

**FACULTAD DE TECNOLOGÍA DE LA SALUD**

# **Estadística Descriptiva y Estadística de Salud**

**Autores:**

**Héctor Bayarre**

**Rubén Hersford**

**Maritza Oliva**

## Introducción

Desde los comienzos de la civilización han existido formas sencillas de estadística, pues ya se utilizaban representaciones gráficas y otros símbolos en pieles, rocas, maderas y paredes de cuevas para contar el número de personas, animales y ciertas cosas. Hacia el año 3000 a.C., los babilonios usaban pequeñas tablillas de arcilla para recopilar datos en tablas sobre la producción agrícola y de los géneros vendidos o cambiados mediante trueque. Los egipcios analizaban los datos de la población y la renta del país mucho antes de construir las pirámides. Los libros bíblicos *Los Números* y *Las Crónicas* incluyen, en algunas partes, trabajos de estadística: aquel contiene dos censos de la población de Israel, y el otro describe el bienestar material de las diversas tribus judías. En China existían registros numéricos similares con anterioridad al año 2000 a.C. Los griegos clásicos realizaban censos cuya información se utilizaba hacia el 594 a.C. para cobrar impuestos.

El Imperio Romano fue el primer gobierno que recopiló una gran cantidad de datos sobre la población, superficie y renta de todos los territorios bajo su control. Durante la Edad Media sólo se realizaron algunos censos exhaustivos en Europa. Los reyes caloringios Pipino el Breve y Carlomagno ordenaron hacer estudios minuciosos de las propiedades de la Iglesia en los años 758 y 762 respectivamente. Después de la conquista normanda de Inglaterra en 1066, el rey Guillermo I de Inglaterra encargó un censo; la información obtenida con el mismo, llevado a cabo en 1086, se recogió en el *Domesday Book*. El registro de nacimientos y defunciones comenzó en Inglaterra a principios del siglo XVI, y en 1662 apareció el primer estudio estadístico notable de población, titulado *Observations on the London Bills of Mortality* (*Comentarios sobre las partidas de defunción en Londres*). Un estudio similar sobre la tasa de mortalidad en la ciudad de Breslau, en Alemania, realizado en 1691, fue utilizado por el astrónomo inglés Edmund Halley como base para la primera tabla de mortalidad. En el siglo XIX, con la generalización del método científico para estudiar todos los fenómenos de las ciencias naturales y sociales, los investigadores aceptaron la necesidad de reducir la información a valores numéricos para evitar la ambigüedad de las descripciones verbales.

Empero, la Estadística como ciencia germina en épocas más recientes, a partir del siglo XVII, cuando surgieron de forma simultánea tres escuelas:

- *La administrativa*, alemana, que considera problemas de información al Estado, cuyos principales exponentes fueron Vito de Seckendorff (1626-1689), Hermann Coring (1600-1689) y Godofredo de Achenwall († 1772).
- *La probabilística*, de origen italiano pero devenida francesa sustancialmente, con figuras como Blaise Pascal (1623-1662), Pierre de Fermat (1601-1655), Pierre Simon, marqués de Laplace (1749-1827), Simeón Denis Poisson (1781-1840), los Bernoulli (Jean, Jacques y Daniel), y el alemán Carl Friedrich Gauss (1777-1855). Considera problemas relacionados con el azar.
- *La demográfica*, inglesa, que considera problemas actuariales, encabezada por Petty, Halley, King, Davenant y Graunt.

Posteriormente, continuó desarrollándose la escuela probabilística, gracias a los trabajos de rusos como Tchebichev, Tchuprov, Markov, Kolmogorov, y franceses como Borel, Levy y Fréchet. No obstante, es en los albores del siglo XIX cuando una segunda escuela inglesa, preocupada por problemas de agricultura y biometría, sienta los cimientos de la ciencia actualmente llamada Estadística, guiada por grandes de la talla de Sir Francis Galton (1822-1911), Carl Pearson (1857-1936), Sir Ronald Fisher (1890-1962), y Gosset (*Student*). Merecen mención los integrantes de la *escuela escandinava* Gram, Thiele, Cramer; de la *norteamericana* Hotelling, Wilks, Wald, Neyman, Hoel Mood; y el indio Mahalanobis, entre muchos otros.

En nuestros días, la estadística se ha convertido en un método efectivo para describir con exactitud los valores de datos económicos, políticos, sociales, psicológicos, biológicos y físicos, y sirve como herramienta para relacionar y analizar dichos datos. El trabajo del experto estadístico no consiste ya sólo en reunir y tabular los datos, sino sobre todo en el proceso de *interpretación* de esa información.

En esta sección te presentamos un conjunto de definiciones y procedimientos empleadas por esta ciencia que te ayudarán a entender los contenidos de temas posteriores.

## 1.1 Estadísticas y Estadística

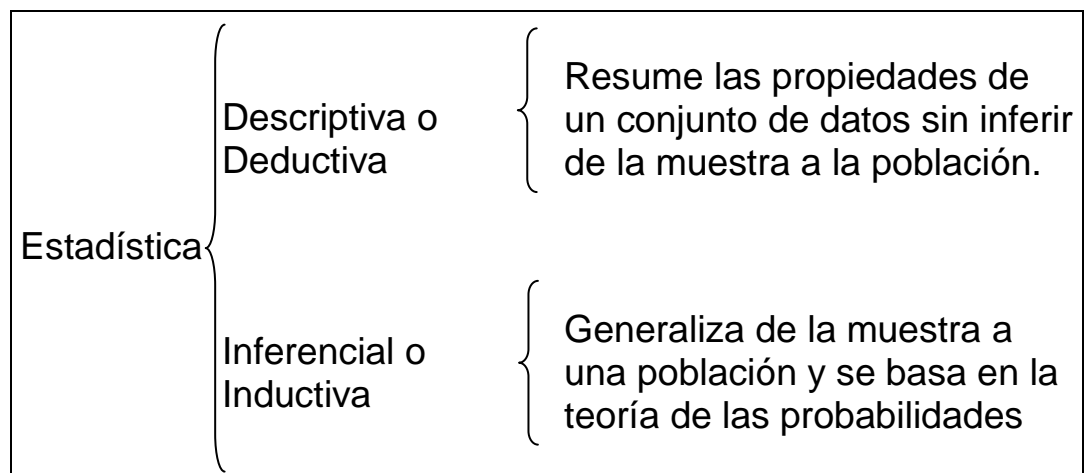
Un primer aspecto que te debe quedar claro es la diferencia entre los términos estadísticas y Estadística. No está muy claro el origen etimológico de la palabra *estadística*, pues algunos la derivan del griego *statera* (balanza), otros del latín *status* (posición, estado, situación), mientras hay quienes afirman que proviene del alemán *staat* (estado, situación). Cualquiera que fuese su fuente, el hecho es que el primer término se utiliza, en este ámbito, para referirse a datos numéricos, en tanto que el segundo se refiere a:



*La **Estadística** es la ciencia encargada de recolectar, organizar, presentar, analizar e interpretar datos numéricos.*

Como habrás notado, la diferencia es sustancial. Aunque, lamentablemente, este último término con mucha frecuencia queda reducido —en opinión de muchos investigadores y profesionales, entre otros— a una simple “herramienta” de la cual se valen otras ciencias para resolver problemas que tienen implícito la medición.

La Estadística suele clasificarse en dos grupos, como verás en el siguiente cuadro:



Por ahora nos limitaremos a abordar la Estadística Descriptiva, quedando la Inferencial para momentos posteriores.

## 1.2 Variables. Definición y clasificación

En la investigación empírica el término **variable** es de uso común, como verás más adelante. Si bien son varias las definiciones que se le han dado, en este contexto la entendemos como:



Una **variable** es *una característica susceptible de ser medida en las unidades de análisis<sup>1</sup> que se estudian, que toma diferentes valores o grados de intensidad, en dependencia de cuál sea la unidad medida.*

Probablemente, estarás murmurando para tus adentros: ¡¿Qué dirá esa definición?! Verás qué sencillo resulta.

Toda característica inherente a los objetos y fenómenos que nos rodean puede ser una variable: los colores de las cosas, la estatura de nosotros, la altura de las edificaciones, el volumen de los recipientes, el sexo, la cantidad de autos que pasan la noche en el parqueo de la esquina, o los países ganadores en las Copas Mundiales de Boxeo; en fin, resultaría interminable la lista. Pero, observa que no son variables por el simple hecho de ser características, sino porque pueden asumir valores diferentes en dependencia de qué objeto o fenómeno midas.

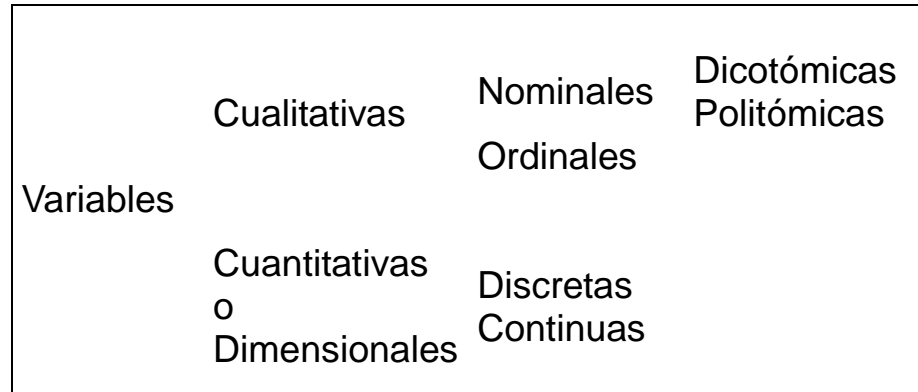
Mira, nosotros los humanos tenemos (todos) una cualidad biológica inherente a cada persona que es el sexo. Ahora bien, en dependencia de a quien escogamos para indagar el sexo, así será el resultado de nuestra pesquisa: nos encontraremos con féminas y hombres indistintamente. ¿Comprendido? Claro que sí.

Prosigamos. Supón que unos investigadores de cierta área de salud cubana investigan la presencia de discapacidad física en ancianos de una comunidad. Para ello, miden en cada anciano escogido las variables sexo, raza, sensación de inactividad, número de consultas médicas recibidas, y la edad, entre muchas otras.

Vamos a analizar minuciosamente cada variable mencionada. ¿Pueden ser medidas de la misma forma? ¿Pueden ser tratadas de igual manera? No, definitivamente no. Para analizar algunas, necesitaremos valores numéricos, y para analizar otras, no los necesitaremos. Estas últimas describen cualidades de los objetos, o sea, atributos que no pueden medirse numéricamente (v.g. el

<sup>1</sup> Aquellos individuos, objetos o fenómenos de los cuales se obtendrá directamente la información del estudio son las **unidades de análisis**.

color, el olor, el sexo, si Juan es o no un alcohólico, entre muchas más), y reciben el nombre de **variables cualitativas**. Por su parte, las otras variables, que sí pueden medirse numéricamente (como el peso, la edad, la talla, el número de hijos, entre otras), se denominan **variables cuantitativas**. Precisamente, las variables se agrupan en torno a estas dos grandes vertientes. Te mostramos una clasificación más detallada a continuación:



Dentro de las variables cualitativas se distinguen las dos categorías señaladas: **nominales** y **ordinales**, las cuales veremos seguidamente.

Para ciertas diferencias de calidad no es necesario que cuentes si deseas distinguir entre una y otra. Nos referimos, por ejemplo, a las diferencias entre un niño y una niña; o entre un auto negro, uno blanco y otro púrpura. Ante *variables que representen atributos que no lleven implícitas diferencias de magnitud*, como es el caso del sexo, el color, el olor, o alguna parecida, estás enfrentándote a **variables cualitativas nominales**.

Estas variables tienen una peculiaridad: algunas *pueden tomar solamente dos valores* (el sexo puede ser femenino o masculino —o, al menos, así lo acepta la mayoría—; si un individuo es fumador o no; si un trabajador del sector Salud está satisfecho con el trabajo que realiza o no; por citar algunas), mientras otras *pueden tomar más de dos valores*, es el caso de la ocupación (estudiante, trabajador, jubilado, desocupado), el estado conyugal (soltero, casado, viudo), entre otras. En la primera situación estás frente a una **dicotomía**, y las variables que así se comportan son **variables dicotómicas**; en la otra situación se trata de una **politomía**, llamándose las que tienen estos rasgos **variables politómicas**.

En la práctica diaria te encuentras variables que, aun cuando describen cualidades de las unidades de análisis, dichas

cualidades pueden ser dispuestas siguiendo un orden lógico. Abundan los ejemplos bastante cercanos a nosotros los médicos: el estado del paciente, que puede ir desde *de cuidado* hasta *crítico extremo*, pasando por niveles intermedios; el estadio de una enfermedad (leve, moderada o severa). Se trata de **variables cualitativas ordinales**<sup>2</sup>, que *sí llevan implícito diferencias de magnitud*, lo que posibilita ordenarlas y así indicar la posición que ocupan dentro del grupo. Ellas brindan mayor información que sus buenas amigas las nominales, aunque las diferencias no puedan medirse numéricamente.

Por otra parte, también tenemos a las variables cuantitativas. Al trabajar con ellas, debes distinguir entre discretas y continuas; lo cual aprenderás en las próximas líneas.

Una variable cuantitativa puede ser el número de hijos de ciertos matrimonios. Observa que puede tomar diferentes valores en cada pareja encuestada, pues seguramente habrá quienes no han tenido descendencia, otros tendrán un bebé, e incluso podrás encontrarte una parejita muy dulce con siete u ocho chiquillos. Una variable como ésta toma valores de tal suerte que *entre dos consecutivos no existen posibilidades prácticas ni teóricas de que haya valores intermedios*. Esto queda claro, nadie puede tener 1½ hijos (tiene un hijo o tiene dos), ni puedes tener en tu casa 4.87 camas (tienes cuatro camas o tienes cinco). Habiendo esclarecido este punto, te diremos entonces que acabas de conocer a las **variables cuantitativas discretas o discontinuas**. Un recurso para identificarlas puede ser el siguiente: por lo general estas variables resultan del conteo: número de pacientes, número de casas en mal estado, y así por el estilo.

Por último, hablemos de las **variables cuantitativas continuas**. Estas son variables cuantitativas que, si toman dos valores posibles, *todos los valores intermedios también lo son*. Por ejemplo, al medir a un individuo en un estudio, los resultados podrían ser los siguientes (en dependencia de la exactitud requerida y del instrumento utilizado): 166.0 cm, 166.0002 cm, 166.000211111 cm, etc. Desde luego, todos son posibles. Digamos esto de otro modo: estas variables aceptan —al menos teóricamente— cualquier valor fraccionario, independientemente de que en muchas ocasiones *sólo se manejen los valores enteros*.

---

<sup>2</sup> Algunos las llaman **cuasicuantitativas**.

Siempre debes tener esto último muy en cuenta, porque con frecuencia tratamos variables cuantitativas continuas como si no lo fuesen. Un ejemplo cotidiano entre tantos: la edad es una variable continua, pero nadie suele decir su edad exacta (26 años, 8 meses, 1 semana, 4 días, etc.); todo lo contrario, la expresamos en números redondos (26 años y ya está). Te alertamos en este punto pues debes estar seguro de la naturaleza de las variables antes de manipularlas; independientemente del tratamiento que les des, ellas no perderán su esencia.

### 1.3 Universo y muestra

Los vocablos que dan nombre a este apartado desempeñan un papel loable en la investigación, pues de su correcta determinación dependerá, en parte, el éxito de la misma. Por **universo o población** entendemos *un grupo, casi siempre numeroso, compuesto frecuentemente —pero no necesariamente— por personas, que tienen en común al menos una característica, susceptible de ser investigada*. De ella se extraen las muestras necesarias para su estudio.

Por **muestra** entendemos *el subgrupo de una población extraído por un investigador para extraer conclusiones de la misma, o para realizar estimaciones sobre ella*. La muestra se obtiene mediante el **muestreo**, importante técnica estadística que no abordamos en este curso por no ser objetivo del mismo.

Estos términos tienen la peculiaridad de no ser absolutos, viéndolos desde el punto de vista de que en un momento el mismo conjunto puede ser población, y en otro puede ser muestra.

Veamos algunos ejemplos, los cuales te ilustrarán lo antedicho:

1. Para determinar si el medicamento X producido en la Fábrica de Medicamentos Y tiene la calidad adecuada, se toman 100 tabletas al azar de la producción de una semana. Aquí la población es el total de tabletas del medicamento X producidas por la fábrica, y la muestra está formada por las 100 tabletas que se estudiaron.
2. Un grupo de investigadores desea estudiar el comportamiento del síndrome anémico en el área que atiende el policlínico Sur de la provincia Guantánamo, para ello decide tomar a los habitantes de los consultorios 66, 50, 17 y 5 a fin de realizarles los exámenes pertinentes. En este caso, la población está



formada por el total de personas que atiende el policlínico, y la muestra por las personas que atienden los cuatro consultorios escogidos.

3. Unos investigadores se proponen estudiar la actitud de las jubiladas de cierto consultorio ante el estrés generado por las tareas del hogar, por lo que estudiaron a 10 de las 37 jubiladas existentes. Obviamente, la población está constituida por las 37 jubiladas, y la muestra por las 10 señoras estudiadas.
4. Los mismos investigadores desean estudiar el fenómeno en el área atendida por el policlínico, de ahí que estudiaran a las jubiladas de cinco consultorios. Ahora, la población es el total de jubiladas que atiende el policlínico, y la muestra las jubiladas de aquellos consultorios escogidos.

La determinación correcta de la muestra entraña el cálculo del tamaño muestral adecuado para lograr resultados fiables, y la selección del método de muestreo apropiado. Te recomendamos que te dirijas al Bioestadístico más cercano para que te brinde su asesoría al respecto.

#### 1.4 Escalas de clasificación de las variables

A menudo sucede que, con el objetivo de facilitar el trabajo, el investigador toma la decisión de agrupar los datos de manera tal que convenga a sus intereses. Ello se logra mediante la construcción de una escala de clasificación, o simplemente escala.

Para lograr una buena escala, toma en cuenta que la misma debe:



Una **escala** debe cumplir los siguientes requisitos:

- Ser **exhaustiva**, o sea, que permita clasificar a todas las unidades de análisis.
- Las clases o categorías que la integran deben ser mutuamente **excluyentes**, esto es, que una unidad de análisis puede estar en una y solo una categoría.

Las escalas pueden ser *cuantitativas* o *cualitativas*, en dependencia de si sus categorías pueden ser numéricamente medidas o no. Hagamos un alto en este emocionante viaje para detallar algunos aspectos.



### Escalas

Cualitativas	- nominales - ordinales
Cuantitativas	- de intervalo - de razón o de proporción

Si así lo deseas, puedes construir una **escala nominal** a cualquier variable (dentro de lo razonablemente lógico, por supuesto), no importa cual sea la naturaleza de dicha variable. Claro que, al hacerlo, no puedes pasar por alto algo muy importante: si estás representando una variable de naturaleza no nominal, sencillamente estás perdiendo información que pudiera resultarte valiosa; es más, en realidad no estás midiendo cosa alguna, *sólo estás clasificando las unidades de análisis en categorías o grupos*. Mira el siguiente ejemplo:

Tabla 1. Distribución de recién nacidos según peso. Hospital "Clodomira Acosta". Febrero, 1999.

Peso	Número	Porcentaje
<b>No normopeso</b>	62	62.0
<b>Normopeso</b>	38	38.0
<b>Total</b>	<b>100</b>	<b>100.0</b>

Fuente: Registro de nacimientos. Hospital "Clodomira Acosta".

Aquí vemos cómo una variable continua, el peso, fue tratada como nominal dicotómica. Es una forma útil de manejar el dato, ya que permite una rápida evaluación de la situación, pero adolece de falta de información, pues si necesitas más datos acerca del peso de esos infantes, no te quedaría más remedio que remitirte a la fuente.

En relación con la utilización cada vez más creciente de los softwares para el tratamiento estadístico de la información, estamos en el deber de alertarte ante un error bastante frecuente

por desgracia. Se trata del mal uso de la *codificación*, un recurso que brindan los paquetes estadísticos, consistente en la asignación de códigos numéricos a las variables en estudio, a fin de facilitar su manejo. Por ejemplo, el sexo puede tratarse como 1 para el femenino y 2 para el masculino (o viceversa), pero esto este tratamiento no le confiere valor cuantitativo a la variable, puesto que es cualitativa; por ende, no se le pueden aplicar procedimientos matemáticos propios de variables cuantitativas<sup>3</sup>.

Otra forma de manipular tus datos es construyendo una **escala ordinal**. En esta situación, estás creando rangos<sup>4</sup> al asignar una posición dentro del grupo al dato en cuestión. Aquí, los datos de una categoría no son simplemente diferentes a otros en otra categoría, sino que guardan una relación entre sí, relaciones que pueden ser «mayor/menor/más complejo que», por mencionar algunas. Sin embargo, no conoces las distancias entre los elementos (el elemento A está a 2.4 unidades del B, por ejemplo), sólo te limitas a manejar las relaciones antedichas.

