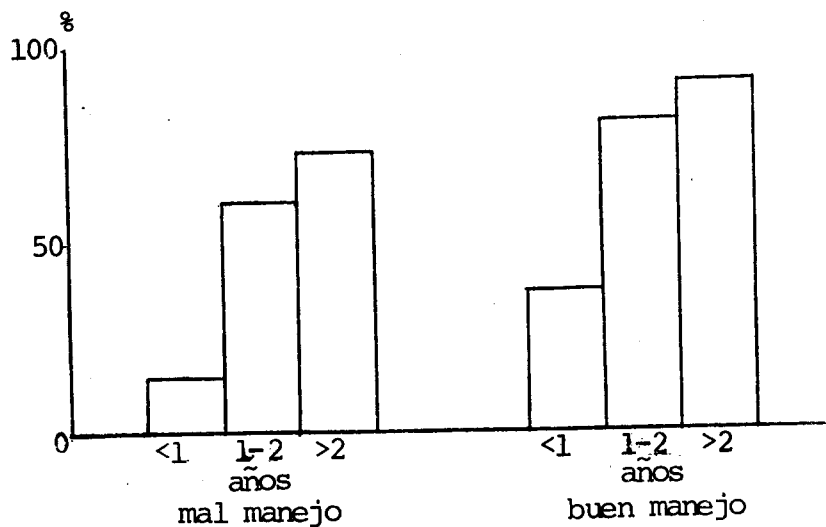


12. Barras agrupadas

Se utilizan fundamentalmente para mostrar alguna relación entre dos hechos. Se distribuyen grupos de barras que corresponden a subdivisiones de una clasificación más general. Las barras de cada grupo deben tener rayados diferentes para cada subdivisión con una indicación de la clave en un lugar apropiado del gráfico. Se prefiere

hacer la clasificación primaria con mayor número de clases para así disminuir el número de claves a colocar en el gráfico, a menos que con esta forma de agrupar se pierda claridad o no se cumpla con los objetivos del gráfico. La comparación de mayor interés es dentro de un grupo de barras.

Relación entre las condiciones de manejo, edad y cantidad de anticuerpos suficientes contra la fiebre aftosa en bovinos. Zona XX. 1970.



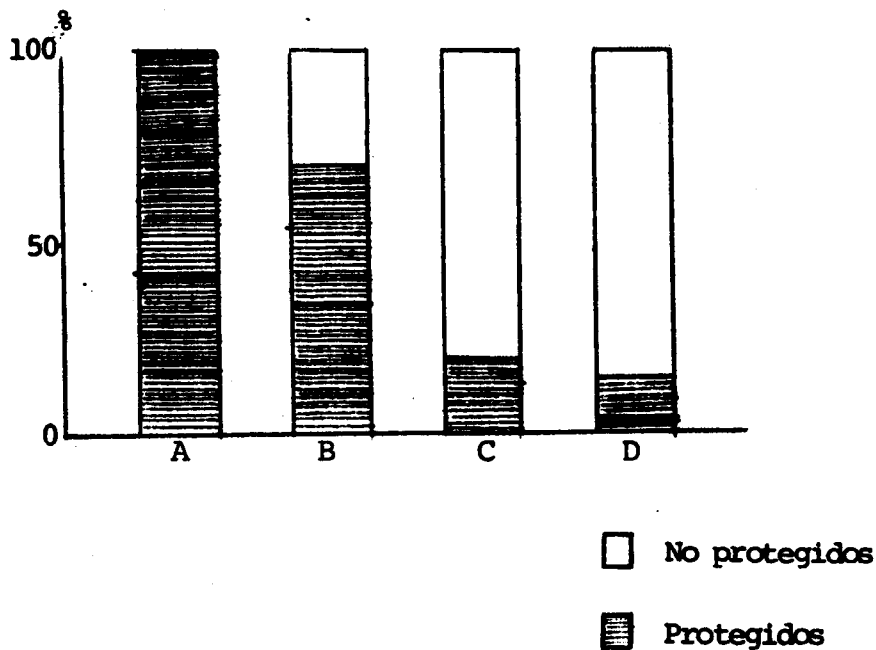
### 13. Barras proporcionales

Este gráfico es muy apropiado para mostrar la composición proporcional de distintas categorías. No es conveniente hacer más de tres subdivisiones en cada barra ya que la comparación se dificulta.



Ejemplo :

Porcentaje de bovinos protegidos  
120 días después de la vacunación antiaftosa con 4 vacunas  
experimentales. Establecimiento YY. 1971.