Puedes representar datos cuantitativos y ordinales en esta escala, pero nunca datos nominales. Tiene la ventaja de que brinda mayor información que la escala nominal. Veamos los datos subsiguientes:

Tabla 2. Distribución de recién nacidos según peso al nacer. Hospital Municipal “Mariano Pérez”, Municipio Bartolomé Masó. Primer trimestre, 1999.

<b>Peso al nacer</b>	<b>Número</b>	<b>Porcentaje</b>
<b>Bajo peso</b>	60	60.0
<b>Normopeso</b>	38	38.0
<b>Sobrepeso</b>	2	2.0
<b>Total</b>	100	100.0

Fuente: Registro de nacimientos. Hospital “Mariano Pérez”.

Indiscutiblemente, ya cuentas con más información que en el ejemplo anterior. Antes sólo sabías que la mayor parte de los recién nacidos no tenían un peso considerado “normal”; ahora sabes que nacieron niños con pesos superiores e inferiores a los estándares, e incluso no es tan difícil imaginar que algo no anda bien con la atención materno-infantil, a juzgar por los datos. Pero,

<sup>3</sup> Puedes consultar las técnicas estadísticas apropiadas para cada tipo de variables en los Temas 3 y 4.

<sup>4</sup> O sea, estás creando un orden jerárquico, ya sea ascendente o descendente.

si necesitaras más información por alguna causa, entonces estarías en el mismo lugar que en el ejemplo precedente.

Para construir una **escala de intervalo**, debes mantener las características de una escala ordinal, sólo que en la presente conoces las distancias entre dos números de la escala. Ella se caracteriza por una unidad de medida común y constante que asigna un número real, pero tanto la unidad de medida como el punto cero son arbitrarios, un ejemplo lo constituye la temperatura, que utiliza dos escalas (Celsius y Fahrenheit) en las que el punto cero difiere entre sí, siendo arbitrario en ambos casos.

Si una escala posee las características antedichas, pero con la diferencia de que se origine en un cero real, entonces la misma es una **escala de razón**. La razón entre dos puntos cualesquiera no depende de la unidad de medida. Escalas construidas al peso, la talla, u otra variable continua en su clara esencia, constituyen ejemplos de este tipo de escala. De hecho, muchas de las escalas consignadas como de intervalo, en realidad son de razón.

Otra forma de manejar las variables cuantitativas<sup>5</sup> es mediante la construcción de **escalas discretas o discontinuas y escalas continuas**, guardando consonancia con las características de las variables de igual nombre. Sólo debes tener en cuenta que puedes manejar una variable continua en escala continua o discreta, pero una discreta no puede ser tratada en escala continua.

Antes de continuar, es necesario que fijes que no siempre la naturaleza de la variable coincide con el tipo de escala que le construirás.

La construcción de **escalas cuantitativas de intervalo** requiere que la analicemos con esmero, por ser de mucha utilidad y necesitar ciertos requisitos adicionales.

Ante todo, debes saber que esta escala de clasificación está compuesta por varias divisiones ordenadas llamadas **intervalos de clase** (IC), los cuales están delimitados por **límites de clase**, que son los valores mayor y menor que los enmarcan.

Las escalas pueden ser *cerradas* o *abiertas*, en virtud de que todos sus intervalos posean o no sus límites de clase. Algunos

---

<sup>5</sup> Quizás esta sea la forma más ampliamente utilizada.

autores denominan *semiabiertas* o *semicerradas* a aquellas escalas que omiten uno de los límites, ya sea el límite inferior del primer intervalo de clase, o el límite superior del último intervalo; reservando el término *abierto* para las escalas que omiten ambos límites.

Para ilustrar las ideas planteadas, te mostramos una escala cuantitativa de intervalos abierta y los intervalos utilizados:

Edad	
< 15 años	← Intervalo de clase abierto
15 – 19 años	
20 – 24 años	
25 y más años	← Intervalo de clase abierto

Ilustración 1. Escala cuantitativa abierta

Los intervalos pueden o no tener igual amplitud, aunque se prefiere lo primero, ya que facilita enormemente el trabajo posterior con los datos.

Otro término que debes manejar es el de **límite real (LR)**. En realidad son dos límites reales para cada intervalo: el **superior (LRS)** y el **inferior (LRI)**, con la particularidad de que el LRS de un intervalo es, a la vez, el LRI del intervalo siguiente. Si bien los LR no se utilizan para la construcción de una escala definitiva, son útiles en la determinación de la llamada amplitud del intervalo y en la construcción de algunos gráficos, como verás con posterioridad. Se calculan mediante la semisuma<sup>6</sup> de los límites de clase superior e inferior de intervalos contiguos.

La **amplitud o recorrido (A)** de un intervalo de clase es la longitud de éste. Su cálculo puede hacerse de distintas maneras:

1. La más utilizada consiste en hallar la diferencia entre los límites reales del intervalo en cuestión. En el ejemplo anterior, los límites reales del segundo IC son 14.5 y 19.5, por lo que la amplitud es 5.

<sup>6</sup> Esto es, sumar ambos números y dividir el resultado por dos. Por ejemplo, la semisuma de 5 y 6 es  $(5 + 6) \div 2 = 5.5$ . ¿Claro?

2. Otra forma estriba en hallar la diferencia de los límites de clase del intervalo de referencia y, luego, adicionarle una unidad al resultado obtenido. Así, para el segundo IC del ejemplo anterior,  $A = (19 - 15) + 1 = 5$ .
3. Por último, puedes calcular A contando los números enteros que se encuentran entre los valores límites, incluyendo éstos. Así, para el segundo IC del ejemplo anterior, la amplitud sería el conteo de 15, 16, 17, 18 y 19, es decir,  $A = 5$ .

Obviamente, los casos 2 y 3 sólo son válidos cuando los límites de las escalas son números enteros.

Por otra parte, la **marca de clase** de un IC es el punto medio de dicho intervalo, que se computa mediante la semisuma de los límites de clase del intervalo referido. Por ejemplo, la marca de clase del tercer intervalo es  $MC = (20 + 24) / 2 = 22$ .

Para construir una escala con intervalos de clase de igual amplitud, debes seguir los siguientes pasos:



1. **Determina el recorrido de la serie (R).** Esto lo logras restando el valor mínimo al máximo.
2. **Fija el número mínimo de intervalos de clase deseado.** Esta decisión va por ti, lo determinarás en dependencia de tus necesidades. No siempre este es el número definitivo de ICs, ocasionalmente requerirás un IC adicional.
3. **Calcula la amplitud (A) de los intervalos.** Para ello, divide el recorrido que obtuviste en el paso 1 por el número que fijaste en el paso anterior.
4. **Delimita los límites inferiores (LI) de los intervalos.** Partiendo del valor mínimo de la serie, añádele la amplitud y tendrás el LI del intervalo siguiente, a este le sumas la amplitud y tendrás el subsiguiente, y así hasta llegar al último LI de la escala.
5. **Delimita los límites superiores (LS).** Lo harás sustrayendo una unidad al LI siguiente. En el caso del LS del último intervalo, lo obtendrás sumándole la amplitud al último LI, y luego restando al resultado una unidad.

Nota: te sugerimos que redondees en el paso 3 para convertir la amplitud en un número redondo, lo que te facilitará la construcción de la escala.

A fin de fijar ideas, construyamos juntos una escala. Supón que tienes una lista con los pesos (en libras) de 20 adolescentes, y deseas agruparlos en una escala cuantitativa con intervalos de igual amplitud.

<b>Peso</b>	<b>Peso</b>	<b>Peso</b>	<b>Peso</b>
1. 160,00	6. 170,54	11. 166,00	16. 150,00
2. 160,36	7. 160,20	12. 156,70	17. 151,78
3. 158,20	8. 163,20	13. 154,50	18. 152,00
4. 174,00	9. 165,80	14. 155,00	19. 154,80
5. 170,00	10. 165,90	15. 155,90	20. 156,70

Para visualizar mejor el recorrido, comencemos por ordenar los pesos:

<b>Peso</b>	<b>Peso</b>	<b>Peso</b>	<b>Peso</b>
1. 150,00	6. 155,00	11. 160,00	16. 165,90
2. 151,78	7. 155,90	12. 160,20	17. 166,00
3. 152,00	8. 156,70	13. 160,36	18. 170,00
4. 154,50	9. 156,70	14. 163,20	19. 170,54
5. 154,80	10. 158,20	15. 165,80	20. 174,00

Como puedes ver, el valor mínimo es 150.00, y el máximo es 174.00, de tal suerte que la escala se construiría de la siguiente manera:

1. El recorrido de la serie es  $R = 174.00 - 150.00 = 24$ .
2. Supongamos que deseas como mínimo 4 intervalos de clase.
3. La amplitud que tendrán los intervalos es  $A = 24 \div 4 = 6$ .
4. Límites inferiores:

<u>IC</u>	<u>LI</u> s
1.	<b>150</b>
2.	$150 + 6 = 156$
3.	$156 + 6 = 162$
4.	$162 + 6 = 168$

### 5. Límites superiores:

#### IC LIs

1.  $156 - 1 = \mathbf{155}$

2.  $162 - 1 = \mathbf{161}$

3.  $168 - 1 = \mathbf{167}$

4.  $174 - 1 = \mathbf{173}$

Hagamos un alto necesario. Con lo que hemos hecho hasta ahora, la escala será 150-155, 156-161, 162-167, 168-173, completando así los cuatro intervalos deseados. Mas, como puedes comprobar, en los datos existe un valor que supera 173, de ahí que sea necesario añadir un intervalo de clase al número predicho con el objetivo de lograr la exhaustividad de la escala:

$$\text{Límite inferior del quinto IC: } 168 + 6 = \mathbf{174}$$

$$\text{Límite superior del último intervalo: } 174 + 6 - 1 = \mathbf{179}$$

Finalmente, la escala que necesitabas construir es la siguiente:

$$\mathbf{150 - 155}$$

$$\mathbf{156 - 161}$$

$$\mathbf{162 - 167}$$

$$\mathbf{168 - 173}$$

$$\mathbf{174 - 179}$$

Observa que esta escala cumple con los requisitos planteados al inicio de este epígrafe. En otras partes del curso verás algunos usos específicos de las escalas.

### 1.5 Distribuciones de frecuencias absolutas, relativas y acumuladas

En la mayoría de las situaciones prácticas de la investigación, se hace imprescindible echar mano a un recurso sumamente útil: la agrupación de las unidades de análisis en dependencia de su frecuencia de aparición, partiendo de las escalas de clasificación que ya viste con anterioridad.

Supón que, a petición del delegado del Poder Popular, estás registrando el estado constructivo de 10 de las casas del área que atiendes, para lo cual te riges por ciertas reglas que te permiten clasificarlo en bueno, regular y malo. Redactas un informe con los resultados que obtuviste, y finalmente entregas esto:



Estado constructivo de las casas  
del Sector 000  
Por: Dr. Ambrosio Fino Delgado

<u>Casa</u>	<u>Estado</u>
01:	<u>constructivo</u>
02:	R
03:	B
04:	M
05:	B
06:	R
07:	M
08:	M
09:	B
10:	B
	R

Cuando el funcionario reciba el solicitado informe, puede que se extrañe un tanto. ¿Por qué? Mira, no es que esté mal lo que le entregaste, la información se ajusta a lo pedido, pero hay algo que no facilita las cosas. ¡Claro! ¡La presentación de los datos! Juzga por ti mismo la diferencia:

Estado constructivo de las casas del  
Sector 000  
Por: Dr. Ambrosio Fino Delgado

<u>Estado constructivo</u>	<u>Casas</u>
Bueno:	4
Regular:	3
Malo:	3

¿No crees que ganaste con esta forma de presentar los datos? Por supuesto, sobre todo porque al **resumir** la información y exponerla de forma asequible, estás garantizando —en gran medida— que un lector abrumado por mucho trabajo dedique un minuto a tu informe, y no lo destine a la gaveta del olvido, entre otras cosas.

Esto que acabas de hacer, o sea, agrupar los datos por frecuencia de aparición, de acuerdo con la escala que más se ajusta a tus necesidades, es lo que se denomina **distribución de frecuencias**, que consiste en:



Una **distribución de frecuencias** es el modo en que se distribuyen las unidades de análisis entre las clases o categorías que conforman la escala de clasificación de la variable en cuestión.

Las distribuciones de frecuencias pueden clasificarse en:



	- Absoluta
<b>Distribuciones de frecuencias</b>	- Relativa
	- Absoluta
<b>Acumulada</b>	- Relativa

Las dos primeras se utilizan cuando se tratan variables en cualquier escala, en tanto que las acumuladas se emplean cuando se estudian variables en una escala cuantitativa o dimensional.

La **frecuencia absoluta** es el resultado de contar los casos u observaciones (o sea, el número de observaciones) que corresponden a cada una de las clases o categorías de la escala de clasificación.

La **frecuencia relativa** es la importancia o peso relativos que tienen las unidades de análisis de una categoría o clase sobre el total de las unidades. Se calcula dividiendo la frecuencia absoluta de la clase en cuestión por el total de observaciones, en cuyo caso obtendrás una proporción. Si multiplicas este resultado por 100, obtendrás un porcentaje.

Las **frecuencias acumuladas** son las frecuencias absolutas o relativas que se acumulan hasta un intervalo de clase dado. Se calculan sumando las frecuencias (absolutas o relativas, en dependencia de la que necesites) hasta la clase deseada. La frecuencia acumulada para el último IC será el total de observaciones, si se tratare de la frecuencia absoluta acumulada; y si fuera el caso de la relativa, entonces será 1 ó 100, en dependencia de si usaste proporción o porcentaje, respectivamente.

A continuación te presentamos mediante un ejemplo cada tipo de distribución de frecuencias.

En un estudio sobre riesgos profesionales en los trabajadores de la industria del cromo en 1995, se encontraron los siguientes datos en una muestra de 897 trabajadores:

Tiempo de trabajo en la industria (años)	Número (frecuencia absoluta)	Frecuencia relativa		Frecuencia acumulada		
		Proporción	%	Absoluta	Proporción	%
< 1 año	41	0.04	4.0	41	0.04	4.0
1 – 3	115	0.13	13.0	156	0.17	17.0
4 – 10	304	0.34	34.0	460	0.51	51.0
> 10	437	0.49	49.0	897	1.000	100.0
<b>Total</b>	<b>897</b>	<b>1.00</b>	<b>100.0</b>	<b>—</b>	<b>—</b>	<b>—</b>

Si observas detenidamente, te podrás dar cuenta que:

- La suma de las frecuencias absolutas es el total de la muestra o población estudiada.
- La suma de las frecuencias relativas es igual a 1 o a 100, en dependencia de que se trate de una proporción o de un porcentaje.
- En las frecuencias acumuladas la última clase o categoría de la variable suma el total, la unidad o el ciento por ciento.

Puede que la suma de los porcentajes no sea 100 exactamente debido a las imprecisiones de la aproximación, de ahí que se admita hasta 99.9 como total; pero debes evitar cifras inferiores ya que denotan errores de cálculo.

## Resumen

En este tema estudiaste que:

1. Es importante distinguir entre los términos **estadísticas**, sinónimo de datos numéricos; y **Estadística**, ciencia que se encarga del manejo de dichos datos, es decir, de la recolección, procesamiento, análisis y presentación de éstos.
2. Dentro de la Estadística se distinguen dos ramas: la **Descriptiva o Deductiva**, que resume las propiedades

de un conjunto de datos sin inferir de la muestra a la población; y la **Inferencial o Inductiva**, que generaliza de la muestra a una población y se basa en la teoría de las probabilidades.

3. Una **variable** es cualquier característica investigada en una población o muestra, que puede asumir diferentes valores o grados de intensidad entre las unidades de análisis. Ellas pueden clasificarse en cualitativas y cuantitativas.
4. La **escala de clasificación** de una variable permite agrupar la información obtenida sobre ésta. En general, se clasifican en cualitativas y cuantitativas.
5. Una **distribución de frecuencias** es el modo en que se distribuyen las unidades de análisis entre las clases o categorías que conforman la escala de clasificación de la variable en cuestión. Entre ellas se encuentran las **absolutas**, las **relativas** y las **acumuladas**. Las dos primeras pueden utilizarse para cualquier tipo de escala, mientras que la última sólo se utiliza para cuantitativas.

## Ejercitación

1. Un grupo de investigadores desea estudiar la satisfacción personal de los médicos de la familia de la provincia Ciudad de La Habana. Para ello, de un total de 6617 médicos que trabajan en ese territorio, estudiarán 600. Desean recoger información sobre la edad, la calificación profesional, el tiempo de graduado, el sexo, el estado civil y la presencia de satisfacción personal.
  - a) Identifique cuál es la población en estudio y cuál es la muestra, en caso de existir.
  - b) Identifique las variables a estudiar y clasifíquelas.
2. A continuación te presentamos algunas variables investigadas en un grupo de estudiantes del 1<sup>er</sup> año de la especialización de Medicina General Integral, con la finalidad de controlar el proceso docente-educativo.
  - Número de guardias mensuales,
  - Porcentaje de asistencia a actividades docentes,
  - Existencia de quejas de la población,
  - Existencia de sanciones,
  - Sexo, y

– Edad.

a) Clasifique cada una de las variables estudiadas.

3. La dirección de un policlínico desea estudiar la nutrición de los adolescentes del área. Una de las variables investigadas es la talla (en centímetros). A continuación te presentamos la talla de 20 jóvenes del estudio mencionado tomados al azar.

1. 168	6. 170	11. 160	16. 175
2. 145	7. 167	12. 159	17. 161
3. 162	8. 163	13. 158	18. 152
4. 155	9. 165	14. 174	19. 147
5. 155	10. 165	15. 171	20. 148

- a) Construye una escala con intervalos de igual amplitud para representar esta información. Te sugerimos que utilices seis intervalos.
- b) Distribuye la información a través de frecuencias absolutas, relativas y acumuladas.
4. A continuación te presentamos la variable temperatura (en grados Celsius) clasificada en una escala de intervalos:

Temperatura

34 - 36

37 - 39

40 - 42

- a) Determina la amplitud del segundo intervalo de clase.
- b) ¿Cuál es la marca de clase del primer intervalo de clase?
- c) ¿Cuál es el límite de clase inferior del tercer intervalo de clase?

## Autoevaluación

1. En 1996, en una investigación sobre la insatisfacción personal en mujeres de edad mediana, pertenecientes a un área de salud del municipio La Lisa (Ciudad de La Habana), del total de mujeres (3006) se estudiaron 600. A todas se les recogió información sobre la presencia de insatisfacción personal según los resultados del instrumento ISP-RELEBA<sup>7</sup>

<sup>7</sup> No te preocupes, sólo es el nombre del instrumento (raro, ¿verdad?).

(autoestima, percepción de salud, proyecto de vida, condicionamiento de género y relación de pareja).

- a) Identifica cuál es la población en estudio y cuál es la muestra, en caso de existir.
- b) Identifica las variables a estudiar y clasifícalas.

2. Las cifras de tensión arterial diastólica de un grupo de pacientes de cierta área de salud aparecen en la siguiente ficha de vaciamiento:

1. 95	6. 75	11. 110	16 80	21. 85
2. 90	7. 100	12. 105	17 80	22. 80
3. 85	8. 75	13. 90	18. 80	23. 100
4. 90	9. 80	14. 80	19. 80	24. 90
5. 80	10. 80	15. 75	20. 80	25. 95

- a) Construye una escala con intervalos (sugerimos siete) de igual amplitud para representar esta información.
- b) Distribuye la información a través de frecuencias absolutas, relativas y acumuladas.

3. De la escala construida en la pregunta anterior, dinos:

- a) Límites reales del segundo intervalo.
- b) Marca de clase del primer intervalo.
- c) Amplitud del tercer intervalo.

## Bibliografía

1. Daniel WW. Bioestadística. Base para el análisis de las ciencias de la salud. 3ª ed. México D.F.:Limusa; 1997.
2. Spiegel MR. Teoría y problemas de Estadística. La Habana:Pueblo y Educación; 1977.
3. Freund J. Estadística elemental moderna. 2ª ed. La Habana:Edición Revolucionaria; 1988.
4. Coolican H. Métodos de investigación y estadística en psicología. México D.F.:El Manual Moderno; 1997.
5. Silva LC. Muestreo para la investigación en salud. Madrid:Díaz de Santos; 1993.
6. Riegelman RK, Hirsch RP. Cómo estudiar un estudio y probar una prueba: lectura crítica de la literatura médica. Washington DC:OPS; 1992. (Pub. Cient. No. 531).

7. Cansado E. Curso de Estadística General. La Habana:Ed. Revolucionaria. 1966.
8. Swaroop S. Estadística Sanitaria. La Habana:Ed. Revolucionaria. 1964.
9. Camel F. Estadísticas médicas y de Salud Pública. La Habana:Pueblo y Educación; 1985.

## Tema 2. Representación Estadística

### Introducción

Cuando investigamos, obtenemos información sobre las variables del estudio. Sin embargo, por lo general es tanto el volumen de información que poseemos, que se necesita utilizar alguna técnica que permita presentarla de forma resumida.

A continuación estudiarás las maneras de representar la información: **la tabla o cuadro estadístico y los gráficos.**

Te sugerimos que prestes mucha atención a este tema, pues a pesar de su sencillez, estos conocimientos se usan con bastante frecuencia, y los desatinos que usualmente se cometen hacen ineficaces el uso de estos recursos que te brinda la Estadística.

#### 2.1 El cuadro estadístico

Es conveniente que sepas, en principio, qué es una **tabla o cuadro estadístico**:



Un **cuadro estadístico** es *un recurso que emplea la Estadística con el fin de presentar información resumida, organizada por filas y columnas.*

El cuadro estadístico tiene la finalidad de representar distribuciones de frecuencias, medidas de resúmenes y series cronológicas.

##### 2.1.1 Partes del cuadro estadístico

Varias son los elementos que integran una tabla estadística. Seguidamente te presentamos cada uno de ellos:





### Partes del Cuadro

- Presentación (Identificación y Título)
- Cuerpo de la tabla
- Fuente
- Notas explicativas

1. **Identificación.** Consiste en otorgar un orden consecutivo a las tablas, comenzando por el número uno, v.g. Tabla<sup>8</sup> 1, Tabla 2, etc.
2. **Título.** Debe ser **completo** y **conciso**. Para ser **completo**, el título debe responder a las preguntas **qué, cómo, dónde** y **cuándo**. Reconozcamos en un ejemplo cada una de estas preguntas:

Tabla 1. Distribución de fallecidos según grupos de edad y sexo. La Habana, 1999.

Un análisis del título anterior permite conocer que:

- Distribución de fallecidos es de **qué** trata la tabla.
- Los grupos de edad y sexo son el **cómo** se midió, es decir, a través de cuales variables.
- La Habana es **dónde** se realizó el estudio.
- 1999 es **cuándo** se realizó el estudio.

Observa que las variables se presentan después del vocablo “según”, aunque alternativamente puedes usar el término “por”.

También es conveniente decirte que, en ocasiones, no es necesario dar respuestas a estas cuatro interrogantes en un título, en cuyo caso sólo deberás responder al qué y al cómo. Ello ocurre cuando estás representando información obtenida en una investigación y, en el informe final o el artículo, ya has consignado

<sup>8</sup> Indistintamente Cuadro, si lo deseas, pero no debieras usarlos alternativamente en un mismo trabajo: uno o el otro.

en algún apartado anterior al de Resultados, dónde y cuándo se realizó el estudio.

El otro elemento, la **concisión**, consiste en escribir justamente lo necesario. Elimina las preposiciones y artículos que no ayuden a la comprensión del título de tu cuadro. No obstante, no sacrifiques el mensaje redactando un título “telegráfico”, no seas lacónico, pues dificultarías la comprensión de lo que quieres decir.

3. **Cuerpo de la Tabla.** Es el cuadro en sí mismo, formado por espacios llamados celdas, las cuales se vertebran en filas y columnas, por ejemplo:

<b>Columna matriz</b>	<b>Fila de encabezamiento</b>	<b>Total</b>
XX	XX	XX
XX	XX	XX
XX	XX	XX
<b>Total</b>	<b>XX</b>	<b>XX</b>

La **columna matriz** se utiliza para consignar la variable con su escala de clasificación. En caso de que el cuadro represente más de una variable, por la columna matriz representarás la que tenga más clases o categorías o la que constituye la causa, en estudios de causalidad.

En la **fila de encabezamiento** se presentan las distribuciones de frecuencias, las medidas de resúmenes o la otra variable. La fila y columna últimas se dedican a los **totales**.

Las tablas estadísticas suelen clasificarse según el número de variables que representan en:



**Tablas**

- Unidimensionales:** una variable.
- Bidimensionales:** dos variables.
- Multidimensionales:** tres o más variables.

Deben ser autoexplicativas, o sea, que se expliquen por sí mismas, por lo que debes evitar presentar demasiada información

en ellas en aras de ganar claridad. En general, como forma de presentación se utilizan cuadros uni y bidimensionales, reservándose el uso de los multidimensionales para fines de trabajo.

4. **Fuente.** Se refiere al documento<sup>9</sup> de donde se extrajo la información presentada. Por lo general, las fuentes de información se clasifican en:

<b>Fuentes</b>	<b>Primaria:</b> aquella de la que el investigador obtiene directamente la información utilizando diversas técnicas y métodos, v.g. la encuesta.
	<b>Secundaria:</b> aquella que existe independientemente del estudio y el investigador sólo la utiliza, v.g. el Registro de Nacimientos, las historias clínicas.

Resulta válido y oportuno aclarar que en la tabla sólo consignarás **Fuentes Secundarias**.

Recuerda que una fuente es un *documento*. Frecuentemente esto se olvida, y se consignan erróneamente como fuentes algunos locales, departamentos, entre otros, como la Oficina Nacional de Estadísticas, el Archivo del policlínico, etc.

5. **Notas explicativas o aclaratorias.** Se utilizan cuando se desea aclarar algo, por lo general del título o del cuerpo de la tabla.

### 2.1.2 Errores más frecuentes

No son pocos los errores que se cometen, voluntaria o involuntariamente, en la confección de los cuadros estadísticos. A continuación te presentamos una lista de los más comunes, que podrás revisar si no deseas incurrir en los mismos.

<sup>9</sup> Cuando decimos documento no hacemos distinción de su soporte, puede estar impreso, en formato digital, etc.

## I. Errores en la presentación.

- Cuadros sin identificación.
- Título o encabezamiento incorrecto o inadecuado:
  - Telegráfico: título demasiado pequeño, carente de claridad.
  - Ampuloso: título demasiado extenso, que incluye vocablos que no aportan nada a la claridad del texto.

## II. Errores del cuerpo.

- Errores de cálculo.
- Disposición incorrecta de los datos.
- Mostrar solamente medidas relativas (frecuentemente porcentajes) u otras medidas de resumen.
- Cuadros sobrecargados.

## III. Errores en la fuente.

- No citar la fuente cuando es secundaria.
- Citar la fuente cuando es primaria.
- Consignar como fuente aquello que no es un documento (oficinas, departamentos, centros, etc.)

### 2.1.3 ¿Cómo leer un cuadro estadístico?

Realmente parece algo tan trivial, que muchas personas lo pasan por alto muchas veces, lo que puede conllevar a interpretaciones erróneas de la tabla —o en el peor de los casos, a no entenderla—. Para que te evites el fiasco, echa una ojeada al orden propuesto:

1. Lee cuidadosamente el título, así sabrás de qué trata el cuadro exactamente.
2. Lee las notas explicativas. Obviamente, las mismas mejoran considerablemente la comprensión de la tabla.
3. Infórmate de las unidades de medida utilizadas.
4. Fíjate en el promedio total o porcentaje general del grupo.
5. Relaciona el promedio total con el porcentaje de cada una de las variables estudiadas.
6. Relaciona los promedios o porcentajes de las variables estudiadas.

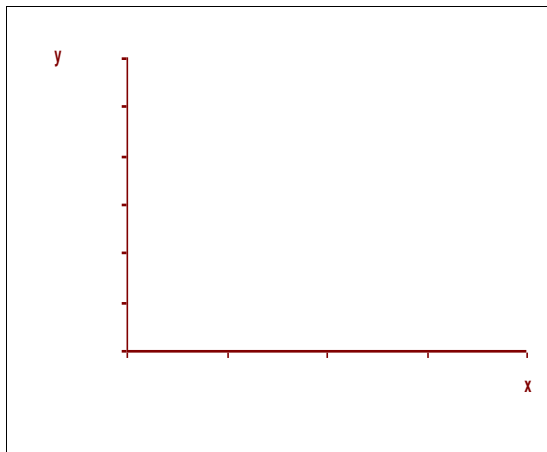
## 2.2 Los gráficos estadísticos

Otra manera de presentar la información estadística es a través de los gráficos. Ellos pueden resultar muy útiles, aunque en ocasiones un uso incorrecto los convierte en instrumentos estériles. Son complemento de las tablas, por ende, deben ser más autoexplicativos que ellas.

### 2.2.1 Características generales

Generalmente se inscriben en los ejes de coordenadas cartesianas o ejes rectangulares, los cuales:

- Deben poseer la misma longitud, aceptándose como máximo que el eje X exceda hasta 1.5 veces al eje Y. Esto evita la introducción de falacias.
- Deben estar rotulados. Por el eje X se presenta(n) la(s) variable(s) con su escala de clasificación; en el eje Y, la distribución de frecuencias o medida de resumen utilizada.
- De ser posible, el origen de los ejes debe ser en el punto (0,0).
- Deben utilizarse números redondos.
- Debe evitarse el exceso de divisiones de los ejes.



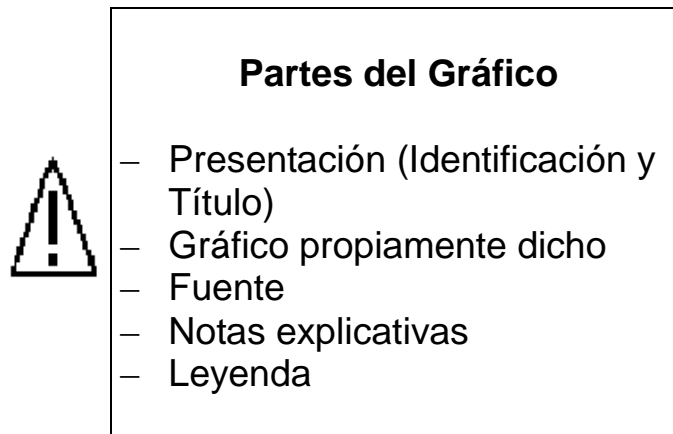
**Figura 1. Ejes cartesianos o rectangulares**

En la actualidad, con el advenimiento de las nuevas tecnologías informáticas, han proliferado los softwares que permiten la construcción de gráficos estadísticos. Al utilizarlos, debes tomar la precaución de analizar cuidadosamente el tipo de información que

quieres representar, pues la mayoría de ellos ofrece varias posibilidades de representación, quedando a tu juicio escoger la más apropiada.

### 2.2.2 Partes del gráfico

Todo gráfico estadístico está constituido por varios elementos, los cuales te mencionamos a continuación.



1. **Identificación:** consiste en numerar los gráficos consecutivamente, por ejemplo: Gráfico 1, Gráfico 2, etc.
2. **Título:** el de la tabla que lo originó.
3. **Gráfico propiamente dicho:** verás los distintos tipos de gráficos en el epígrafe siguiente.
4. **Fuente:** la tabla que lo originó.
5. **Notas explicativas:** su uso es similar a lo descrito en las tablas.
6. **Leyenda:** su fin es identificar los elementos del gráfico (barras, sectores, etc.) con su correspondiente origen.

### 2.2.3 Gráficos para representar variables en escalas cualitativa y cuantitativa discreta

A continuación te presentamos un grupo de gráficos que se estudiarán en este epígrafe, atendiendo al número de variables que representan.



<b>Variable s</b>	<b>Gráfico</b>
<b>1</b>	Barras simples, Pastel
<b>2</b>	Barras múltiples y Barras compuestas

### 2.2.3.1 Gráfico de barras simples

#### Uso

Es un gráfico formado por barras separadas que representan a las categorías de la variable en estudio. Se utiliza cuando queremos representar una variable cualitativa o cuantitativa discreta, y la información se dispone en frecuencias absolutas o relativas, o en medidas de resumen.

#### Elementos a considerar en su construcción

1. Dispón las barras separadas entre sí, para dar la idea de discontinuidad de la variable representada.
2. El ancho de las barras será opcional, pero debe ser el mismo para todas.
3. La separación entre barras debe ser igual a la mitad del ancho de ellas.
4. Si la variable es nominal, ordena las barras en orden creciente o decreciente, en dependencia de tus gustos.
5. Utiliza tantas barras como categorías tenga la variable.
6. Puedes colocar las barras en el eje vertical o en el horizontal. Comúnmente se utiliza el eje horizontal.
7. Este gráfico se origina a partir de tablas unidimensionales.

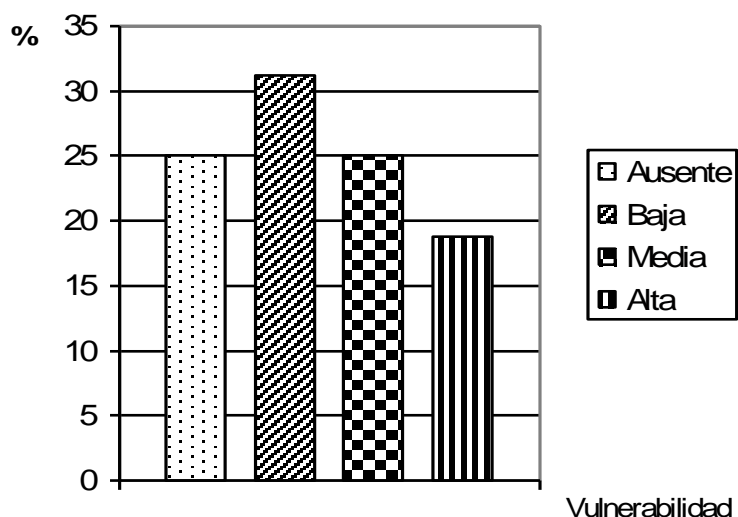
#### Ejemplo

Un grupo de investigadores desea conocer el comportamiento de la vulnerabilidad psicosocial en ancianos de un área de salud. Para ello aplica el cuestionario de vulnerabilidad-bienestar psicosocial del Dr. R. Pérez y obtiene los siguientes resultados:

Tabla 1. Distribución de ancianos según vulnerabilidad psicosocial. Municipio Playa, 1994.

Vulnerabilidad Psicosocial	Número	Porcentaje
<b>Ausente</b>	20	25.00
<b>Baja</b>	25	31.25
<b>Media</b>	20	25.00
<b>Alta</b>	15	18.75
<b>Total</b>	<b>80</b>	<b>100.00</b>

Gráfico 1. Distribución de ancianos según vulnerabilidad psicosocial.



Municipio Playa, 1994.  
Fuente: Tabla 1.

### 2.2.3.2 Gráfico de pastel, de sectores o circular

#### Uso

Este gráfico se utiliza cuando queremos representar una variable cualitativa o cuantitativa discreta, y la información se dispone en porcentaje. Básicamente, es un círculo dividido en sectores que representan las categorías de la variable.



## Elementos a considerar en su construcción

1. La totalidad de la información se representa por el número total de grados de un círculo (360°).
2. Para obtener los grados correspondientes a cada categoría, se multiplica 3.6° por la frecuencia relativa utilizada.

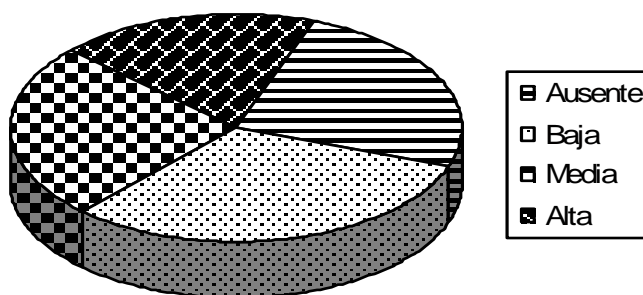
## Ejemplo

Utilizando la información del ejemplo anterior, y haciendo los cálculos pertinentes, el gráfico quedaría de la siguiente forma:

Cálculos previos:

	Porcentaje (%)	3.6 x %
<b>Ausente</b>	25.00	90.0
<b>Media</b>	31.25	112.5
<b>Baja</b>	25.00	90.0
<b>Alta</b>	18.75	67.5
	100.00	360.0

Gráfico 2. Distribución de ancianos según vulnerabilidad psicosocial. Municipio Playa, 1994.



Fuente: Tabla 1.

**Nota:** Con toda intención representamos los mismos datos por dos formas gráficas diferentes (barras simples y pastel), así te confirmamos la posibilidad de utilizarlos indistintamente, aunque si la variable tiene más de cinco categorías, es preferible usar las barras simples.

### 2.2.3.3 Gráfico de barras múltiples

#### Uso

Este gráfico se utiliza cuando queremos representar dos variables, las cuales pueden ser: cualitativas o cuantitativas discretas ambas, o una cualitativa y la otra cuantitativa discreta; y la información se dispone en frecuencias absolutas o relativas, o en medidas de resumen. Los datos se representan mediante barras agrupadas, como verás a continuación.

#### Elementos a considerar en su construcción

1. Dispondrás grupos de dos, tres o más barras, es decir, barras dobles, triples, etc.
2. El número de grupos a formar dependerá del número de categorías consignadas en la columna matriz o en la fila de encabezamiento, según tu gusto.
3. La separación entre cada grupo de barras es aproximadamente la mitad del ancho del grupo.
4. Este gráfico se origina a partir de tablas bidimensionales.

#### Ejemplo

El siguiente gráfico resume la información de 300 niños de un Círculo Infantil atendido por un médico de familia, atendiendo a las variables sexo y raza.

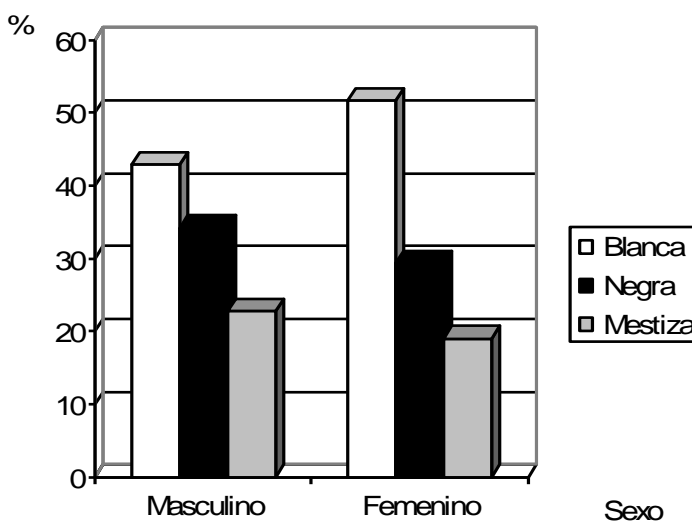
Tabla 2. Distribución de niños según raza y sexo. Círculo Infantil “El Camarón Encantado”. Municipio Playa, 1999.

Raza	Sexo			
	Masculino		Femenino	
	No.	%	No.	%
<b>Blanca</b>	79	42.9	60	51.7
<b>Negra</b>	63	34.2	34	29.3
<b>Mestiza</b>	42	22.8	22	19.0
<b>Total</b>	<b>184</b>	<b>100.0</b>	<b>116</b>	<b>100.0</b>

Fuente: Libro de matrícula del Círculo Infantil “El Camarón Encantado”. Curso académico 1999-2000.

Nota: el porcentaje se calculó por columnas.

Gráfico 3. Distribución de niños según raza y sexo. Círculo Infantil “El Camarón Encantado”. Municipio Playa, 1999.



Fuente: tabla 2.

### 2.2.3.4 Gráfico de barras compuestas

Uso

Al igual que el gráfico anterior, utiliza este cuando quieras representar dos variables: ambas cualitativas o cuantitativas discretas, o una cualitativa y la otra cuantitativa discreta; y dispongas la información en frecuencias relativas. Aquí, la información perteneciente a una variable se representa en su totalidad en una sola barra.

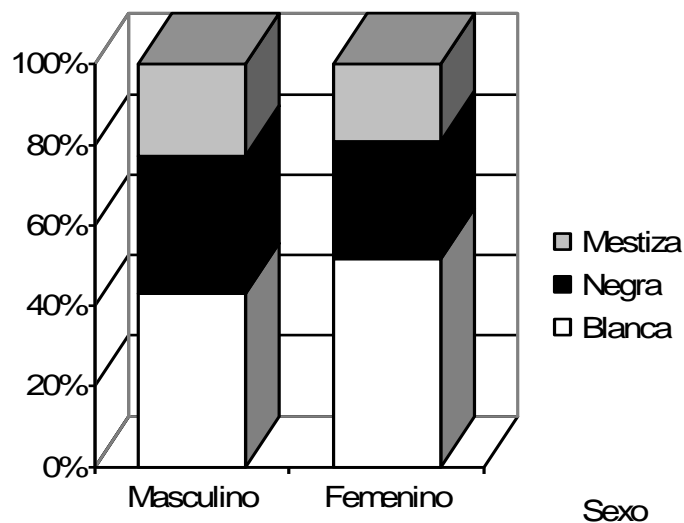
### Elementos a considerar en su construcción

1. Cada barra representa el ciento por ciento de la información del grupo representado.
2. El ancho de las barras queda a tu gusto, pero debe ser el mismo para todas.
3. La separación entre las barras es aproximadamente la mitad del ancho.
4. Lo originan tablas bidimensionales.