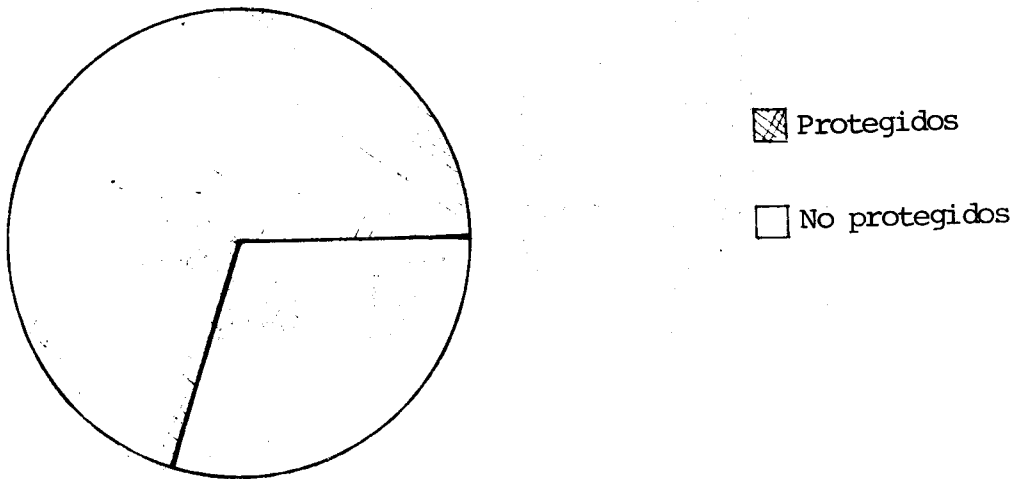


#### 14. Diagramas sectoriales

El principal objetivo de este tipo de gráfico es comparar la frecuencia de una clase con el total, es decir, hacer resaltar el peso relativo o la participación de una clase en el total. Este gráfico sólo se debe emplear cuando las clases consideradas son pocas, no más de 5 ó 6 clases como máximo. En él hay que transformar los porcentajes a grados en un círculo .

Si se toman los resultados de la vacuna antiaftosa experimental B, del ejemplo anterior te nemos que

$$\frac{360^\circ}{100\%} = \frac{x}{70\%} \quad x = \frac{360 \times 70}{100} = 252^\circ$$



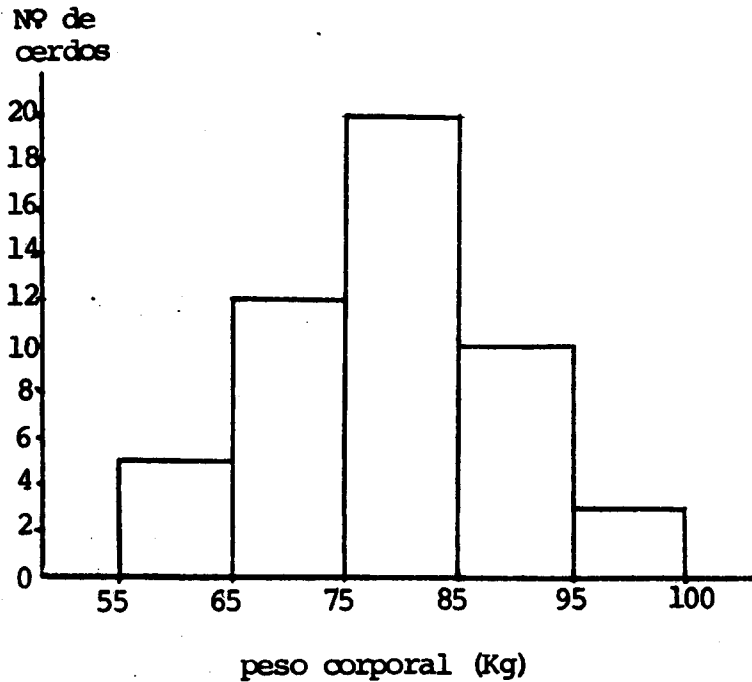
En este tipo de gráfico varía el ángulo y se mantiene constante el radio.

### 15. Histograma

Se utilizan en la representación gráfica de series de variables continuas. Consisten en una serie de rectángulos adyacentes cuyas áreas representan la frecuencia absoluta correspondiente al intervalo de clase. Cuando los intervalos de clases tienen todos la misma am plitud, todos tienen la misma base, y por tanto su altura representa directamente la frecuencia correspondiente al intervalo.

Ejemplo :

Distribución de cerdos mestizos de 6 meses de edad según clases de peso corporal (kg). Zona YY, 1965.



El primer rectángulo representa la clase 55 — 65 kg. En este caso puesto que todos los intervalos de clases tienen una misma amplitud, la altura de cada rectángulo varía según la frecuencia de cada clase. La

amplitud de cada intervalo es igual a 10 kg. De esta manera, también el área de cada rectángulo es proporcional a la frecuencia de clase. En cuanto al área total del histograma es proporcional a la frecuencia total o tamaño de la muestra .

Cuando los intervalos son desiguales en su amplitud esto debe reflejarse en la confección del histograma.