### Ejemplo

Utilizando la información del ejemplo anterior, el gráfico quedaría de la siguiente forma:

Gráfico 4. Distribución de niños según raza y sexo. Círculo Infantil “El Camarón Encantado”. Municipio Playa, 1999.



Fuente: tabla 2.

## 2.2.4 Gráficos para representar variables en escala cuantitativa continua

A continuación te presentamos un grupo de gráficos que se estudiarán en este epígrafe, atendiendo al número de variables que representan.



Variables	Gráfico
1	Histograma de frecuencias, ojiva*
2	Polígono de frecuencias

\*: No abordaremos este gráfico por su infrecuente uso en nuestro ámbito.

### 2.2.4.1 Histograma

#### Uso

Este gráfico consiste en barras adyacentes, y se utiliza cuando queremos representar una variable cuantitativa continua, y la información se dispone en frecuencias absolutas o relativas, o en medidas de resumen.

#### Elementos a considerar en su construcción

1. Las barras o rectángulos se disponen unidos para dar idea de continuidad.
2. El ancho dependerá de la amplitud de los intervalos de clase en que se clasifica la variable en estudio.
3. La altura de cada IC se obtiene mediante el cociente frecuencia absoluta/amplitud.
4. Por el eje X se consigna el límite de clase inferior o real de cada intervalo.
5. Lo originan tablas unidimensionales.

## Ejemplo

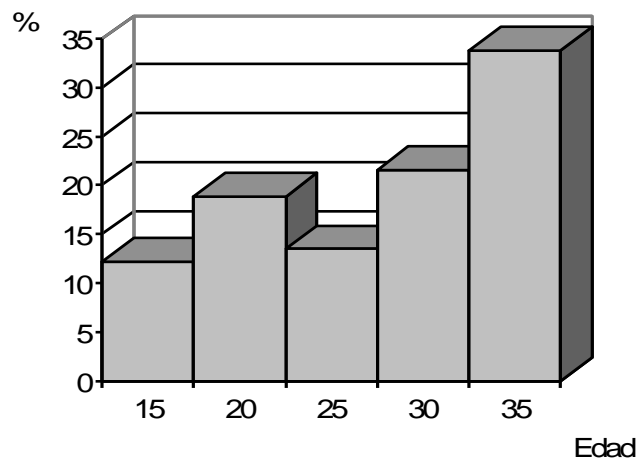
A continuación te presentamos los resultados de un estudio relacionado con las edades maternas.

Tabla 3. Distribución de recién nacidos según edad materna. HGO “E. Hernández, julio, 1999.

<b>Edad materna</b>	<b>Número</b>	<b>Porcentaje</b>
<b>15 – 19</b>	45	12.2
<b>20 – 24</b>	70	18.9
<b>25 – 29</b>	50	13.5
<b>30 – 34</b>	80	21.6
<b>35 – 39</b>	125	33.8
<b>Total</b>	<b>370</b>	<b>100.0</b>

Fuente: Registro de nacimientos. HGO “E. Hernández”, 1999.

Gráfico 5. Distribución de recién nacidos según edad materna. HGO “E. Hernández, julio, 1999.



Fuente: tabla 3

### 2.2.4.1 Polígono de frecuencias

## Uso

Este gráfico se utiliza cuando queremos representar hasta dos variables, de las que al menos una debe ser cuantitativa continua, y la información se dispone en frecuencias absolutas o relativas, o en medidas de resumen. Está formado por una o dos curvas que representan a cada variable estudiada.

## Elementos a considerar en su construcción

1. Se pueden construir histogramas inicialmente, y luego marcar los puntos medios de cada IC (marca de clase), los cuales al unirse forman una curva.
2. Habrá tantas curvas como categorías tenga la variable discontinua.
3. Lo originan tablas uni o bidimensionales.

## Ejemplo

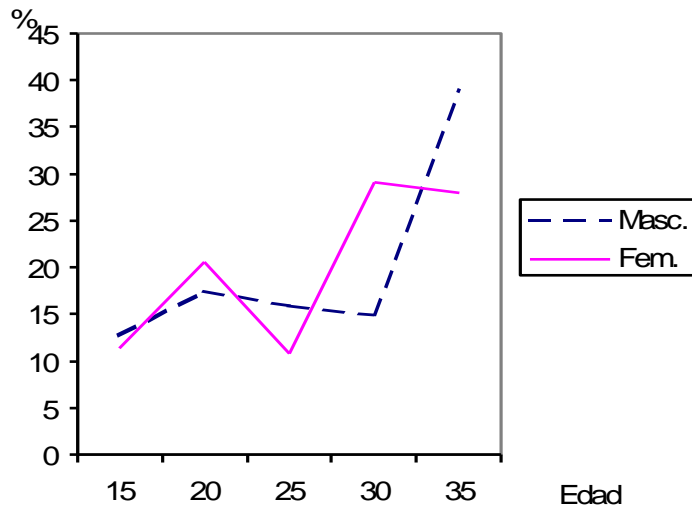
A continuación te presentamos los resultados de un estudio relacionado con las edades maternas y el sexo de los recién nacidos.

Tabla 4. Distribución de recién nacidos según edad materna y sexo. HGO “E. Hernández, julio, 1999.

Edad materna	Sexo			
	Masculino		Femenino	
	No.	%	No.	%
<b>15 – 19</b>	25	12.8	20	11.4
<b>20 – 24</b>	34	17.4	36	20.6
<b>25 – 29</b>	31	15.9	19	10.8
<b>30 – 34</b>	29	14.9	51	29.1
<b>35 – 39</b>	76	38.9	49	28.0
<b>Total</b>	<b>195</b>	<b>100.0</b>	<b>175</b>	<b>100.0</b>

Fuente: Registro de nacimientos. HGO “E. Hernández”, 1999.

Gráfico 6. Distribución de recién nacidos según edad materna y sexo. HGO “E. Hernández, julio, 1999.



Fuente: tabla 4

## 2.2.5 Gráficos para representar variables en el tiempo

A continuación te mostramos el gráfico aritmético simple, utilizado en el estudio de las series cronológicas o temporales, como algunos las llaman. No obstante, es bueno que sepas que hay quienes utilizan como gráfico de trabajo el de barras simples para el tratamiento de fenómenos vistos en el tiempo.

### 2.2.5.1 Gráfico aritmético simple (GAS)

#### Uso

Este gráfico se utiliza para representar una variable a través del tiempo.

#### Elementos a considerar en su construcción

1. Cada categoría o clase de la variable se representa por una curva.
2. En el eje de las abscisas se consignará el año, mes, semana, etc., según la unidad en que se mida el tiempo.
3. En ocasiones, cuando los ejes no ajustan, se utiliza una escala semilogarítmica para su construcción.



## Ejemplo

A continuación te presentamos la mortalidad perinatal de Cuba desde 1990 hasta 1998.

Tabla 5. Mortalidad perinatal según componentes. Cuba, 1990–1998.

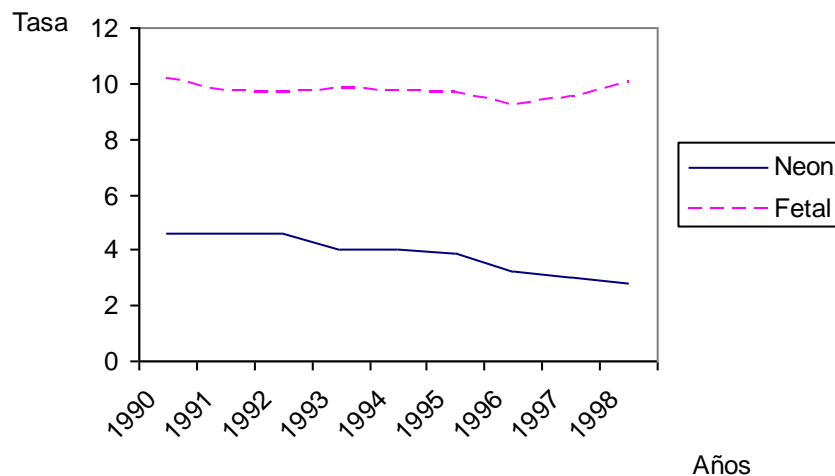
Años	Mortalidad			
	Neonatal		Fetal*	
	No.	Tasa	No.	Tasa
1990	861	4.6	1897	10.1
1991	811	4.6	1708	9.7
1992	720	4.6	1532	9.6
1993	601	4.0	1506	9.8
1994	598	4.0	1442	9.7
1995	586	3.9	1424	9.6
1996	456	3.2	1304	9.2
1997	461	3.0	1462	9.5
1998	435	2.8	1519	10.0

Tasa por 1 000 NV y defunciones fetales de 1 000 gramos y más

\*: 1 000 gramos y más

Fuente: Anuario Estadístico de Cuba. 1998.

Gráfico 7. Mortalidad perinatal según componentes. Cuba, 1990–



1998.

Gráfico 7: Mortalidad perinatal según componentes. Cuba 1990-1998

Fuente: tabla 5

## Resumen

En este tema estudiaste que:

1. La información puede presentarse de forma resumida en **tablas y gráficos estadísticos**.
2. La **tabla o cuadro estadístico** es un recurso que se emplea con el fin de presentar información resumida, organizada por filas y columnas.
3. La estructura del cuadro estadístico incluye: Identificación, Título, Cuerpo de la tabla, Fuente y Notas explicativas.
4. Los errores cometidos con mayor frecuencia en la confección de tablas estadísticas son:

### I. Errores en la presentación.

- Cuadros sin identificación.
- Título o encabezamiento incorrectos o inadecuados:
  - Telegráfico: título demasiado pequeño, carente de claridad.
  - Ampuloso: título demasiado extenso, que incluye vocablos que no aportan nada a la claridad del texto.

### II. Errores del cuerpo.

- Errores de cálculo.
- Disposición incorrecta de los datos.
- Mostrar solamente medidas relativas (frecuentemente porcentajes) o medidas de resúmenes.
- Cuadros sobrecargados.

### III. Errores en la fuente.

- No citar la fuente cuando es secundaria.
- Citar la fuente cuando es primaria.
- Consignar como fuente aquello que no es un documento (oficinas, departamentos)

5. El gráfico estadístico complementa la información previamente presentada en tablas.
6. La estructura del gráfico estadístico incluye: Identificación, Título, Gráfico propiamente dicho, Fuente, Notas explicativas y Leyenda.
7. Los gráficos más comúnmente usados en el ámbito sanitario son: de barras, pastel, histograma y polígono de frecuencias.

## Ejercitación

1. A continuación te presentamos parte de la información obtenida en un estudio realizado en médicos de familia del municipio Cerro, sobre el comportamiento del síndrome de burn out. Construye el gráfico apropiado para representar esta información.

Tabla 1. Distribución de médicos de familia según síndrome de burn out y sexo. Municipio Cerro, 1998.

Síndrome de burn out	Sexo			
	Masculino		Femenino	
	No.	%	No.	%
<b>Ausente</b>	15	42.8	12	34.3
<b>Ligero</b>	10	28.6	10	28.6
<b>Moderado</b>	5	14.3	8	22.8
<b>Grave</b>	5	14.3	5	14.3
<b>Total</b>	<b>35</b>	<b>100.0</b>	<b>35</b>	<b>100.0</b>

2. A continuación te presentamos los resultados de la medición de creatinina sérica en 20 personas del municipio Playa, atendidos en el policlínico "26 de Julio" en enero de 1999.

Creatinina Sérica			
160,00	170,54	166,00	150,00
160,36	160,20	156,70	151,78
158,20	163,20	154,50	152,00
174,00	165,80	155,00	154,80
170,00	165,90	155,90	156,70

- a) Con la información anterior, construye una escala de intervalos de igual amplitud.
- b) Construye un cuadro estadístico e interpreta los resultados.
- c) Construye un gráfico estadístico.

3. Para los siguientes títulos de tablas, escoge el gráfico que mejor la represente.
- Distribución de ancianos según sexo. La Habana, 1999.
  - Distribución de médicos de familia según edad y sexo. Municipio Playa, 1998.
  - Mortalidad materna. Cuba, 1990-1996.
  - Distribución de recién nacidos según raza. HGD "Dr. A. Neto", 1998.
  - Distribución de ancianos según peso. La Habana, 1999.
4. El siguiente cuadro muestra información resultante de cierto estudio. Analízalo cuidadosamente y emite tu criterio al respecto.

Tabla # 1. Muestra de las madres del Consultorio del Médico de la Familia # 1925 según el número de hijos tenidos en el último matrimonio. CMF # 1925, municipio Contramaestre, provincia Santiago de Cuba, enero de 1962.

# de hijos	Edad materna				Total	
	< 20		> 20			
	#	%	#	%	#	%
<b>0</b>	8	8.0	14	14.0	22	22.0
<b>1-3</b>	20	20.0	6	6.0	26	26.0
<b>3-4</b>	10	10.0	19	19.0	29	29.0
<b>5-6</b>	10	10.0	6	6.0	16	16.0
<b>6 y más</b>	2	2.0	7	7.0	9	9.0
<b>Total</b>	<b>50</b>	<b>100.0</b>	<b>50</b>	<b>100.0</b>	<b>100</b>	<b>100.0</b>

Fuente: Encuesta

### Autoevaluación

- La información siguiente es parte de los resultados de un estudio realizado en escolares del municipio Marianao, provincia Ciudad de La Habana, durante 1997-1999.

<u>Nombre</u>	<u>Sexo</u>	<u>Edad</u>
YHP	F	8
HHD	M	7
GTR	F	6
TRF	M	5
ADE	F	14
LOI	F	11
YGF	M	5
HTR	M	9
SDE	F	10
GTW	F	8
JFY	F	9
BHG	M	8
FRW	M	12
RTL	F	8
ABV	F	9
JUY	M	10

a) Construye una tabla con estos datos, y represéntalos en un gráfico de ser posible.

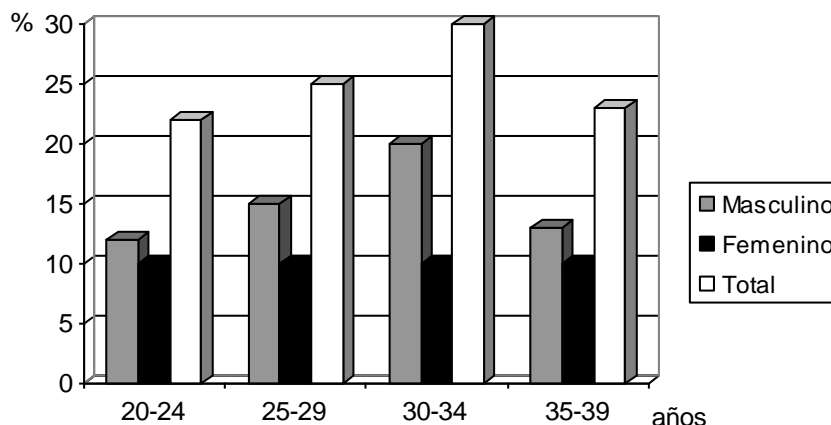
2. Analiza cuidadosamente la información siguiente y di tu opinión al respecto.

Tabla: Sexo de los pacientes de 20 – 40 años según la edad en la muestra estudiada. Hospital General Provincial. 1998.

<b>Edad (años)</b>	<b>Masculino</b>		<b>Sexo Femenino</b>		<b>Total</b>	
	<b>No.</b>	<b>%</b>	<b>No.</b>	<b>%</b>	<b>No.</b>	<b>%</b>
<b>20 – 24</b>	12	12.0	10	10.0	22	22.0
<b>25 – 29</b>	15	15.0	10	10.0	25	25.0
<b>30 – 34</b>	20	20.0	10	10.0	30	30.0
<b>35 – 39</b>	13	13.0	10	10.0	23	23.0
<b>Total</b>	60	60.0	40	40.0	100	100.0

Fuente: Encuesta.

Gráfico: Sexo y edad de la casuística estudiada. Hospital General



Provincial. 1998.  
Fuente: Encuesta.

### Bibliografía

1. Daniel WW. Bioestadística. Base para el análisis de las ciencias de la salud. 3ª ed. México D.F.:Limusa; 1997.
2. Anuario Estadístico de la República de Cuba. 1998. La Habana:UNICEF-FNUAP; 1999.
3. Spiegel MR. Teoría y problemas de Estadística. La Habana:Pueblo y Educación; 1977.
4. Camel F. Estadísticas médicas y de Salud Pública. La Habana:Pueblo y Educación; 1985.

## Tema 3. Medidas de resumen para variables cualitativas

### Introducción

En la investigación científica en la Atención Primaria de Salud, con frecuencia se utilizan variables cualitativas, bien por su naturaleza, o por la escala empleada. Por supuesto, una vez que la información se recogió, es necesario calcular alguna medida de resumen cuyo resultado es un indicador que deberá analizarse en un momento posterior.

En este tema te presentamos las medidas de resumen para variables cualitativas que se utilizan con mayor frecuencia en los estudios que realizas en el nivel primario de atención de salud.

#### 3.1 Razón e Índice. Definición. Cálculo e interpretación

Por **razón** puede entenderse:



Una **razón** es la relación por cociente que se establece entre las unidades de análisis que pertenecen a un grupo o categoría (a) y las unidades de análisis que pertenecen a otra categoría (b) de la misma variable. Su expresión general es:  $\frac{a}{b}$ .

¡Uhhm! ¿Ésa es la definición? No te desanimes, es una medida de fácil comprensión. Te la explicaremos con un ejemplo:

Supongamos que de los 400 recién nacidos (RN) de un municipio en cierto período, 300 presentaron los ojos oscuros (OO), en tanto que sólo 100 los tenían claros (OC). Aplicando la expresión general, la razón OO/OC es:

$$R = \frac{a}{b} = \frac{\text{RN con OO}}{\text{RN con OC}} = \frac{300}{100} = 3$$

Nota: Utilizamos la letra **R** por razones didácticas, realmente la razón no tiene símbolo propio.

La razón ojos oscuros/ojos claros es de 3; o lo que es lo mismo, 3:1.

Pero, ¿qué significa este resultado? **Expresa que hay tres recién nacidos con ojos oscuros por cada recién nacido de ojos claros en ese municipio y en ese período.**

Fíjate que el numerador y el denominador son disjuntos, es decir, no se interceptan, no están contenidos uno en el otro. Ello te ayudará a establecer las diferencias con las medidas de resumen que estudiarás a continuación.

Si multiplicas el resultado obtenido por 100, entonces el nuevo número se denomina **índice**, de tal suerte que en el ejemplo anterior el índice sería 300. En otras palabras, **en el municipio de referencia, en el período estudiado, por cada 100 bebés de ojos claros hay 300 de ojos oscuros.**

### 3.2 Proporción y Porcentaje. Definición. Cálculo e interpretación

Por **proporción** se entiende:



Una **proporción** es la relación por cociente que se establece entre las unidades de análisis que pertenecen a un grupo o categoría ( $a$ ) de una variable y el total de las unidades de análisis estudiadas ( $a + b$ ). Su expresión general es:  $\frac{a}{a + b}$ . Si se multiplica su resultado por 100, se obtendrá el **porcentaje**.

Seguiremos utilizando el ejemplo del epígrafe anterior. ¿Lo recuerdas? Por supuesto que sí. Pues bien, determinemos la proporción de niños con ojos oscuros (300) en la población de recién nacidos (400):

$$P = \frac{a}{a + b} = \frac{\text{RN con OO}}{\text{Total de RN}} = \frac{300}{400} = \frac{3}{4}$$



Alternativamente, puedes calcular el porcentaje:

$$P \times 100 = \frac{\text{RN con OO}}{\text{Total de RN}} \times 100 = \frac{300}{400} \times 100 = 75.0\%$$

Nota: Usamos la **P** con fines ilustrativos, pues la proporción carece de simbología.

Los resultados anteriores significan que tres de cada cuatro recién nacidos tienen los ojos oscuros; o que el 75 por ciento de los recién nacidos tiene los ojos oscuros (y, obviamente, el 25% los tiene claros).

¿No te resultan familiares estas nuevas medidas, o sea, la proporción y el porcentaje? Ya debes estarte preguntando la diferencia que existe entre éstas y la distribución de frecuencias relativas que ya estudiaste. Nada más claro: no es que sean parecidas, *son exactamente las mismas*, pero restringidas a variables cualitativas.

Observa que el porcentaje te permite analizar el aporte, el peso específico o la importancia relativa de cada categoría respecto al total.

Otro elemento que debes conocer es el siguiente: si la variable es dicotómica, puedes utilizar tanto razones como proporciones; pero si es politómica, entonces calcula sólo proporciones<sup>10</sup>.

### 3.3 Tasas

Siempre que necesites medir el riesgo de que acontezca cierto fenómeno en una población determinada, dispones de un indicador valioso y único: las tasas.

---

<sup>10</sup> De hecho, la interpretación y cálculo de razones en una politomía está proscrito.



Una **tasa** es una relación por cociente que expresa el riesgo de que ocurra cierto evento en una población y período determinados. Está compuesta por tres elementos, a saber:

$$\text{Tasa} = \frac{\text{numerador}}{\text{denominador}} \times k$$

Veamos cuáles son esos elementos:

- El numerador contiene al número de veces que ocurrió determinado fenómeno en un área geográfica y en un período determinados.
- El denominador indica el número de habitantes de la población en la cual puede ocurrir el fenómeno.
- k es un múltiplo de 10 cuyo uso está justificado por el hecho de que habitualmente el resultado del cociente es un número fraccionario, y al multiplicarlo por una potencia de 10 se facilita enormemente la lectura y comprensión del indicador.

Esta es una medida que expresa el **riesgo de ocurrencia** del evento estudiado en el numerador en la población involucrada, en el tiempo y lugar establecidos.

Las tasas que más importancia revisten para nuestro desempeño en el campo de la Salud Pública son las siguientes:



### **Tasas de importancia Relevante en Salud Pública**

- Tasas relacionadas con la natalidad
- Tasas relacionadas con la mortalidad
- Tasas relacionadas con la morbilidad

Una particularidad realmente útil de las tasas es que puedes calcularlas tanto para la totalidad de la población, como para parte de ella (por ejemplo, para el grupo de edad de cinco a nueve años, para los estudiantes, para los residentes del área rural, y así por el estilo); por otra parte, puedes calcular las tasas para todas las causas, o

solamente para una de ellas (o un grupo de ellas). De este modo, tendrás calculadas **tasas brutas, crudas, generales o globales** si se tratara de tasas que involucren a toda la población o al total de causas; al tiempo que habrás calculado **tasas específicas** si incluían a una parte de la población o a una causa o grupo de ellas.

Así las cosas, estarás en plena facultad de hallar tasas brutas de mortalidad, de natalidad, o bien específicas por edad, por sexo, por edad y sexo a la vez, entre muchas otras. Teniendo a tu disposición los datos adecuados, podrás hallar una tasa tan específica como desees.

Existe en punto cardinal en el manejo de las tasas: **la población expuesta al riesgo en cuestión**. Como ya sabes, este es el denominador de la ecuación, y de su correcta determinación depende la fidelidad del cálculo. Nunca serán suficientes las medidas que tomes para asegurarte que estás empleando el dato acertado. No creas que es muy difícil saber que estás errado o en lo cierto, el problema radica en que muchas veces se pasa por alto este “*detalle*” de forma involuntaria.

Probablemente te habrás preguntado: «Bueno, ¿y qué tanto problema con el denominador?» ¡Ah! Es que ahí radica el quid de la cosa. Recuerda que calculas una tasa para medir el riesgo de ocurrencia de un evento o fenómeno en una población, pero *no en cualquier población*, sino en la población **expuesta** a ese riesgo. Esto quiere decir que sólo podrás calcular la tasa de mortalidad por cáncer de útero en las **mujeres** de cierta ciudad, puesto que sería imposible calcularla en los **hombres**; del mismo modo que no puedes calcular la tasa de morbilidad por cáncer de pulmón de los habitantes de **Pueblo Mocho** en **1999**, utilizando para ello a los habitantes que tenía el pueblo en **1979**, o a los habitantes de **Palma Mocha** en 1999. ¿Satisfecha tu inquietud?

También haz de saber que las poblaciones están sometidas a constantes cambios en lo que a su número atañe, determinados por los nacimientos y defunciones y por los movimientos migratorios (emigración e inmigración), que provocan que no sea la misma a lo largo de todo el año. De ahí que, por convenio, se tome la población

existente **a mediados del período**<sup>11</sup> **o población media** para el cálculo de las tasas.

Por otra parte, debes tener especial cuidado al calcular tasas para poblaciones pequeñas, como la que usualmente manejan los Consultorios, pues suelen volverse inestables, ya que cualquier evento “mueve” mucho la tasa, y a veces no guarda relación el resultado obtenido con la magnitud del evento acontecido.

Bueno, ya estamos en condiciones de particularizar en las tasas más relevantes en la práctica diaria.

### 3.3.1 Tasas relacionadas con la natalidad

El estudio de la natalidad está relacionado con el número de nacimientos ocurridos en una población y tiempo determinados, así como la distribución que siguen de acuerdo con ciertas características. Como ves, todo gira en torno a la medición de la misma, y una de las formas de conseguirlo es utilizando las tasas.

Ahora nos tropezamos con una contrariedad: la población expuesta al riesgo es muy difícil de definir, ya que tener un hijo no involucra a toda la parte femenina de la población, y va más allá, pues otros factores de índole psicosocial actúan en tal decisión. Por estas razones, verás que se han buscado soluciones alternativas a esta situación.

#### 3.3.1.1 Tasa bruta de natalidad

Comencemos por la **tasa bruta de natalidad**. La misma expresa cómo se comportan los nacimientos en un área y tiempo determinados. Su cálculo es sencillo:

$$\text{TBN} = \frac{\text{total nacidos vivos (NV) en lugar y tiempo X}}{\text{población total en el lugar y tiempo X}} \times 1000$$

Por ejemplo, la tasa cruda de natalidad de Cuba en 1998 fue:

Total de nacidos vivos en Cuba durante 1998: 151 080<sup>12</sup>.

<sup>11</sup> Corrientemente se toma la población del 30 de junio.

Total de habitantes en Cuba durante 1998: 11 122 308.

$$\text{TBN} = \frac{151\,080}{11\,122\,308} \times 1000 \approx 13.6$$

Bien, ya tienes el número calculado. Pero, ¿es suficiente con eso? Claro que no, necesitas saber qué significa, a fin de manejarlo apropiadamente. En primer lugar, debes informar el resultado de la siguiente forma: **«La tasa bruta de natalidad de Cuba en 1998 fue de 14 nacidos vivos por cada 1000 habitantes»**, ello significa que **durante 1998 en Cuba nacieron como promedio 14 niños por cada 1 000 habitantes.**

Esta tasa tiene el inconveniente de no tomar en cuenta a las personas realmente expuestas al evento, pero por su sencillez y facilidad de comprensión es la medida más generalmente utilizada.

De ahora en adelante nos limitaremos a enseñarte cómo calcular e interpretar el indicador. Continuemos entonces.

### 3.3.1.2 Tasa general de fecundidad

Este indicador mide la natalidad, pero tomando en cuenta solamente a la población femenina en edad reproductiva o fértil (15 a 49 años). El hecho de que se restringe el denominador no inyecta especificidad a la tasa, pues continúa siendo una mezcla de diversos grupos de edades con situaciones diversas; amén de que se mueve a la par de la tasa cruda de natalidad. Se calcula de la siguiente forma:

$$\text{TGF} = \frac{\text{total nacidos vivos en el lugar y tiempo } X}{\text{mujeres en edad fértil en el lugar y tiempo } X} \times 1000$$

Así, la tasa de Cuba en 1998 fue:

$$\text{TGF} = \frac{151\,080}{3\,055\,907} \times 1000 \approx 49.4$$

Interpretación: **En Cuba, durante 1998, nacieron como promedio 49 niños por cada 1000 mujeres en edad fértil (15 a 49 años).**

<sup>12</sup> A menos que se indique lo contrario, todos los datos que aparecen en este tema proceden del Anuario Estadístico de la República de Cuba, 1998.

### 3.3.1.3 Tasa de fecundidad específica por edad

Esta es una tasa específica, que usualmente se calcula para grupos quinquenales comprendidos entre 15 y 49 años.

$$TEF_{\text{edad}} = \frac{\text{NV de mujeres de edad } i \text{ en lugar y tiempo } X}{\text{mujeres de edad } i \text{ en lugar y tiempo } X} \times 1000$$

$$TEF_{15-49} = \frac{19\,476}{345\,942} \times 1000 \approx 56.3$$

Interpretación: **Durante 1998 en Cuba nacieron como promedio 56 niños por cada 1000 mujeres de 15 a 49 años de edad.**

### 3.3.2 Tasas de mortalidad

La medición de la mortalidad tiene como fin conocer el número de defunciones ocurridas en cierta población durante un período dado, a la vez que se estudia su distribución relacionándolas con diversas características de dicha población.

Entrando en la materia que nos ocupa, te decimos a continuación las tasas que podrás calcular.

#### 3.3.2.1 Tasa bruta de mortalidad

Esta tasa expresa el riesgo que tienen todos los habitantes de cierta población, en un momento dado, de morir por cualquier causa.

$$TBM = \frac{\text{fallecidos en lugar y tiempo } X}{\text{población en lugar y tiempo } X} \times 1000$$

En 1998, en nuestro país esta tasa fue:

$$TBM = \frac{77\,558}{11\,122\,308} \approx 7.0$$

Esto significa que **en 1998, en Cuba fallecieron como promedio 7 personas por cada 1 000 habitantes.**

### 3.3.2.2 Tasa de mortalidad por edad

Ahora te presentamos una tasa de mortalidad específica, que solo mide el riesgo de morir que tienen las personas del grupo de edad analizado. Su cálculo se logra restringiendo el denominador a las personas de la edad deseada, e incluyendo en el numerador a los fallecidos en esa edad.

$$TME_{\text{edad}} = \frac{\text{fallecidos de edad } i \text{ en lugar y tiempo } X}{\text{población de edad } i \text{ en lugar y tiempo } X} \times 1000$$

Por ejemplo, en 1998, en Cuba, la tasa de mortalidad en personas de 60 años y más fue:

$$TME_{60 \text{ y } +} = \frac{52\ 558}{1\ 440\ 127} \approx 50$$

**Así, durante 1998, en Cuba fallecieron como promedio 50 individuos de 60 y más años por cada 1 000 personas de ese grupo de edad.**

### 3.3.2.3 Tasa de mortalidad por sexo

El cálculo de esta tasa es muy similar a la anterior, con la diferencia de que te restringes a un sexo en particular. Expresa el riesgo de morir de las personas de ese sexo en esa población, en el período especificado. Para calcularla, sustituye el numerador por el total de defunciones del sexo analizado, y el denominador por el total de habitantes de ese sexo en el lugar y momento deseados.

$$TME_{\text{sexo}} = \frac{\text{fallecidos de sexo } i \text{ en lugar y tiempo } X}{\text{población sexo } i \text{ en lugar y tiempo } X} \times 1000$$

En nuestro país, durante 1998 la tasa de mortalidad del sexo femenino fue:

$$TME_{\text{fem}} = \frac{34\ 692}{5\ 550\ 426} \times 1\ 000 \approx 6$$

Interpretación: **en Cuba, en 1998 fallecieron como promedio 6 mujeres por cada 1000 féminas.**

### 3.3.2.4 Tasa de mortalidad por causa

Análogamente, puedes conocer el riesgo a que están sometidos los habitantes de cierto lugar, en un momento definido, de morir por una causa de muerte dada. Ahora el numerador está formado por las defunciones debidas a la causa en cuestión, mientras que el denominador incluye al total de población.

$$TME_{causa} = \frac{\text{fallecidos por causa } i \text{ en lugar y tiempo } X}{\text{población en lugar y tiempo } X} \times 100\,000$$

En Cuba, durante 1998 la tasa de mortalidad por enfermedades del corazón fue:

$$TME_{enf. \text{ cor}} = \frac{21\,467}{11\,122\,308} \times 100\,000 \approx 193.0$$

Interpretación: **en Cuba, en 1998 fallecieron como promedio 193 personas por enfermedades del corazón por cada 100 000 habitantes.**

### 3.3.2.5 Tasa de mortalidad infantil

Arribamos a un punto de suma importancia al analizar la situación de salud de una comunidad. Este indicador es una especie de diana hacia la cual se dirigen los ojos de todo aquel que, avezado o no, se tome interés en el estudio de las características de una población.

Es un indicador que toma como población expuesta al riesgo a los nacidos vivos en período estudiado, y se calcula de la siguiente forma:

$$TMI = \frac{\text{fallecidos menores de 1 año en lugar y tiempo } X}{\text{NV en lugar y tiempo } X} \times 1\,000 \text{ NV}$$

A partir del triunfo revolucionario, este indicador ha mostrado una tendencia descendente, llegando a alcanzar en los dos últimos años



cifras inferiores a 8, incluyéndose de esta forma entre los países de más baja tasa a escala mundial. En 1998, la tasa cubana fue:

$$TM = \frac{1070}{151080} \times 1\,000 \approx 7.1$$

Ello significa que **en 1998, en Cuba fallecieron como promedio 7 niños por cada 1000 nacidos vivos.**

Este indicador tiene la singularidad de que puede descomponerse en varios indicadores, que miden con más especificidad el comportamiento de la mortalidad en el menor de un año. Estos componentes son los siguientes:

### 1. Tasa de mortalidad neonatal precoz

Al calcular esta tasa conocerás el riesgo de morir de los bebés con menos de siete días de nacidos. Su cálculo estriba en sustituir el numerador de la TMI por las defunciones ocurridas en recién nacidos de menos de siete días en el período y lugar estudiados. En 1998 tuvimos una TMNP de 2.9 por 1000 NV. De este modo, puedes decir que **en Cuba, durante 1998, fallecieron como promedio 3 niños de menos de 7 días por cada 1000 nacidos vivos.**

### 2. Tasa de mortalidad neonatal tardía

Conforme calculaste el riesgo de muerte de los bebitos menores de siete días, puedes conocer también el de siete en adelante y menores de 28 días, cerrando así el diapasón en la etapa neonatal de la vida. Sólo tienes que sustituir el numerador de la tasa anterior por las defunciones de niños de 7 - 27 días en la población de tu interés, durante el período que necesites.

Para nuestro país la TMNT en 1998 fue de 1.4 por cada 1000 nacidos vivos, lo que quiere decir que **en 1998, en Cuba falleció como promedio 1 niño de 7 a 27 días por cada 1000 nacidos vivos.**

### 3. Tasa de mortalidad posneonatal

Ahora determinarás el riesgo de muerte de los niños mayores de 28 días y menores de un año. Con sólo sustituir el numerador de la TMI por las defunciones acaecidas en los bebés de 28 días a 11 meses, 29 días y 23:59 horas, habrás cumplido tu cometido.

El que el denominador, de los tres componentes de la mortalidad infantil, sea el mismo le imprime a estas tasas una peculiaridad: se puede obtener la mortalidad infantil mediante la simple suma de sus componentes, o lo que es lo mismo, los componentes de la mortalidad infantil son sumables.

En Cuba, durante 1998, tuvimos una TMP de 2.8 por 1000 NV. Dicho sea con otras palabras: **en 1998, en Cuba fallecieron como promedio 3 niños mayores de 28 días y menores de un año por cada 1000 nacidos vivos.**

### 3.3.2.6 Tasa de mortalidad perinatal

Esta es una tasa especial que mide el riesgo de morir en los momentos cercanos al nacimiento. Se calcula de la siguiente forma<sup>13</sup>:

$$TMP = \frac{DFT + DNP}{NV + DFT} \times k$$

donde:

- DFT: defunciones fetales tardías (edad materna igual o superior a las 28 semanas, o peso fetal de 1000 gramos o más).
- DNP: defunciones neonatales precoces (defunciones en el menor de siete días).
- NV: nacidos vivos

### 3.3.2.6 Tasa de mortalidad materna

Aquí tienes otro de los indicadores más celosamente cuidados por todo el personal de salud, bien sabes de ello. La lógica aspiración de todo país interesado realmente en exhibir indicadores de salud ejemplares, es mantener esta medida en niveles bajos, juntamente con la tasa de mortalidad infantil, entre otros. Su cálculo comprende algo que puede inducir extrañeza: el denominador está formado por los nacidos vivos del lugar y tiempo escogidos. Al analizarlo con

<sup>13</sup> Esta es la forma más comúnmente utilizada para su cálculo.

detenimiento verás que resulta lo más indicado, ya que brinda una estimación mejor del riesgo puesto que este indicador solamente toma en cuenta las defunciones maternas producidas por complicaciones del embarazo, parto o puerperio (entendido como los 42 días siguientes al parto).

$$TMM = \frac{\text{defunciones maternas directas en lugar y tiempo } X}{\text{nacidos vivos en lugar y tiempo } X} \times 100\,000$$

La TMM fue de 47.7 por 100 000 NV en 1998 para nuestro país. Esto quiere decir que **por cada 100 000 nacidos vivos, murieron en promedio 48 mujeres por causas directamente relacionadas con el embarazo, parto y puerperio durante 1998 en Cuba.**

### 3.3.3 Tasas de morbilidad

La morbilidad, entendida como el patrón de enfermedades que sufren los habitantes de alguna región, puede ser estudiada numéricamente mediante las tasas de morbilidad. Ellas son la ***tasa de incidencia, la tasa de prevalencia y la tasa de letalidad.***

La ***tasa de incidencia*** (TI) mide el riesgo que tiene una persona que habita en un lugar y tiempo determinados, de contraer o adquirir cierta enfermedad, visto esto en función del tiempo. Por su lado, la ***tasa de prevalencia*** (TP) mide el riesgo de tener la enfermedad, o sea, de estar enfermo; y la ***tasa de letalidad*** (TL) expresa la gravedad de la enfermedad.

$$TI = \frac{\text{casos nuevos en lugar y tiempo } X}{\text{población total en lugar y tiempo } X} \times 10\,000$$

$$TP = \frac{\text{casos nuevos + casos viejos en lugar y tiempo } X}{\text{población total en lugar y tiempo } X} \times 10^n$$

$$TL = \frac{\text{fallecidos por enfermedad } i \text{ en lugar y tiempo } X}{\text{enfermos por enfermedad } i \text{ en lugar y tiempo } X} \times 100$$

### 3.3.4 Comparación de tasas

En ocasiones, pretendemos comparar los riesgos de morir, de enfermar, etc. entre distintas poblaciones o entre distintas categorías o clases de una variable. Para ello, lo más conveniente es utilizar la **tipificación**, bien por el método directo o por el indirecto; técnicas que no se expondrán en este curso, pues se abordarán en cursos posteriores. Esta técnica solo sirve para comparar, sus resultados no miden en modo alguno el riesgo de ocurrencia de los eventos estudiados en la población.

## Resumen

En este tema estudiaste que:

1. Las medidas de resumen para datos cualitativos más frecuentemente utilizadas son **las razones, las proporciones y las tasas**.
2. Cada uno de esos indicadores tiene diferente interpretación. Así, los más refinados son las tasas, pues expresan el riesgo de ocurrencia del evento consignado en su numerador.
3. Debes tener cuidado al calcular las tasas para poblaciones pequeñas, por ejemplo, en el Consultorio Médico de la Familia, porque suelen ser inestables.
4. Las tasas pueden dividirse en **generales y específicas**.
5. En el ámbito sanitario, las tasas más usadas son las de natalidad, mortalidad y morbilidad.

## Ejercitación

- I. De los siguientes indicadores, di su nombre y qué datos necesitarías para calcularlos.
  1. Riesgo de morir por enfermedades del corazón en Guantánamo, 1998.
  2. Riesgo de morir por cualquier causa en Ciego de Ávila, 1997.
  3. Contribución del grupo de 60 años y más a la mortalidad general en Cuba, 1998.
  4. Contribución del sexo femenino a la estructura poblacional de Cuba, 1998.
  5. Riesgo de morir en la primera semana de vida extrauterina.
  6. Riesgo de morir de los enfermos de tuberculosis.

II. Completa e interpreta los datos siguientes:

Nombre del indicador	Numerador	Denominador	Resultado
Tasa de mortalidad general, Cuba, 1998.	77 558		7.0
Tasa de natalidad, Cuba, 1997.		152 491	13.8
Tasa de mortalidad infantil, Ciudad de La Habana, 1998.	211	28 133	

Nota: todas las tasas están calculadas utilizando  $k = 1\ 000$ .

### Autoevaluación

1. Un grupo de investigadores recogió algunos datos relacionados con la población cubana del año 1998, con el objetivo de confeccionar indicadores que reflejaran la situación del país. Debido a un virus informático, se estropeó parte de la información almacenada. A continuación te mostramos los datos que se pudieron recuperar. A partir de los mismos, ¿podrías ayudarnos a completar las partes faltantes? Para ello, calcula e interpreta los indicadores solicitados.