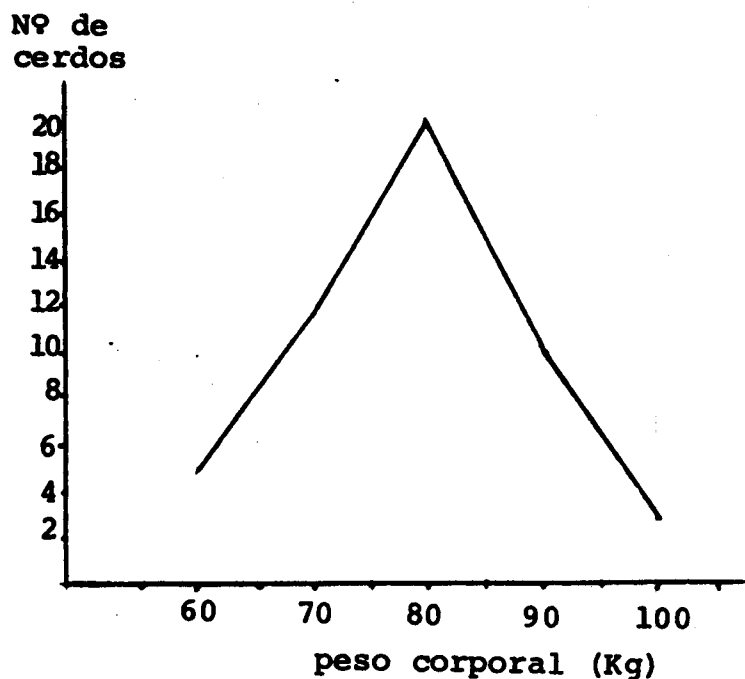
Este tipo de gráfico sólo debe utilizarse para variables continuas.

#### 16. Polígono de frecuencia

Cuando se comparan dos distribuciones de frecuencia pueden ocurrir que los histogramas respectivos se superpongan haciendo poco clara la comparación. Una solución a esta dificultad está dada por otro diagrama de área denominado polígono de frecuencias. Es una forma de representación gráfica muy semejante al histograma y también sólo debe utilizarse para variables continuas. Los contornos del histograma se reemplazan por un polígono que une los puntos medios de sus bordes superiores, de manera que las áreas de los distintos rectángulos se compensan aproximadamente. El polígono puede cerrarse uniendo su primer y último punto con la línea horizontal del gráfico correspondiente al punto medio de las clases adyacentes. Esto debe hacerse de una manera que el área bajo el polígono resulte aproximadamente igual al área bajo el histograma. En ocasiones la escala horizontal parte de cero y al extenderse a la clase próxima se pierde el sentido.

Hay que advertir que la mejor manera de comparar distribuciones de frecuencias es a través de las frecuencias relativas o porcentuales.

Distribución de cerdos mestizos de 6 meses de edad según clases de peso corporal (kg). Zona YY. 1965.



17. Gráficos lineales

Pueden ser aritméticos y logarítmicos. Están indicados cuando se debe representar la relación entre dos variables, por ejemplo : concentración sanguínea de una droga en función de dosis inyectada, tasa de mortalidad infantil a través de los últimos 10 años, etc.

La variable independiente se inscribe en el eje horizontal y la variable dependiente en el eje vertical. La escala en el eje vertical debe comenzar con 0. Si esto implica que un segmento importante del eje no se utiliza y que la escala pierde detalle, se podrá cortar este eje mediante dos líneas.

Frente al valor de la variable independiente se inscribirá con un punto, el valor correspondiente de la variable dependiente para una misma unidad de observación. Los puntos contiguos se unen por líneas rectas.