#### Información recogida por los investigadores

- Nacidos vivos bajo peso: 10 145
- Población total: 11 122 308
- Nacidos vivos: 151 080
- Defunciones totales: 77 558
- Total de hombres: 5 571 882
- Total de consultas médicas y estomatológicas: 100 819 793
- Fallecidos de 15 a 49 años: 10 057
- Total de mujeres: 5 550 426
- Fallecidos menores de un año: 1 070
- Fallecidos de la provincia Guantánamo: 2 722
- Casos diagnosticados por enfermedad meningocócica: 44
- Fallecidos mayores de 28 días y menores de 12 meses: 417
- Población de Guantánamo: 508 864
- Hombres fallecidos por tumores malignos: 9 126
- Total de nacidos vivos en Sancti Spíritus: 5 642

- Mujeres fallecidas: 34 692
- Fallecidos menores de 7 días: 435
- Población de 15 a 49 años: 6 117 424
- Fallecidos mayores de 7 días y menores de 27 días: 218
- Nacidos vivos de la provincia Guantánamo: 7 939

### **Indicadores solicitados:**

- Relación entre hombres y mujeres.
- $\frac{5\,550\,426}{11\,122\,308} = 0.49$
- Riesgo de morir de la población cubana.
- Riesgo de morir de la población guantanamera.
- Riesgo de morir de las mujeres.
- $\frac{9\,126}{5\,571\,882} = 163.8$
- Riesgo de contraer la enfermedad meningocócica.
- Índice de bebés bajo peso al nacer.
- Relación entre el total de consultas médicas y de Estomatología y la población total.
- Natalidad en la provincia Guantánamo.
- Natalidad en la provincia Sancti Spíritus.
- Riesgo de morir en el menor de 7 días.
- $\frac{218}{151\,080} = 1.4$
- Riesgo de morir en los mayores de 28 días y menores del año de vida.
- Riesgo de morir de los menores de un año de vida.
- $\frac{10\,057}{6\,117\,424} = 1.6$

### **Bibliografía**

1. Daniel WW. Bioestadística. Base para el análisis de las ciencias de la salud. 3ª ed. México D.F.:Limusa; 1997.
2. Spiegel MR. Teoría y problemas de Estadística. La Habana:Pueblo y Educación; 1977.
3. Freund J. Estadística elemental moderna. 2ª ed. La Habana:Edición Revolucionaria; 1988.
4. Coolican H. Métodos de investigación y estadística en psicología. México D.F.:El Manual Moderno; 1997.
5. Anuario Demográfico de Cuba 1998. ONE. La Habana, 1999.
6. Swaroop S. Estadística Sanitaria. La Habana:Ed. Revolucionaria. 1964.

7. Camel F. Estadísticas médicas y de Salud Pública. La Habana:Pueblo y Educación; 1985.

## Tema 4 Medidas de resumen para variables cuantitativas

### Introducción

En el tema anterior viste cómo resumir las variables cualitativas, ya lo fuesen por naturaleza o porque manejaste los datos en escala cualitativa. La información cuantitativa también es dable de ser resumida, con las consabidas ventajas que de ello se generan, pues corrientemente es necesario representar un conjunto de datos por un número que —en la medida de lo posible— logre describir a dicho conjunto. Para obtenerlo, podrás disponer de tres grandes grupos de medidas de resumen: de tendencia central, de posición, y de variación, las cuales verás en los apartados siguientes.

### Antes de comenzar

En este Tema nos valdremos de algunas notaciones matemáticas para representar las fórmulas de cálculo de los distintos indicadores.

Con el ánimo de refrescarte la memoria, te mencionamos los elementos propios del lenguaje matemático que empleamos.

Para representar las distintas variables utilizamos las letras del alfabeto. Por ejemplo, si medimos la edad y la talla de cinco individuos, pudiésemos referirnos a la primera como *la variable X*, y a la segunda como *la variable Y*, de ahí que cada vez que hagamos alusión a *X*, sabemos que en realidad estamos hablando de la edad de las personas estudiadas (si fuera *Y*, entonces hablamos de la talla). Te aclaramos que para escoger las letras no hay una regla específica, eso queda a tu decisión.

Supongamos que los resultados del estudio fueron los siguientes:



Individuos	Edad (años) X	Talla (cm) Y
Ana	24	130.0
Juan	26	120.0
Rosa	27	140.0
Pedro	25	150.0
Teresa	23	110.0

Por otra parte, pudieras referirte a la edad por las letras *ed*, o algo por el estilo. Lo que debe quedar claro, es a qué te refieres con la simbología utilizada, pues **X** puede significar edad en un estudio, pero sexo en otro, por citar un ejemplo.

Ahora bien, llegamos a otro punto que necesita ser definido. Para referirte a las edades de los cinco individuos, y suponiendo que las representaste por la letra X, pudieras escribir entre otras tantas formas:

$$X_{(\text{Ana})} = 24$$

$$X_{(\text{Juan})} = 26$$

$$X_{(\text{Rosa})} = 27$$

$$X_{(\text{Pedro})} = 25$$

$$X_{(\text{Teresa})} = 23$$

Como habrás visto, es un procedimiento bastante tedioso el utilizar los paréntesis u otra forma similar de identificación de los datos. En su lugar, puedes usar los **subíndices**, resultando algo por el estilo:  $X_1 = 24$ ,  $X_2 = 26$ ,  $X_3 = 27$ ,  $X_4 = 25$ ,  $X_5 = 23$ . De la misma forma, las tallas serían:  $Y_1 = 130$ ,  $Y_2 = 120$ ,  $Y_3 = 140$ ,  $Y_4 = 150$ ,  $Y_5 = 110$ .

También resultan muy útiles los **subíndices generales** (por lo general son letras), los cuales hacen alusión a un grupo de valores de la variable. Por ejemplo, para decir que hay cinco valores de la edad, puedes escribir:  $X_i$  ( $i = 1, 2, 3, 4, 5$ ); o lo que es lo mismo:  $X_i$  ( $i = 1, \dots, 5$ ). Observa que lo anterior no dice que la variable toma los valores 1, 2, 3, 4 ó 5, sino que hay cinco mediciones de la misma, ¿claro?.

Supón ahora que sumas todos los valores de la variable peso, entonces escribirías lo siguiente:

$$X_1 + X_2 + X_3 + X_4 + X_5 = 650$$

En el caso que nos ocupa, resulta fácil y rápida esta forma de escritura, pero si fuesen, digamos, ¡200! valores...

En esta situación, se utiliza la letra griega sigma mayúscula  $\Sigma$ , que representa el símbolo de **sumatoria**, el cual antecede a la variable en cuestión y se acompaña de dos anotaciones: una encima y otra debajo, como lo siguiente:  $\sum_{i=1}^n X_i$ . Esto se lee como “la suma de las Xs desde i hasta n”, o sea, las Xs cuyo subíndice van desde los valores especificados en i hasta n.

Retomando el ejemplo anterior, hubieses escrito  $\sum_{i=1}^5 X_i$ . Si quisieras sumar solamente los tres valores del centro, entonces escribirías:

$$\sum_{i=2}^4 X_i.$$

De la misma forma,  $\sum_{i=1}^5 X_i Y_i$  significa  $X_1 Y_1 + X_2 Y_2 + X_3 Y_3 + X_4 Y_4 + X_5 Y_5 = (24 \cdot 130) + (26 \cdot 120) + (27 \cdot 140) + (25 \cdot 150) + (23 \cdot 110) = 16\,300$ .

Al mismo tiempo, te recordamos que **elevar al cuadrado** un número es multiplicarlo por sí mismo, y se representa por el supraíndice 2, o sea,  $13^2 = 169$  (porque  $13 \cdot 13 = 169$ ). Entonces,  $\sum_{i=1}^5 X_i^2$  es la representación matemática de lo siguiente —utilizando los datos del ejemplo—:  $X_1^2 + X_2^2 + X_3^2 + X_4^2 + X_5^2 = 24^2 + 26^2 + 27^2 + 25^2 + 23^2 = 576 + 676 + 729 + 625 + 529 = 3125$ .

La operación inversa de elevar al cuadrado no es dividirlo por sí mismo, sino **extraerle su raíz cuadrada**, que se representa por el símbolo de radical  $\sqrt{\quad}$ , quedando bajo la barra horizontal a lo que se le extrae la raíz cuadrada. De ahí que  $\sqrt{4} = \pm 2$ , porque  $2 \cdot 2 = 4$ , pero  $-2 \cdot -2 = 4$  también, por lo que debe especificarse en buena lid el símbolo  $\pm$ . Para algunos datos del ejemplo anterior:  $\sqrt{24} = \pm 4.89$ ;  $\sqrt{26} = \pm 5.09$ ;  $\sqrt{27} = \pm 5.19$ .

Por último, seguramente recordarás que el **valor absoluto o modular** de un número es él mismo sin el signo asociado, esto es, se toma la magnitud positiva del número. Se representa por dos

barras verticales que enmarcan al número deseado, v.g.  $|3| = 3$ , y  $|-3| = 3$ .

## 4.1 Medidas de tendencia central

Seguramente, lo primero que estarás preguntándote es: «¿Por qué de tendencia central?». Bueno, estriba en que ellas están constituidas por un número al que se acercan o “tienden” la mayoría de las observaciones. Con otras palabras, alrededor de él se agrupan las observaciones de la serie de datos, puesto que es en sí el centro de la serie, aunque ello no significa que este número tiene que estar representado en la serie (de hecho, muchas veces no ocurre así). Veamos en detalle cada una de estas medidas.

### 4.1.1 La media aritmética

Puedes encontrar la medida que te vamos a presentar con diversos nombres, entre ellos los más utilizados son **promedio**, **promedio aritmético** e incluso simplemente **media**. Al respecto debemos aclararte que existen otras medias que no son aritméticas, pero cuando decimos *media* a secas, nos referimos a la aritmética.

La media aritmética es una cifra que obtienes al sumar todos los valores observados y dividirlos por el número de valores. ¿No te resulta familiar? Claro, estás muy acostumbrado al cálculo del promedio. Se denota por los símbolos  $\mu$  (letra griega mu) ó  $\bar{x}$  (se lee equis media), la distinción entre estos dos símbolos se hará importante en cursos sucesivos, cuando abordemos temas de Estadística Inferencial. Por lo pronto, utilizaremos el último de ellos. La media conserva las unidades de medida de la variable en su estado original, o sea, que la media de un grupo de edades en años se expresará asimismo en años.

El cálculo de la media dependerá de cómo aparezcan los datos, de tal suerte que, para **datos simples**<sup>14</sup>, la fórmula es:

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

<sup>14</sup> O sea, datos no agrupados ni distribuciones de frecuencias.

Veamos un ejemplo. Supón que deseas conocer la estatura media de cinco adolescentes de tu consultorio. Estas son las observaciones (datos) de la medición de cada uno de ellos (en cm):

170.0            150.0            130.0            160.0            140.0

$$\bar{X} = \frac{170 + 150 + 130 + 160 + 140}{5} = 150 \text{ cm}$$

Este resultado indica que, **en promedio, los adolescentes miden 150 centímetros**. Sencillamente, no hemos hecho otra cosa que decir: «*más o menos, los muchachos miden 150 centímetros*».

Si los datos aparecen en una **distribución de frecuencias**, entonces calcula la media de la siguiente forma:

$$\bar{X} = \frac{\sum_{i=1}^n f_i \cdot X_i}{n}$$

donde:

- $f_i$  son las frecuencias absolutas;
- $X_i$  son las categorías de la variable;
- $n$ : total de observaciones.

Por ejemplo, estos son los pesos (en kilogramos) de 10 individuos representados por sus frecuencias absolutas:

Peso ( $X_i$ )	Número ( $f_i$ )	$f_i \times X_i$
60	3	180
61	1	61
62	2	124
63	3	189
64	1	64
<b>Total</b>	<b>10</b>	<b>618</b>

$$\bar{X} = \frac{618}{10} = 61.8 \text{ Kg}$$

Con este resultado puedes decir que **los 10 sujetos pesan 61.8 kilogramos** (o mejor, 62) **como promedio**.

Ahora bien, puede que los datos estén **agrupados** en una escala de intervalos, entonces el cálculo se realiza como verás. A continuación te mostramos los resultados de la frecuencia cardíaca (FC), en latidos cardíacos por minuto, de 100 pacientes ingresados en el Servicio de Medicina Interna de cierto hospital:

FC	Pacientes ( $f_i$ )
21 – 40	1
41 – 60	17
61 – 80	38
81 – 100	40
101 – 120	4
<b>Total</b>	<b>100</b>

La media aritmética de una serie de datos agrupados está dada por:

$$\bar{X} = \frac{\sum_{i=1}^n MC_i \times f_i}{n}$$

donde:

- $MC_i$  son las marcas de clase de los intervalos,
- $f_i$  son las frecuencias absolutas de los intervalos.
- $n$ : total de observaciones.

Procedamos pues, a calcularla:

1. Ante todo, necesitarás conocer las **marcas de clase** ( $MC_i$ ) de cada intervalo, para lo cual suma el límite inferior (LI) al superior (LS) del intervalo en cuestión, y luego divídelo por dos.
2. Luego, multiplicarás la marca de clase por la frecuencia absoluta ( $f_i$ ) del intervalo.
3. Sumas todos los números resultantes del paso anterior y, finalmente, divides la suma por el total de observaciones ( $n$ ).

Los cálculos pertinentes por intervalos son:

	Marcas de clase	MC x f <sub>i</sub>
1 <sup>er</sup> intervalo:	$[21 + 40]/2 = 30.5$	$30.5 \times 1 = 30.5$
2 <sup>o</sup> intervalo:	$[41 + 60]/2 = 50.5$	$50.5 \times 17 = 858.5$
3 <sup>er</sup> intervalo:	$[61 + 80]/2 = 70.5$	$70.5 \times 38 = 2679.0$
4 <sup>o</sup> intervalo:	$[81 + 100]/2 = 90.5$	$90.5 \times 40 = 3620.0$
5 <sup>o</sup> intervalo:	$[101 + 120]/2 = 110.5$	$110.5 \times 4 = 442.0$
	Total =	7630.0

Por último, tenemos:  $\bar{x} = \frac{7630}{100} = 76.3$ , por lo que puedes decir que, en promedio, los pacientes estudiados tenían una frecuencia cardiaca de 76 latidos por minuto.

Hay ocasiones en que los datos numéricos se hacen difíciles de manejar, bien porque sean muchos, o porque constan de varios dígitos. En esta situación, te aconsejamos agruparlos primero y luego calcular la media. Por otra parte, es bueno que recuerdes que el mayor monto de la información estadística aparece agrupada, por lo que te verás obligado a utilizar la fórmula estudiada.

Cometeríamos un grave error si no habláramos de las propiedades de la media. Ciertamente, entre las más notables tenemos que:



1. Es fácilmente entendible por la mayoría de las personas (o, al menos, es fácil de explicar su significado);
2. Siempre existe, y puede calcularse para cualquier dato numérico;
3. Es única, o sea, un grupo de datos sólo tiene una media;
4. Toma en cuenta a todos los valores de la serie de forma individual, esto es, recorre la serie completa.

Esta última resulta ser sumamente importante, pues la media calculada representa a todos los valores de la serie, siendo precisamente lo que se quería lograr. Ahora bien, no siempre esto resulta beneficioso, como verás en este ejemplo: imagínate que se deseaba saber la edad promedio de las personas reunidas en un salón de cierto Círculo Infantil, para lo cual se escogió al azar uno de

los que poseía dicho centro. En el momento de la medición, se encontraban presentes en el salón escogido siete bebés y la educadora que los cuidaba, siendo sus edades las siguientes (m: meses, a: años):

18m 10m 12m 16m 20m 12m 14m 34a

Edad media:  $510 \text{ meses} / 8 = 63.75 \text{ meses} = 5.3 \text{ años}$

¿Viste eso? Ahora tenemos que, en promedio, las personas allí reunidas tenían 64 meses de edad (¡Bueno, 5 años es algo más fácil de entender!) ¿Crees que sea cierto ese dato? Claro que no, está bastante lejos de la realidad, mas no está mal hecho el cálculo. Matemáticamente es impecable, pero la lógica dice que algo falló.

El motivo por el que apareció un resultado tan dispar es la presencia de un dato discordante en el conjunto: la edad de la educadora. Cuando en una serie de datos encuentras algún dato que se aparta de los demás de forma llamativa, entonces puedes nombrarlo *dato(s) aberrante(s)*. Si calculásemos la media con las edades de los pequeños solamente, entonces hubiese sido de 15 meses.

En resumen, si los datos son relativamente homogéneos, la media aritmética es una buena medida de resumen; pero si existen valores muy alejados de la mayoría (datos aberrantes), entonces se distorsiona mucho y deja de reflejar la realidad existente.

### 4.1.2 La mediana

Aquí tienes otra de las medidas de tendencia central. Al igual que la media, puedes utilizarla para describir el “centro” de un grupo de datos. No tiene un símbolo específico que la denote; nosotros usaremos *med* o *mediana* en lo adelante.



La **mediana** es la observación que divide a una serie ordenada de datos en dos partes iguales, o sea, es la observación que ocupa la posición central de una serie ordenada.

De lo antedicho se deduce que lo primero que tienes que hacer para calcular la mediana es *ordenar* la serie, ya sea en orden creciente o decreciente. Luego, buscarás cuál de los valores es la mediana, lo cual dependerá del número total de observaciones o datos que tengas.

De tener los **datos simples**, si tienes un número **n impar** de observaciones, la del centro es la medida buscada, como lo es 32 en esta serie: 41, 40, 36, **32**, 26, 21, 20. Fíjate que a ambos lados de la mediana hay la misma cantidad de números.

En este caso, por simple observación llegaste al resultado, pero puedes valerte de calcular  $\frac{n+1}{2}$  para saber **la posición** de la mediana, comenzando a contar por cualquiera de los dos extremos de la serie. En el ejemplo anterior el resultado es  $(7+1)/2 = 4$ , y el cuarto puesto lo ocupa el 32, no importa por cuál extremo comienzas a contar.

La contrapartida ocurre cuando el total de datos es un número **par**, entonces la mediana es la media aritmética de los valores del centro de la serie, como sucede en el ejemplo: 20, 24, 33, **39, 45**, 51, 75, 80. Los valores del centro son 39 y 45, su media es 42, y es este el valor de la mediana de esa serie.

No debe causarte extrañeza tal proceder, pues si aplicásemos la fórmula de la posición, entonces la mediana ocuparía el lugar  $(8+1)/2 = 4.5$ , esto es, la mitad entre los números 4 y 5 de la serie. Siendo el 39 y el 45 los lugares 4<sup>o</sup> y 5<sup>o</sup> respectivamente, entonces 42 es el centro entre ellos. ¿De acuerdo?

Se te puede presentar la situación de que tengas una serie con varios valores iguales, como 50, 54, 56, 56, 56, 56, 60, 62. Aquí la mediana es 56, claro está. Recuerda que ella es el valor central del grupo, y sería un atentado abierto a la lógica cuestionarse cuál de los 56s es la mediana.

También puedes calcular la mediana para **datos agrupados**. Supón que tienes las edades de 100 individuos de 20 a 54 años de tu área de salud:



Intervalo	Edad (años)	Frecuencia Absoluta	Frecuencia Absoluta Acumulada
1	20 – 24	25	25
2	25 – 29	12	37
3	30 – 34	14	51
4	35 – 39	9	60
5	40 – 44	10	70
6	45 – 49	12	82
7	50 – 54	18	100
	<b>Total</b>	<b>100</b>	—

Nota: la primera columna cumple funciones didácticas solamente, y la última contiene datos de trabajo.

Para el cálculo de la mediana en series agrupadas, se utiliza la siguiente expresión:

$$\text{med} = \text{LRI} + \left[ \frac{\frac{n}{2} - \text{FAA}}{f} \right] \cdot c$$

donde:

- LRI es el límite real inferior del intervalo que contiene a la mediana.
- n es el total de observaciones de la serie de datos.
- FAA es la frecuencia absoluta acumulada del IC que antecede al que contiene a la mediana.
- f es la frecuencia absoluta del IC que contiene a la mediana.
- c es la amplitud del IC que contiene a la mediana.

Ilustremos, a través del ejemplo anterior, los pasos que seguirás para hallar la mediana de la serie ordenada:

1. Calcula la mitad de las observaciones:  $n/2 = 100/2 = 50$ .
2. Determina el IC que contiene a la mediana: busca el IC cuya frecuencia absoluta acumulada (FAA) sea la primera que excede el número que acabas de obtener. En el caso que nos ocupa, será el 3<sup>er</sup> intervalo, ya que las FAA de los intervalos 1 y 2 son menores que 50.