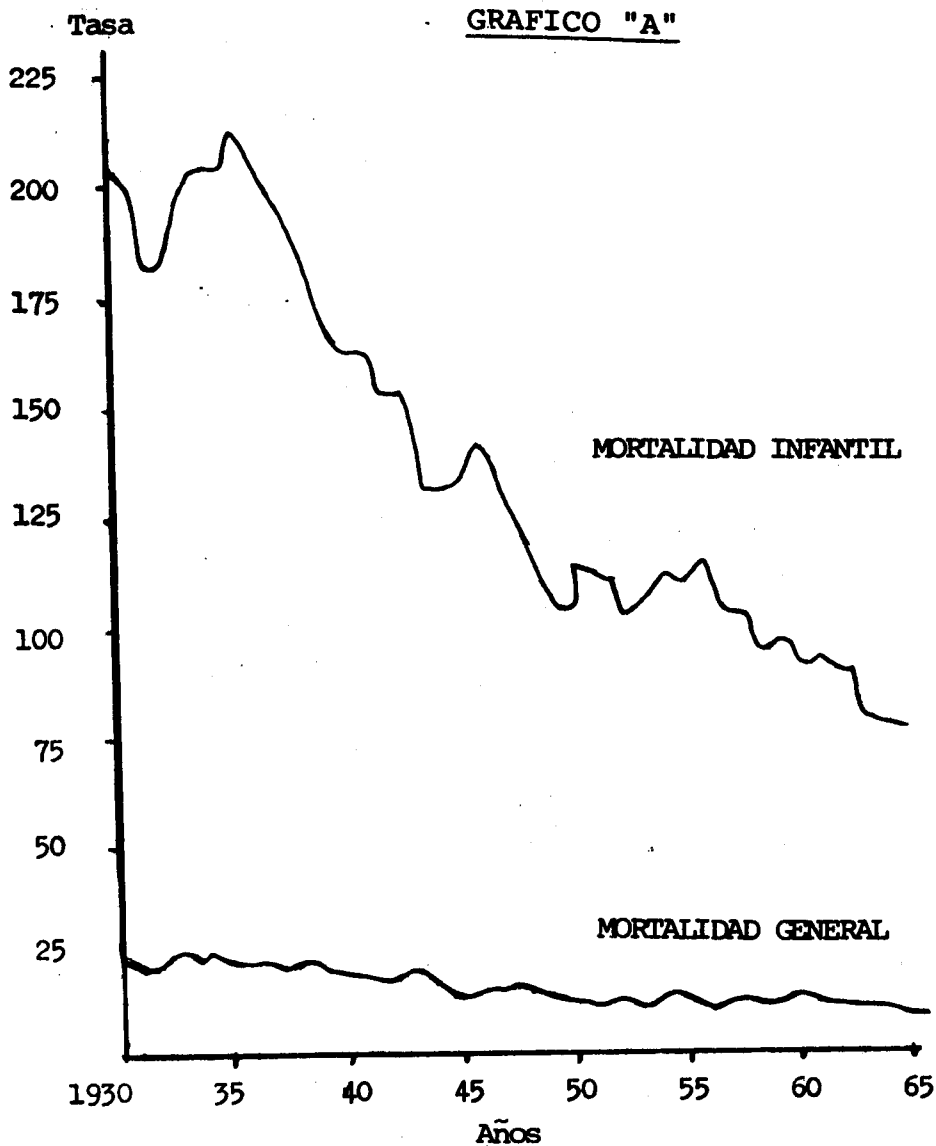
Cuando se tiene una serie en intervalos iguales y por algún motivo se desconoce la información frente a algunos de los valores de la variable independiente, conviene indicar este hecho con una interrupción de la línea. Esto vale sobre todo para series cronológicas en que los datos de un período a otro suelen sufrir grandes fluctuaciones.

Si para la misma escala de la variable independiente se quieren mostrar varios fenómenos, cada uno se inscribirá con líneas de diferente trazo o color. Si estos fenómenos estuvieren en escalas dependientes diferentes se puede inscribir cada escala como eje vertical paralelo al comienzo o al final del eje horizontal.

Cuando se desea conocer la conducta del fenómeno en términos absolutos se usarán escalas aritméticas en ambos ejes. Si se desea investigar cambios relativos de la variable dependiente es útil usar el gráfico semilogarítmico, con el eje horizontal en escala aritmética y el vertical en escala logarítmica. El último objetivo también se logra expresando cada valor de la variable dependiente en relación a un valor base, por ejemplo, la morbilidad por fiebre aftosa en 10 años en un país se puede expresar en términos porcentuales, relacionando la morbilidad de cada año con la morbilidad en 1962.

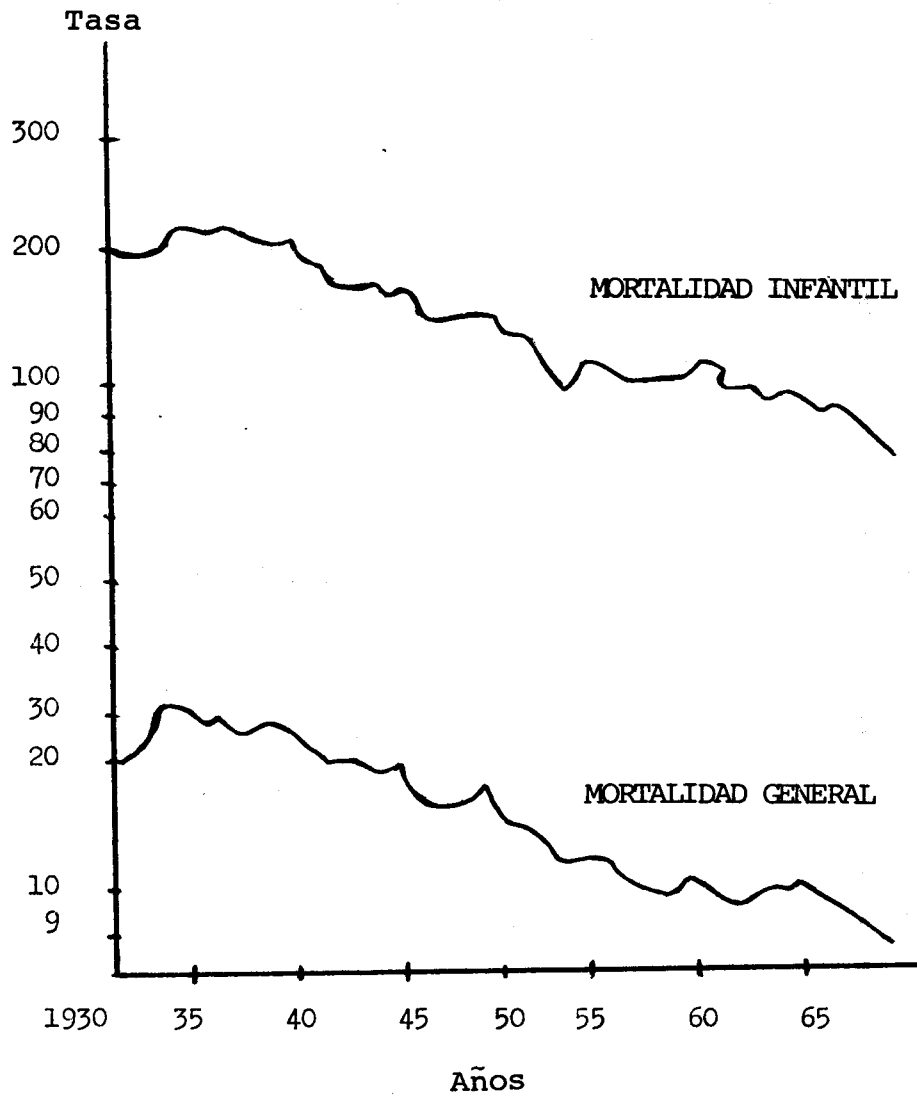
Los gráficos A y B se basan ambos en los mismos datos : Tasas de mortalidad infantil y de mortalidad general en Chile de 1930 a 1969(+)

El gráfico A en escalas aritméticas, muestra los cambios absolutos, el B, en escala logarítmica para las tasas, muestra la reducción relativa de ambos índices.



(+) Tomado de documentos docentes de la unidad de Bioestadística y Biomatemáticas del Depto. de Salud Pública y Medicina Social. Facultad Medicina. Universidad de Chile, 1971.

GRAFICO "B"





El trazado aritmético tiene distancias iguales entre las líneas coordinadas. La distancia entre 2 y 5 es la misma que entre 17 y 20, y por lo tanto en el trazado estas distancias tienen igual magnitud.

La progresión aritmética implica la existencia de una diferencia constante 1, 3, 5, 7, 9, 11, ... En este caso, esa diferencia constante entre cualquier par de valores sucesivos es 2. En la progresión aritmética entonces hay una diferencia constante entre los valores sucesivos. La progresión aritmética da una gráfica rectilínea sobre el papel aritmético.

A veces un conjunto de valores sucesivos tienen entre sí una razón constante. En este caso se habla de una progresión geométrica, como es el caso de la siguiente serie de valores sucesivos 3, 9, 27, 81, 243, 729. En este caso la razón constante entre cualquier par de valores sucesivos es 3. Se puede mostrar que en donde hay una razón constante de cambio entre dos valores sucesivos, la diferencia entre los logaritmos de los valores será la misma.

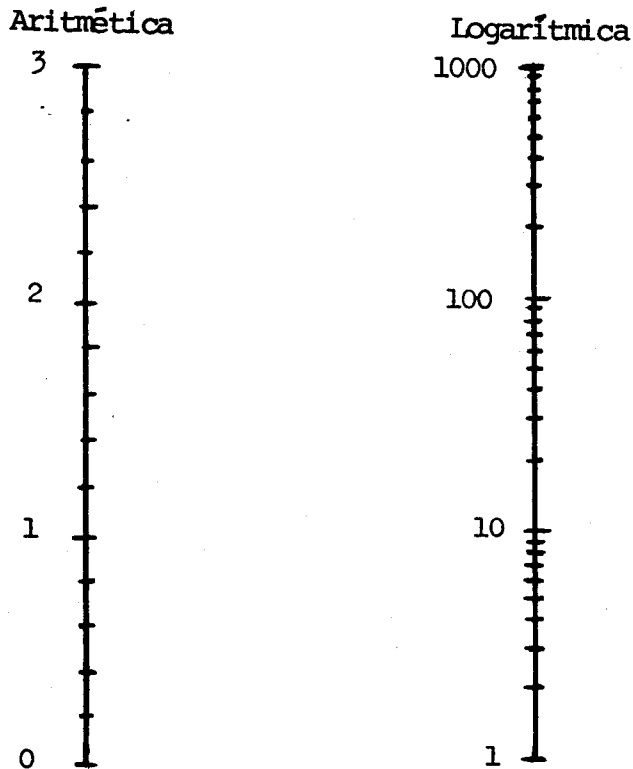
<u>Números</u>	<u>Logaritmos</u>
3	0,4771
9	<u>0,9542</u>
Diferencia	0,4771
27	1,4313
81	<u>1,9084</u>
Diferencia	0,4771

Si se grafican los logaritmos de los valores, en lugar de los valores originales, las diferencias constantes (0,4771) serán iguales a razones constantes (3).