3. Halla el límite real inferior (LRI) del intervalo de clase que contiene a la mediana, calculando la semisuma entre el límite superior del intervalo que le antecede (LSa) —que es el 2º intervalo— y el límite inferior del propio intervalo (Lip).

$$\text{LRI} = \frac{\text{LSa} + \text{Lip}}{2} = \frac{30 + 29}{2} = 29.5$$

4. Determina la FAA del IC que antecede al que contiene a la mediana. En este caso es 37.  
 5. Resta el resultado anterior a la mitad de las observaciones:  $50 - 37 = 13$ .  
 6. Ahora, calcula la amplitud del intervalo de clase que contiene a la mediana:

$$c = \text{límite real superior (LRS)} - \text{límite real inferior (LRI)}$$

$$\text{LRS} = [\text{LS propio intervalo} + \text{LI intervalo siguiente}] / 2 = (34 + 35) / 2 = 34.5$$

$$c = 34.5 - 29.5 = 5$$

7. Por último,  $f = 14$ .

Ahora, ya estás en condiciones de saber quién es la mediana del ejemplo:

$$\text{med} = \text{LRI} + \left[ \frac{\frac{n}{2} - \text{FAA}}{f} \right] \cdot c = 29.5 + \frac{13}{14} \cdot 5 \approx 34$$

Así las cosas, ya puedes decir que **la mediana de la serie en cuestión es de 34 años de edad, o que la edad mediana de la serie es 34 años.**

Esta medida de resumen posee las propiedades siguientes:



1. Su cálculo es sencillo;
2. Siempre existe, y puedes calcularla a cualquier conjunto de datos numéricos;
3. Es única;

4. Se puede calcular en series con límites abiertos, excepto cuando la propia mediana caiga en un límite abierto, pero esto es sumamente improbable; y
5. No se afecta fácilmente por valores extremos.

La cuarta y quinta propiedades hacen que se prefiera esta medida sobre la media en situaciones en que la escala sea abierta o que existan valores aberrantes. Ahora bien, en la mayoría de los casos — lógicamente, salvo los citados— se prefiere conocer la media como medida de tendencia central.

Para ilustrar lo planteado en la quinta propiedad, volvamos al ejemplo de las edades de los niños del Círculo Infantil y su educadora. Si calculamos la mediana de esos datos, ésta sería:

Datos ordenados: 10, 12, 12, 14, 16, 18, 20, 34

Mediana:  $(14 + 16) / 2 = 15$  meses, resultado que sí refleja con certeza el centro de los datos.

Quizá una desventaja imputable a la determinación de la mediana radica en el ordenamiento previo de las observaciones, faena que pudiese devenir tediosa y hasta impracticable de ser un número considerable de datos, pero recuerda las potencialidades que te brindan los softwares existentes en el mercado actual, que facilitan enormemente el trabajo<sup>15</sup> al calcular la mayoría de estos indicadores.

### 4.1.3 La moda

Ahora conocerás una medida realmente sencilla, tanto de determinar como de interpretar. Es muy intuitiva, y consiste en el valor, clase o categoría que aparece con más frecuencia en una serie de datos; o sea, es el que más se repite. Por ejemplo, si de seis pacientes, tres tienen 20 años, y los otros tienen 18, 21, y 25 respectivamente, entonces dirías que 20 años es la moda, o edad modal.

La mayor ventaja de la moda radica en que no requiere cálculo alguno, para beneplácito de algunos que no cuentan a las

<sup>15</sup> No tienes que acudir obligatoriamente a un paquete estadístico especializado: por ejemplo, Microsoft® Excel realiza cálculos como el que nos ocupa.

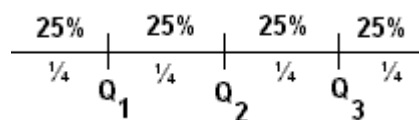
Matemáticas entre su círculo de amistades. Sin embargo, puede que no exista, e incluso puede no ser única. Por ejemplo, la serie 2, 5, 6, 6, 6, 7, 8, 8, 8, es una serie **bimodal**, pues cuenta con el seis y el ocho como modas.

## 4.2 Medidas de posición

Como recordarás, las medidas estudiadas con anterioridad se nombran de tendencia central por representar el centro del conjunto de observaciones. Verás ahora un grupo de medidas que establecen una posición en una serie ordenada de datos, son una referencia a partir de la cual podrás decir: por encima de este valor están las  $\frac{3}{4}$  partes de las observaciones, o algo por el estilo. De estas medidas, llamadas *cuantiles* por algunos, veremos a los **cuartiles**, **deciles** y **percentiles**.

### 4.2.1 Los cuartiles

Los cuartiles, representados por la letra Q, son valores que dividen una serie ordenada de datos en cuatro partes iguales (cuartos), de tal suerte que por debajo del primer cuartil ( $Q_1$ ) se encuentra  $\frac{1}{4}$  parte (el 25%) de los datos, y por ende, el 75% (las  $\frac{3}{4}$  partes) está por encima de ese cuartil. La mitad de las observaciones cae por debajo del segundo cuartil —y, lógicamente, la otra mitad está por encima de él—; y las  $\frac{3}{4}$  partes de los datos están por debajo del tercer cuartil, como muestra la figura.



**Figura 2. Los cuartiles**

Indudablemente, uno de estos nuevos amigos tiene un aire familiar con cierta conocida. ¿Ya te diste cuenta? Claro, es  $Q_2$ , quien coincide con la mediana.

Para hallar los cuartiles de una serie, seguiremos un procedimiento similar al utilizado para conseguir la mediana.

Para una serie de **datos simples**, el cuartil en cuestión  $Q_i$  será el valor que ocupe la posición  $\frac{n+1}{4}q_i$ , siendo  $q_i$  el cuartil deseado y  $n$  el total de observaciones.

Si los datos están **agrupados**, entonces el cuartil será el resultado de calcular:

$$Q_i = LRI + \left( \frac{\frac{nq_i}{4} - FAA}{f} \right) c$$

donde:

- $q_i$  es el cuartil deseado ( $i = 1, 2, 3$ );
- LRI: límite real inferior del intervalo que contiene al cuartil;
- $n$ : total de observaciones;
- FAA: frecuencia acumulada absoluta de la clase que antecede a la del cuartil;
- $f$ : frecuencia absoluta de la clase que contiene al cuartil;
- $c$ : amplitud del IC que contiene al cuartil.

Para calcular el primer cuartil, volvamos a auxiliarnos del ejemplo de las frecuencias cardiacas en ciertos pacientes ingresados.

Frecuencia Cardíaca	Pacientes	Frecuencia Absoluta Acumulada
21 – 40	1	1
41 – 60	17	18
61 – 80	38	56
81 – 100	40	96
101 – 120	4	100
<b>Total</b>	<b>100</b>	<b>-</b>

1. Busca en cual clase está el cuartil:  $\frac{nq_i}{4} = \frac{100 \cdot 1}{4} = 25$ ; con este valor y siguiendo el procedimiento que seguramente recuerdas de la mediana, rápidamente sabrás que el cuartil está en el tercer intervalo (porque su frecuencia acumulada absoluta es la primera en sobrepasar 25).

2. Los LR de ese IC son:  $LRI = (60 + 61) / 2 = 60.5$ ;  $LRS = (80 + 81) / 2 = 80.5$ . Luego:  $c = 80.5 - 60.5 = 20$ .
3.  $FAA = 18$ ,  $f = 38$ .
4. Determina el cuartil:

$$Q_1 = 60.5 + \left( \frac{25 - 18}{38} \right) 20 = 64.2$$

Bien, ya tienes el valor del cuartil, solo tienes que interpretarlo adecuadamente, lo cual lograrás si leíste el comienzo del acápite. **La cuarta parte de los pacientes, o el 25% de ellos, tiene 64 o menos latidos cardiacos por minuto; o puedes decir que el 75% ( $\frac{3}{4}$  partes de los pacientes) tiene más de 64 latidos cardiacos por minuto.**

#### 4.2.2 Los deciles

Veamos ahora los deciles. Muy parecidos a sus parientes los cuartiles, ellos son nueve ( $D_1, D_2, \dots, D_9$ ) que dividen a una serie ordenada de datos en diez partes iguales (décimos). Por ejemplo, por encima de  $D_6$  hay cuatro décimos, quedando seis décimos debajo de él, y así por el estilo. Curiosamente, el quinto decil coincide con el segundo cuartil, y, por consiguiente, con la mediana.

La posición del decil está dada por la expresión  $\frac{n+1}{10}d_i$  en una serie de **datos simples**, siendo  $n$  el total de observaciones, y  $d_i$  el decil deseado ( $i = 1, 2, 3, 4, 5, 6, 7, 8, 9$ ).

Para **datos agrupados**, usemos nuevamente el ejemplo de la frecuencia cardiaca. Los pasos para calcular el decil son en esencia iguales que si fueses a calcular el cuartil, haciendo los cambios siguientes:

$$D_i = LRI + \left( \frac{\frac{nd_i}{10} - FAA}{f} \right) c$$

donde:

- $d_i$  es el decil deseado ( $i = 1, 2, \dots, 9$ );
- LRI: límite real inferior de la clase o intervalo que contiene al decil;
- $n$ : total de observaciones;

- FAA: frecuencia acumulada absoluta de la clase que antecede a la del decil;
- f: frecuencia absoluta de la clase que contiene al decil;
- c: amplitud del IC que contiene al decil.

Calculemos el primer decil:

1.  $\frac{nd_1}{10} = \frac{100 \cdot 1}{10} = 10$ , por tanto, el decil está en el segundo intervalo.
2. LRI = 40.5; c = 20; FAA = 17.

$$D_1 = 40.5 + \left( \frac{10 - 1}{17} \right) 20 = 51.01$$

Y ya puedes decir que la décima parte de los pacientes (el 10% de ellos) tenía 51 o menos latidos cardiacos por minuto; o lo que es lo mismo, el 90% tenía una frecuencia cardíaca de más de 51 latidos por minuto.

¿Eso era todo? Mira, todo resultado, por muy bien calculado que esté, por muy laborioso o sencillo que haya sido llegar hasta él, la interpretación debe estar avalada por un conocimiento previo del problema en cuestión. Realmente, no es improbable que alguien te replicara, no sin cierta sorna: «*Médico, ¿pudiese Ud. explicarnos cómo nos las arreglamos para manejar frecuencias cardíacas impares?*» A lo que tendrías que hacer de tripas corazón y rectificar la cifra, escogiendo para ello el número par precedente o siguiente.

### 4.2.3 Los percentiles

Los percentiles son observaciones que dividen a una serie ordenada de datos en cien partes iguales, motivo por el que hay 99 de ellos. Sí, ya sabemos que el término te resultó conocido apenas lo viste. Casi de la familia. Pues sí, se trata de los mismos percentiles que utilizas, por ejemplo, para conocer la evaluación nutricional de uno de tus bebitos. Bueno, ahora verás cómo se construye un percentil, procedimiento que conoces en parte porque es muy análogo al utilizado para determinar las medidas de posición estudiadas.

Para una serie de datos simples, la posición del percentil estará dada por  $\frac{n+1}{100}p_i$ , siendo  $n$  el total de observaciones, y  $p_i$  el percentil deseado ( $i = 1, 2, \dots, 99$ ).

Para datos agrupados, el percentil deseado será el resultado de computar:

$$P_i = LRI + \left( \frac{\frac{np_i}{100} - FAA}{f} \right) c$$

donde:

- $p_i$  es el percentil deseado ( $i = 1, 2, \dots, 99$ );
- LRI: límite real inferior de la clase o intervalo que contiene al percentil;
- $n$ : total de observaciones;
- FAA: frecuencia acumulada absoluta de la clase que antecede a la del percentil;
- $f$ : frecuencia absoluta de la clase que contiene al percentil;
- $c$ : amplitud del IC que contiene al percentil.

Haciendo uso de nuestro fiel amigo, el ejemplo de las frecuencias cardiacas, calculemos el percentil 95 de la serie:

1.  $\frac{np_{95}}{100} = \frac{100 \cdot 95}{100} = 95$ , por lo que el percentil está en el 4º intervalo.
2. LRI = 80.5; FAA = 40;  $c = 20$ .

$$P_{95} = 80.5 + \left( \frac{95 - 56}{40} \right) 20 = 100$$

Con este dato, puedes afirmar que **el 95% de los pacientes examinados tenía una frecuencia cardíaca de 100 o menos latidos por minuto, o lo que es lo mismo, el 5% tenía frecuencias superiores a los 100 latidos por minuto.**

#### 4.3 Medidas de variabilidad



Hasta el momento has visto algunas medidas que proporcionan información de una serie de datos numéricos, con la característica que un solo número es el encargado de esto. Quizá podría parecerte que con esas medidas sería suficiente para resumir y describir los conjuntos de los cuales proceden, sin embargo, múltiples circunstancias exigen la descripción de otros rasgos de los datos existentes. Verás por qué planteamos esto.

Volviendo al ejemplo de la medición de la frecuencia cardíaca (FC, epígrafe 4.1.1), imagina que el médico a cargo de la sala (que no fue quien hizo el estudio) realice el siguiente análisis: «*tomando en cuenta la FC media, no tengo motivos para preocuparme por la salud de los pacientes, pues, en general, ostentan cifras dentro de límites normales; por ende, tenemos que encaminar nuestros esfuerzos hacia otros problemas*».

¿Qué opinas acerca de esto? Sí, estamos de acuerdo contigo: en principio, no está nada mal, ya que interpretó correctamente el indicador. Pero, hay algo que debes recordar: la realidad es que, hasta donde él sabe, puede que *la mayoría* de los pacientes sea la que tiene frecuencias cardíacas en los límites considerados normales; pero puede que sea solamente la mitad, mientras la otra mitad permanezca en franca bradicardia (o taquicardia). Claro, este es uno de los casos “extremos”, mas el hecho de que sea algo *poco probable* no significa que sea *imposible*. De hecho, nota que en el ejemplo, el intervalo de clase que ostenta mayor frecuencia absoluta (40) es el de 81 a 100 latidos por minuto, o sea, no es precisamente el que contiene a la media.

Veamos otra situación. Fíjate en estas dos series cuyas medias y medianas coinciden (53). Sin embargo, no podríamos decir que son semejantes, si tomamos en cuenta sus datos:

Serie A: 50, 51, 52, **53**, 54, 55, 56

Serie B: 23, 33, 43, **53**, 63, 73, 83

Por supuesto que no podemos. En la serie A los datos están más juntos, están más cercanos (distan una unidad entre sí), lo que no ocurre en la serie B, cuyos datos están mucho más alejados entre sí. Si analizas las distancias entre las observaciones y su media, en la serie A la primera y última observaciones están a tres unidades del

centro de la serie, mientras que en la otra esas observaciones distan 30 unidades de su media. ¿Claro?

Ante una situación parecida, se necesitan otros parámetros acerca de la serie, como aquellos que miden cuán alejadas o agrupadas están las observaciones unas de otras o de la media.

El grado de agrupación o alejamiento de los datos de una serie es lo que recibe el nombre de **variabilidad, variación, esparcimiento o dispersión** de los datos, la cual puede ser **absoluta o relativa**. Con el fin de estudiarla, esto es, de conocer hasta qué punto las observaciones están agrupadas o esparcidas, la Estadística nos facilita las **medidas de variabilidad, variación, esparcimiento o dispersión**, las cuales verás en este epígrafe.

### 4.3.1 Medidas de variabilidad absoluta

#### 4.3.1.1 La amplitud

Esta es la más sencilla de las medidas de dispersión. Consiste en **la diferencia entre el mayor valor de la serie y el menor**, o sea, restar ambos valores. Por ejemplo, la amplitud de los datos de la serie A —vista en el acápite anterior— es de  $56 - 50 = 6$ , y la de la serie B es de 60. Ya conoces esta medida

Como ves, es muy fácil de determinar. Da una descripción *rápida* de la variabilidad de un grupo de observaciones. No es muy descriptiva de la misma porque sólo toma en cuenta los valores extremos de la serie, mira el siguiente ejemplo:

Serie A: 150, 160, 170, 180, 190, 200

Serie B: 150, 190, 197, 198, 199, 200

Resulta obvio que ambas tienen la misma amplitud (50), pero salta a la vista que ambas *no tienen la misma dispersión*. En la primera, los valores se sitúan de forma bastante dispersa entre los extremos; en la segunda la mayoría está cercana al valor mayor de la serie.

Si queremos describir de manera más eficiente la dispersión de un cúmulo de datos, tendremos que echar mano a otra medida. Veamos qué sorpresa nos depara la Sra. Estadística.

### 4.3.1.2 La desviación media

Al hablar de dispersión, lo hacemos la mayoría de las veces tomando en cuenta a la media aritmética del conjunto de observaciones. Así, cuando decimos que la dispersión de una serie es *pequeña*, es porque los datos están agrupados en la cercanía de su media, siendo *grande* si los datos están alejados de ella. Esto sienta las bases para utilizar como referencia a las **distancias** para referirnos a la dispersión, o sea, que sería procedente definirla en función de las **distancias que existen entre los números y su media**. Ahora bien, estas distancias pueden ser vistas como la **desviación** entre los elementos en cuestión; con otras palabras, si hay mucha distancia, decimos que se desvió mucho el número de la media, y así por el estilo.

Si promediásemos las desviaciones entre cada número y la media, o sea, sumarlas y dividir las por la cantidad de números de la serie, obtendríamos una medida de la variación promedio del conjunto de datos dada por:

$$\frac{\sum_{i=1}^n (X_i - \bar{X})}{n}$$

Por desgracia, realizar este cálculo te sería tan provechoso como no hacer ninguno, pues el resultado final siempre es cero<sup>16</sup>, debido a razones matemáticas establecidas. Por ejemplo, considera la siguiente serie: 2, 3, 4, 5, 6. Su media es 4, y si calculásemos lo planteado:

$$\frac{(2-4)+(3-4)+(4-4)+(5-4)+(6-4)}{5} = \frac{(-2)+(-1)+0+1+2}{5} = \frac{0}{5} = 0$$

La solución a este inconveniente es hallar la diferencia modular de las desviaciones, de esa manera sólo tomarás en cuenta la magnitud de dichas desviaciones, esto es, hallar el módulo de las diferencias. De esta forma, estarás calculando la **desviación media (DM) o desviación promedio**, cuya fórmula para datos simples es:

<sup>16</sup> Estriba en que la media de las desviaciones es justamente la media misma. Como escapa a los propósitos de este curso entrar en formalismos matemáticos, si estás interesado puedes acudir a la literatura especializada para indagar al respecto.

$$DM = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n}$$

donde:

- $X_i$  son las observaciones de la serie ( $i = 1, \dots, n$ );
- $\bar{X}$ : media aritmética de la serie;
- $n$ : total de observaciones.

Ilustremos lo antedicho calculando la desviación media de las siguientes mediciones del peso (en libras) que corresponden a cinco estudiantes de tu área de salud.

150.5, 180.8, 145.3, 127.9, 130.5

Ante todo, calcula la media:

$$\bar{X} = \frac{150.5 + 180.8 + 145.3 + 127.9 + 130.5}{5} = 147 \text{ libras}$$

Luego halla las desviaciones de cada observación con respecto a su media:

$X_i$	$ X_i - \bar{X} $
150.5	3.5
180.8	33.8
145.3	1.7
127.9	19.1
130.5	<u>16.5</u>
	<u>74.6</u>

Finalmente,  $DM = 74.6 \div 5 = 14.92$  libras.

Calculada la medida deseada, puedes decir que, en promedio, el peso de los estudiantes se desvía casi 15 libras del promedio general de 147 libras.

Si los datos de que dispones están agrupados, entonces calcularás la desviación media de la siguiente forma:

$$DM = \frac{\sum_{i=1}^n |MC_i - \bar{X}| \cdot f_i}{n}$$

donde

- $MC_i$ : marca de clase de cada IC;
- $f_i$ : frecuencia absoluta de cada intervalo de clase;
- $\bar{x}$ : media aritmética de la serie;
- $n$ : total de observaciones.

Utilizando nuevamente el ejemplo del acápite 4.1.1, calculemos la desviación promedio de esos datos.

( $\bar{x} = 76.3$  latidos/minuto)

FC	Pacientes ( $f_i$ )	$MC_i$	$ MC_i - \bar{x} f_i$
21 – 40	1	30.5	$(30.5 - 76.3) = 45.8$
41 – 60	17	50.5	$(50.5 - 76.3) = 438.6$
61 – 80	38	70.5	$(70.5 - 76.3) = 220.4$
81 – 100	40	90.5	$(90.5 - 76.3) = 568.0$
101 – 120	4	110.	$(110.5 - 76.3) = 136.8$
		5	
Total	100		1409.6

Entonces,  $DM = 1409.6 \div 100 = 14.09$  latidos por minuto, lo que significa que **las frecuencias cardiacas de los pacientes de alejan, en promedio, 14 latidos por minuto de la media de 76.**

Particularicemos en algo: si se tratase de una serie con valores aberrantes, debiste calcular la mediana en vez de la media; entonces debes usar la mediana para calcular la desviación media, sustituyendo la media en la fórmula por la mediana.