Dado que no siempre se está en condiciones de convertir los valores originales a logaritmos , un procedimiento más conveniente es tomar los datos originales de manera que los logaritmos se puedan graficar directamente a través de una escala especial.

Otro procedimiento sería obtener el logaritmo de cada valor original y graficarlos en una escala aritmética en la forma corriente.

Ejemplos de escalas



Si se utiliza un rayado logarítmico tanto en el eje de abscisas como en el de ordenadas, el papel se llama logarítmico. Si se utiliza escala logarítmica sólo en el eje vertical entonces se habla de papel semilogarítmico.

En el trazado logarítmico de una gráfica no existe línea base cero, de ahí que la gráfica logarítmica no permite la interpretación de curvas en término de distancias con respecto a la línea base como en los gráficos aritméticos. Cuando en un trazado logarítmico se grafica una progresión geométrica, se forma una línea recta, puesto que los logaritmos de una progresión geométrica forman una progresión aritmética. En la escala logarítmica los  aumentos o disminuciones iguales indican cambios iguales en porcentaje.

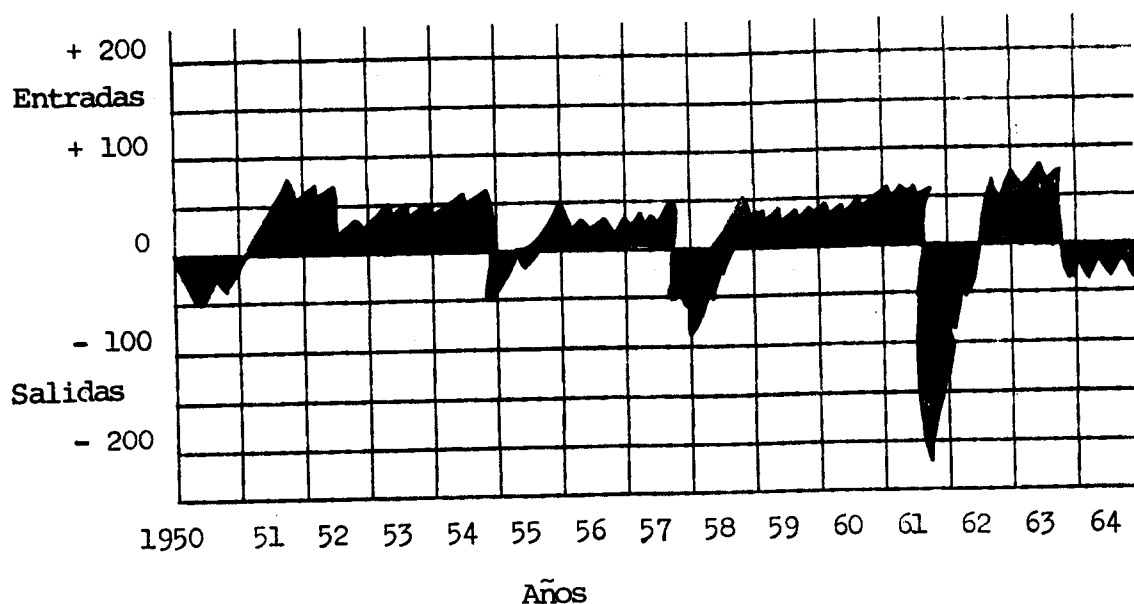
	<u>Cantidades</u>		<u>Logaritmos</u>
	10		1.0000
	20		<u>1.3010</u>
Aumento	100%	Diferencia	0,3010
	<u>Cantidades</u>		<u>Logaritmos</u>
	40		1.6020
	80		<u>1.9030</u>
Aumento	100%	Diferencia	0,3010

Si dos rectas tienen, en una gráfica logarítmica, pendientes iguales entonces la razón de cambio es la misma.

El trazado logarítmico puede ser utilizado cuando se comparan razones de cambio.

También se puede utilizar el trazado logarítmico para comparar las fluctuaciones de dos series que difieren en magnitud. También podría ser que no se consideraran dos series, sino que se estudiara la fluctuación de una sola serie que durante un período ha tenido fluctuaciones alrededor de valores relativamente pequeños y en otro período alrededor de valores mucho más grandes.

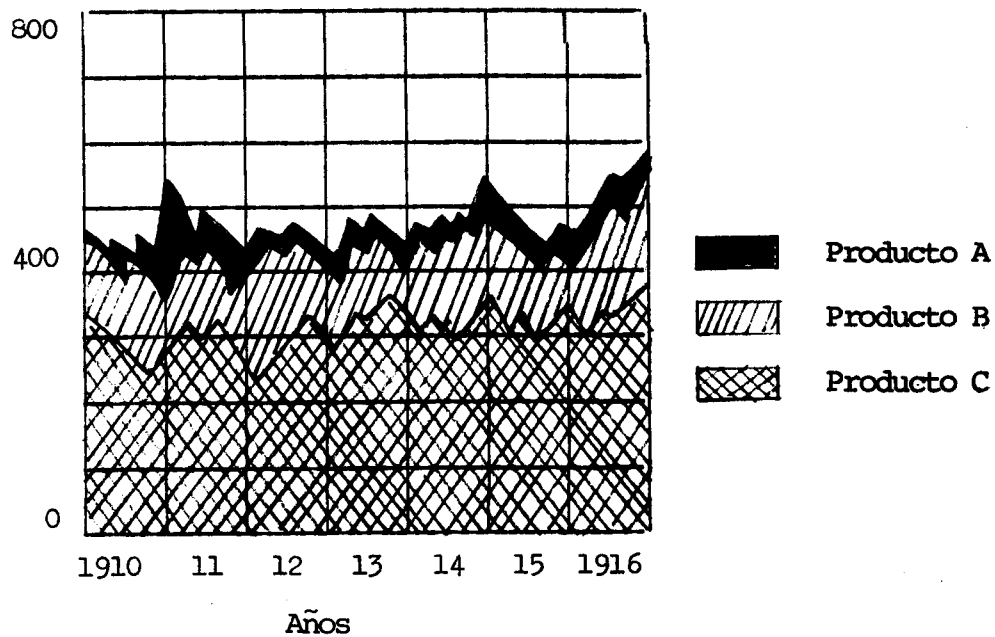
Existen tipos especiales de gráficas lineales. En este curso mencionaremos dos. Una es la gráfica de silueta. Ella expresa desvíos positivos y negativos con respecto a la línea base cero. Se debe llenar el área entre la curva y la línea base.



La otra forma es la gráfica de banda que sirve para expresar variaciones de partes componentes, así como el total. Primero se representa la variación de la parte componente mayor del total. En seguida, la siguiente parte componente se suma al primer componente y se grafica el re

sultado. El proceso acumulativo continúa hasta incluir todos los componentes. Las fluctuaciones de la línea superior re presentan las variaciones del total.

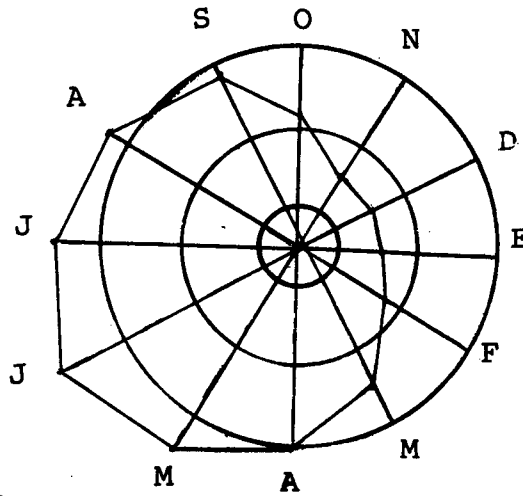
Producción en t.



18. Gráfico polar

Este gráfico es utilizado especialmente para representar variaciones mensuales de un carácter. En este gráfico el ángulo es constante y fluctúa la coordenada lineal de acuerdo con el valor del mes correspondiente . Posteriormente se ligan los extremos de esas coordenadas dano

do el contorno del gráfico que contrasta con el círculo trazado como base, cuyo radio es igual al promedio de los valores de la serie.



Cuando la frecuencia mensual es grande se acostumbra dividirla por 10 o por 100 u otro número que permita llegar a una medida gráfica adecuada.

### 19. Estereogramas

Los gráficos sólidos o de volumen, consistente en formas geométricas (cubos, esferas, cilindros) son confeccionados para ilustrar comparaciones de magnitudes a través de volúmenes en las figuras. No se juzga la altura o longitud sino que el volumen. No es una forma de comparación precisa y sólo se usa para ilustraciones generales en ausencia de otro tipo de gráficos.

En 1960, la producción de un producto X era de 10 millones de toneladas. En 1965, la producción había subido a 16 millones de toneladas y en 1970, a 24 millones de toneladas.

Si la producción de 1960 se representa por un cubo de 2cm por dimensión, el volumen de ese cubo será de  $8 \text{ cm}^3$ , que representa 10 millones de toneladas.

El volumen para 1965 y de 1970 se obtiene :

$$\left. \begin{array}{l} \frac{16}{10} = 1,6 \\ \frac{24}{10} = 2,4 \end{array} \right\} \text{ del volumen del cubo de 1960}$$

Para 1965 :

$$1,6 \times 8 \text{ cm}^3 = 12,80 \text{ cm}^3 \quad \sqrt[3]{12,80} = 2,3 \text{ cm}$$

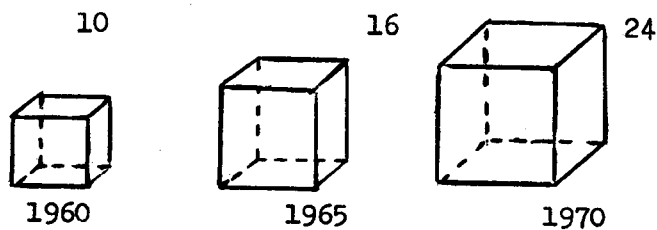
En 1965 el cubo debe tener 2,3 cm por dimensión.

Para 1970 :

$$2,4 \times 8 \text{ cm}^3 = 19,20 \text{ cm}^3 \quad \sqrt[3]{19,20} = 2,67 \text{ cm}$$

En 1970 el cubo debe tener 2,67 cm por dimensión.

MILLONES DE TONELADAS



20. Pictograma

Representación gráfica que utiliza figuras que simbolizan la serie estadística estudiada. Estas figuras pueden ser de variadas dimensiones. Una vez que se ha decidido utilizar una figura determinada se debe reproducir en esa presentación la misma figura. Cada figura representa una cantidad. Si la figura de un bovino representa 100.000 bovinos, la existencia de 350.000 bovinos en una localidad se debe representar por 3 figuras de bovinos y la mitad de una.



Generalmente son empleados para presentar series cronológicas. Existen algunas reglas que es conveniente considerar en la construcción de pictogramas estadísticos. Ellas son:

- a) Los símbolos alusivos deben explicarse por sí mismos.
- b) Las cantidades mayores se indican con un número mayor de símbolos y no por símbolos más grandes.
- c) Sólo se representan cantidades aproximadas y no detalles.

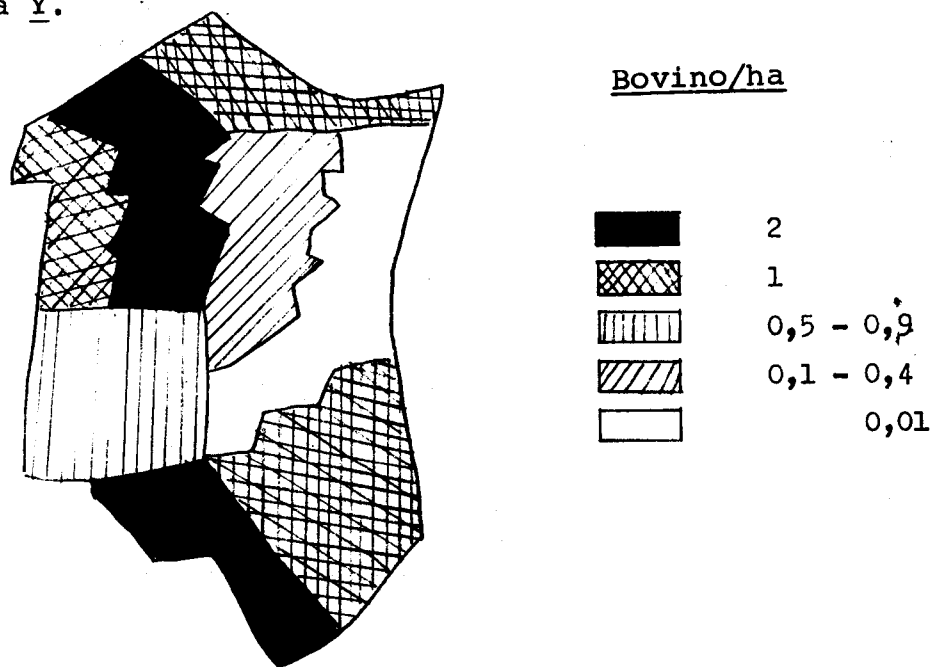


21. Cartogramas estadísticos (mapas)

Presentan información cuantitativa sobre series geográficas.

Los mapas sombreados o achurados, permiten presentar para cada región geográfica que se considera, la magnitud del carácter en estudio. Las variaciones de esta magnitud se expresan gráficamente por diferencia de intensidad de sombreado o de rayado, que puede variar desde el negro sólido hasta el blanco.

Densidad bovina en una región X  
y época Y.



Los mapas achurados o sombreados se utilizan para representar promedios o razones, como la que se da en el ejemplo.

Los mapas punteados, se utilizan más para señalar distribución geográfica de ciertos even

tos, aunque también se pueden utilizar para indicar razones,

## 22. Diagramas de dispersión

Obedecen a los mismos principios de los gráficos lineales, pero en lugar de tener una unidad de observación frente a cada valor de la variable independiente, pueden tener varias unidades de observación. Permiten estudiar series bivariadas. Los ejes deberán comprender aquellos espacios numéricos en que existen valores para ambas variables.

En este caso los puntos no se unen entre sí. Ellos aparecen como una nube de puntos. Si esta nube se asemeja a una línea recta, más estrecha es la relación entre las dos variables estudiadas.