La desviación media es una medida que se utiliza poco en la práctica, sobre todo si son muchos datos o si éstos tienen muchos lugares decimales, pero principalmente debido a razones de índole matemática que no abordaremos en este curso. De todas formas, optamos por dártela a conocer con el fin de que lograras entender cabalmente las medidas que verás a continuación.

### 4.3.1.3 La varianza y la desviación estándar

Como recuerdas, para calcular la desviación media te viste obligado a utilizar las diferencias modulares para obtener un resultado válido, pues de lo contrario no hubieses obtenido nada. Pues bien, existe otra medida que se vale de elevar al cuadrado las desviaciones de los datos con respecto a su media. Dicha medida recibe el nombre de **varianza o variancia**. Se denota por los símbolos  $S^2$  ó  $\sigma^2$  (letra griega sigma minúscula al cuadrado), al igual que con la media, la distinción entre ellos se hará importante en temas de Estadística Inferencial. De momento, usaremos el primero. Su cálculo (para **datos simples**) se verifica según la fórmula:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

donde:

- $X_i$  son las observaciones de la serie ( $i = 1, \dots, n$ );
- $\bar{x}$ : media aritmética de la serie;
- $n$ : total de observaciones.

Esta medida logra describir adecuadamente la dispersión del conjunto de datos, pero tiene un inconveniente: su resultado se expresa en unidades cuadradas, algo harto engorroso y difícil de entender en la mayoría de las situaciones prácticas, y por demás disonante en relación con la medida de tendencia central utilizada. Sería algo así como años cuadrados, o pesos cuadrados (¿?).

A fin de eliminar este aparente escollo, puedes hallar la raíz cuadrada positiva del número obtenido, con lo que tendrás de vuelta a las unidades originales, obteniendo así una medida denominada **desviación típica o estándar**<sup>17</sup>, y es la medida de variación más ampliamente utilizada en el mundo de las estadísticas. Su símbolo es **S** (por ser la raíz cuadrada de la varianza), aunque se utiliza también **DS** (desviación standard) o **SD** (standard deviation). Tiene, además, la ventaja de que hasta las calculadoras de bolsillo —las científicas, claro está— la calculan, y casi la totalidad de los paquetes estadísticos existentes en el mercado del software.

<sup>17</sup> Aunque en muchos textos se utiliza standard, preferimos utilizar la traducción al castellano del término.

Vamos a calcular la desviación estándar de la serie del ejemplo de los pesos (en libras) del acápite anterior.

150.5, 180.8, 145.3, 127.9, 130.5

$\bar{x} = 147$  libras

$X_i$	$(X_i - \bar{x})^2$
150.5	12.25
180.8	1142.44
145.3	2.89
127.9	364.81
130.5	<u>272.25</u>
	1794.64

$$S^2 = 1794.64 \div 5 = 358.93 \text{ libras}^2$$

$$S = \sqrt{358.93} = 18.94 \text{ libras}$$

Con el resultado obtenido puedes decir que, en promedio, la mayoría de los datos se desvían de la media en casi 19 libras.

En gran parte de las situaciones del mundo biomédico, y basándose en elementos de la Estadística Inferencial<sup>18</sup>, se puede utilizar la desviación estándar y la media para construir intervalos en los que se mueve la mayor parte de los datos. Por ejemplo, en el intervalo cuyo extremo inferior sea  $\bar{x} - SD$  y el superior sea  $\bar{x} + SD$ , o sea,  $[\bar{x} - SD; \bar{x} + SD]$ , se encuentra cerca del 68% del total de las observaciones. Si construyes el intervalo con el duplo de la SD:  $[\bar{x} - 2 \cdot SD; \bar{x} + 2 \cdot SD]$ , entonces ahí estará cerca del 95% de los datos; y si utilizas el triplo de la SD, el intervalo contendrá entonces a más del 99% (99.73%) de las observaciones.

Si los datos están **agrupados**, el cálculo de la varianza se realiza mediante la fórmula:

$$S^2 = \frac{\sum_{i=1}^n (C_i - \bar{x})^2}{n}$$

donde

<sup>18</sup> Nos referimos a la normalidad de las poblaciones. Este es un tema que va más allá de los objetivos del presente curso.

- $MC_i$ : marca de clase de cada IC;
- $f_i$ : frecuencia absoluta de cada intervalo de clase;
- $\bar{x}$ : media aritmética de la serie;
- $n$ : total de observaciones.

Los elementos utilizados en el cálculo son los mismos que ya dijimos. El cálculo de la desviación estándar se limita a extraer la raíz cuadrada del número obtenido arriba.

Calculemos la desviación estándar para el ejemplo de las frecuencias cardiacas (media = 76.3 años).

FC	Pacientes	$(MC - \bar{x})^2 f_i$
21 – 40	1	$(30.5 - 76.3)^2 = 2097.64$
41 – 60	17	$(50.5 - 76.3)^2 = 11315.88$
61 – 80	38	$(70.5 - 76.3)^2 = 1278.32$
81 – 100	40	$(90.5 - 76.3)^2 = 8065.60$
101 – 120	4	$(110.5 - 76.3)^2 = 4678.56$
Total	100	27436.00

$$S^2 = 27\,436 \div 100 = 274.36$$

$$S = 16.5$$

Esto significa que, **en promedio, la mayoría de los pacientes se aleja de la media en 16 latidos por minuto.**

Suponiendo que a esta población se le pueden aplicar los porcentajes mencionados, entonces podríamos decir que aproximadamente el 95% de los pacientes tenían una frecuencia cardiaca que oscilaba entre 44 y 109 latidos por minuto.

Si la serie posee valores aberrantes, te viste obligado a utilizar la mediana, por lo que ahora debes sustituir la media por la mediana en la fórmula para calcular la varianza y la desviación típica.

#### 4.3.2 Medidas de variabilidad relativa

En muchas ocasiones es necesario comparar la dispersión entre dos o más conjuntos de datos, y sucede que las variables tienen



diferentes unidades de medida. Con las medidas de dispersión estudiadas no podrás llegar a una conclusión válida acerca de las desviaciones de los datos. Incluso, aún cuando se trate de una sola unidad de medida, las mediciones pueden variar considerablemente: si comparas la desviación típica de la estatura de los niños de 5-14 años de tu área con la de los estudiantes de preuniversitario, es muy probable que esta última sea mayor que la primera, debido a que las tallas sean mayores *per se*, y no porque la variabilidad sea mayor precisamente.

#### 4.3.2.1 El coeficiente de variación

Ante casos como el descrito con anterioridad, es imprescindible contar con una medida de variabilidad relativa, como el **coeficiente de variación (CV)**, que expresa a la desviación típica como porcentaje de la media, y su cálculo se realiza mediante:

$$CV = \frac{SD}{\bar{X}} 100$$

Observa que, por tener la desviación estándar y la media las mismas unidades de medida, quedan canceladas dichas unidades, de ahí que el coeficiente de variación no tenga unidades propias<sup>19</sup>, lo que facilita la comparación.

En el ejemplo siguiente, si comparas las desviaciones estándares de los dos grupos, pudieras creer que ambos tienen igual dispersión:

Grupo 1: media = 60 cm; SD = 4 cm  
Grupo 2: media = 170 cm; SD = 4 cm

Sin embargo, si calculas la medida recién conocida, entonces:  $CV_1 = 6.6$ ,  $CV_2 = 2.3$ . Al contrastarlos, ves algo bien diferente, pues en realidad el grupo 1 tiene casi tres veces más dispersión que el grupo 2.

## Resumen

En este tema aprendiste que:

---

<sup>19</sup> Con otras palabras, el coeficiente de variación es adimensional.

1. Las medidas de resumen para variables cuantitativas suelen dividirse en: **de Tendencia Central, de Posición y de Variabilidad**; resultando conveniente no utilizar las primeras de forma aislada, sino acompañadas de alguna medida de la dispersión.
2. La **media aritmética** es la medida de tendencia central más utilizada; en tanto que la **desviación estándar** sobresale por su uso entre las medidas de variabilidad.
3. Cuando las distribuciones son asimétricas o hay intervalos abiertos, se prefiere el cálculo de la mediana al de la media.
4. Cuando las distribuciones son simétricas, coinciden la media, la mediana, la moda, el segundo cuartil, el quinto decil y el percentil cincuenta.
5. Se prefiere utilizar como medida de dispersión a la **desviación media** cuando se usa como medida de tendencia central la mediana. Si utilizaste la media, generalmente se elige la **desviación estándar**.

## Ejercitación

1. Las cifras siguientes corresponden a la estatura (en centímetros) de los estudiantes de un aula de cierta escuela. La Dra. Fonseca, a cargo de dicho centro, desea resumir esta información con miras a redactar un informe, en el cual deben aparecer consignadas las interpretaciones correspondientes a los resultados de la media aritmética, la mediana, la moda, el primer cuartil, los deciles cuatro y ocho, los percentiles 25, 75 y 93, el rango, la desviación media, la desviación típica y el coeficiente de variación. ¿Podrías ayudar a la atribulada doctora en su empeño?

135	134	136	136	134
146	146	147	147	146
123	117	125	124	122
151	148	153	152	148
138	138	138	138	138
140	138	143	141	139
131	219	132	132	129
145	144	146	146	145
128	126	128	128	127
137	137	137	137	137

2. En una investigación sobre la incidencia de la hipertensión arterial en jóvenes de 20-24 años de un municipio de la provincia Ciego de Ávila, se estudiaron 350 jóvenes, de ellos 200 eran varones. A continuación te presentamos un cuadro con la información obtenida.

Tabla 1. Distribución de jóvenes según tensión arterial diastólica (TAD). Municipio Ciego de Ávila, 1998.

<b>TAD (mm Hg)</b>	<b>No.</b>	<b>%</b>
<b>&lt; 70</b>	136	38.9
<b>70 – 79</b>	73	20.9
<b>80 – 89</b>	126	36
<b>90 – 99</b>	10	2.9
<b>100 y más</b>	5	1.4
<b>Total</b>	350	100.0

- a) Determina la media aritmética, la mediana, la desviación media y la desviación estándar.  
b) Interpreta los resultados obtenidos.

3. Del mismo estudio, se obtuvo la siguiente información por sexos:

Tabla 2. Jóvenes según tensión arterial diastólica (TAD) y sexo. Municipio Ciego de Ávila, 1998.

<b>TAD (mm Hg)</b>	<b>Sexo</b>				<b>Total</b>	
	<b>Masculino</b>		<b>Femenino</b>		<b>No.</b>	<b>%</b>
	<b>No.</b>	<b>%</b>	<b>No.</b>	<b>%</b>	<b>No.</b>	<b>%</b>
<b>&lt; 70</b>	68	34.0	68	45.3	136	38.9
<b>70 – 79</b>	46	23.0	27	18.0	73	20.9
<b>80 – 89</b>	76	38.0	50	33.3	126	36
<b>90 – 99</b>	5	2.5	5	3.3	10	2.9
<b>100 y más</b>	5	2.5	0	0	5	1.4
<b>Total</b>	200	100.	150	100.	350	100.
		0		0		0

- a) Determina e interpreta la media aritmética, la mediana, la desviación media y la desviación estándar para cada grupo.  
b) ¿Qué puedes decir acerca de la variabilidad de cada grupo?

- c) ¿Para qué valor se encuentra el 70% de los varones por encima de él?
- d) ¿Para qué valor el 25% de las hembras está por debajo de él?
- e) Calcula el segundo cuartil, el quinto decil y el cincuenta percentil partiendo de los datos de la tabla 1. Compáralos con la mediana que ya calculaste. ¿A qué atribuyes los resultados?

**Autoevaluación**

1. El Dr. Quevedo afirma que casi el 95% de las 50 trabajadoras de su área de salud tienen entre los 26 y los 37 años de edad, y que el 45% tiene más de 30 años. A juzgar por los datos siguientes, ¿crees que nuestro docto amigo está en lo cierto?

25	55	49	41	50
29	54	50	35	47
44	19	47	22	28
47	26	43	19	20
60	24	26	42	53
48	35	28	44	19
21	33	48	47	26
20	70	60	60	24
19	24	55	48	35
22	38	49	53	54

2. Un grupo de investigadores obtuvo información sobre la glicemia en ancianos de una comunidad, y desean calcular algunas medidas que les permitan resumir la información. A continuación te presentamos lo que obtuvieron:

Tabla 1. Ancianos según edad y glicemia. Comunidad X, 1999.

Edad	Glicemia (mmol/L)						Total	
	3 – 5.5		6 – 10		> 10		No.	%
	No.	%	No.	%	No.	%		
<b>60 – 69</b>	7	28.0	11	35.5	14	28.0	32	30.2
<b>70 – 79</b>	5	20.0	9	29.0	28	56.0	42	39.6
<b>80 y más</b>	13	52.0	11	35.5	8	16.0	32	30.2
<b>Total</b>	25	100.0	31	100.0	50	100.0	106	100.0

- a) Calcula e interpreta las medidas de tendencia central y dispersión que puedan ser aplicadas.
- b) Calcula e interpreta  $Q_3$ ,  $D_4$  y  $P_{75}$ .

## **Bibliografía**

1. Daniel WW. Bioestadística. Base para el análisis de las ciencias de la salud. 3ª ed. México D.F.:Limusa; 1997.
2. Anuario Estadístico de la República de Cuba. 1998. La Habana:UNICEF-FNUAP; 1999.
3. Spiegel MR. Teoría y problemas de Estadística. La Habana:Pueblo y Educación; 1977.
4. Camel F. Estadísticas médicas y de Salud Pública. La Habana:Pueblo y Educación; 1985.

## Tema 5. Estadísticas de Salud Pública

### Introducción

La información es un elemento esencial en la toma de decisiones. En el sector salud, constituye un recurso clave y un requisito previo esencial para el suministro eficaz y la gestión de la atención de salud. El acceso adecuado a la información se reconoce como un ingrediente fundamental para los servicios de salud y para la planificación, funcionamiento, control y supervisión de programas; además es una herramienta para la evaluación de las actividades en salud y los resultados de las intervenciones clínicas y de gestión.

La información en salud puede considerarse como aquella relacionada de alguna manera con la salud, de aquí que es de índole sumamente variada. Comprende, por ejemplo, datos demográficos; información sobre determinantes sociales, culturales, económicos y ambientales de la salud; perfiles de morbilidad y mortalidad por enfermedades específicas; resultados provenientes de la práctica clínica o investigación biomédica y epidemiológica; estadísticas sobre las actividades del servicio de atención de salud, acciones del personal de salud y cobertura de programas de salud. Pero también comprende cosas como registros y archivos de pacientes, etc.

En tu labor como médico de familia necesitas usar y generar mucha de esta información, por ejemplo, debes conocer la comunidad que atiendes, sus características, composición por edad y sexo, sus problemas, los recursos con que cuentas para enfrentarlos y las posibles soluciones; además debes registrar, resumir y analizar gran cantidad de información numérica que te permita, entre otras cosas, diagnosticar el estado de salud de la población. Todos estos indicadores, unido al conjunto de métodos y procedimientos para su elaboración, análisis e interpretación, integran el contenido de las **estadísticas de salud**.

¿ Alguna vez te has sentido sobrecargado de trabajo burocrático?, ¿ Has pensado que la información que debes notificar es excesiva e incluso innecesaria?, quizás no conoces que la mayor parte de las estadísticas del sistema de salud pública se genera precisamente en la Atención Primaria, donde casi la totalidad de los programas de

salud son por ti ejecutados y por ende tú eres uno de sus eslabones principales, el registrador por excelencia de lo que detectas, ejecutas y modificas, además desconoces en muchas ocasiones, la importancia real y la utilidad de esa información.

En este capítulo te presentamos una serie de definiciones, técnicas y procedimientos acerca de las estadísticas de salud, que te serán muy útiles en el quehacer diario, en la realización de un trabajo investigativo e incluso para la toma de decisiones, en el caso que te desempeñes como cuadro administrativo.

### **5.1 - Estadísticas sanitarias: Concepto e Importancia**

En el curso anterior abordamos el término Estadística, en singular, como la ciencia encargada de recolectar, procesar, presentar, analizar e interpretar datos numéricos; ahora nos referiremos a la segunda acepción, estadísticas en plural, la cual se refiere a datos numéricos de cualquier índole. El caso particular de las estadísticas de salud se define como:



#### **Estadísticas de salud**

Es la información numérica, cuantificable, que sirve para conocer el estado de salud de la población, con la finalidad de planificar, evaluar y controlar programas y acciones que realiza el Sistema Nacional de Salud.

Como puedes ver, del concepto anterior se desprende la importancia y utilidad de las estadísticas sanitarias. Éstas son necesarias para todas las etapas del proceso administrativo, desde la etapa de planificación y organización de los programas hasta el control y la evaluación, son útiles además, en la docencia tanto de pregrado como de postgrado, para la investigación e incluso para otras ramas de la economía.

En la etapa de planificación de un programa de salud, facilitan el diagnóstico del problema, su magnitud, posibles causas y los recursos con que contamos para resolverlo. En la etapa de ejecución facilitan la aplicación de las medidas previamente planificadas. En la etapa de evaluación permiten valorar el cumplimiento de los planes originales, perfeccionar los programas y la distribución de recursos, de acuerdo a los resultados obtenidos.

Cuando se conciben y utilizan de manera adecuada, las estadísticas sanitarias, tienen el potencial de producir valiosa información clínica y administrativa orientada a la gestión para el funcionamiento de servicios y programas y para la atención del paciente. Dicha información es indispensable para muchas áreas a saber: decisiones eficaces y eficientes para las actividades, planificación, operación, supervisión y control de la atención de la salud individual y colectiva, en la evaluación y monitoreo de la situación de salud de grupos humanos, en la medición de los resultados de las intervenciones clínicas y de gestión, en la ampliación de la cobertura de los servicios y mejor calidad de la atención, en el mejoramiento y apoyo de las actividades de promoción de la salud, en un mayor acceso al conocimiento y como herramienta en la educación para la salud.

Las estadísticas sanitarias tienen uso individual y estadístico. Cuando una persona acude a una institución de salud, en los registros médicos de dicho centro queda recogida la información referente a los datos personales, historia de la enfermedad, tratamientos empleados, etc. El uso de estos registros médicos y otros como los de defunción y de nacimientos con fines personales, constituye el uso individual. Mientras, que el uso estadístico se refiere al manejo de datos agrupados, es decir, la suma de los registros individuales, los cuales pueden ser usados con diferentes objetivos, entre los más frecuentes se encuentran los siguientes:

- Describir el nivel de salud de una comunidad
- Encontrar soluciones a los problemas de salud
- Determinar prioridad para los programas de salud
- Dirigir o mantener el control durante la ejecución de estos programas
- Determinar el alcance y restricción de las demandas de salud
- Difundir información confiable acerca de la situación y los programas de salud



- Determinar éxito o fracaso de los programas específicos de salud
- Hacer evaluaciones generales del trabajo de salud pública

Interrogantes como las que te muestro a continuación, a las que puedes dar respuestas con el uso adecuado de las estadísticas de salud, constituyen ejemplos claros que te ayudarán a comprender mejor lo necesarias que éstas son.

- ¿Cuáles son las principales causas de muerte en un área?
- ¿A qué edad resulta más alta la mortalidad?
- ¿En qué áreas determinada enfermedad presenta una incidencia mayor que la incidencia promedio? ¿Qué condiciones prevalecen en dicha área?
- ¿Qué reducción cabría esperar como resultado de la aplicación de un programa de salud dirigido a modificar aquellas condiciones?
- Son suficientes los recursos invertidos para llevar a cabo los planes de salud concebidos?

A éstas y muchas otras interrogantes que te surjan como resultado de tu actividad cotidiana, podrás responder con el uso adecuado de las estadísticas sanitarias, como te dijimos anteriormente.

## 5.2 Clasificación de las Estadísticas de Salud.

Dentro de las estadísticas sanitarias se agrupa un conjunto de otras estadísticas particulares que pueden ser consideradas como su clasificación, éstas son:

**Estadísticas vitales:** Información numérica sobre los hechos vitales que ocurren en la población, es decir, aquellos eventos relacionados con el inicio y fin de la vida, así como con los cambios en el estado civil. Por ejemplo: nacimientos, defunciones, matrimonios.

**Estadísticas de morbilidad:** Información numérica sobre las enfermedades, secuelas, accidentes, en fin, los padecimientos que afectan a la población, y de la medida y forma en que lo hacen. Por ejemplo: número de hipertensos de un área, fumadores, discapacitados.

**Estadísticas de población:** información numérica acerca de la composición y principales características de las poblaciones

humanas. Por ejemplo: estructura poblacional según edad, sexo, raza, distribución espacial o geográfica.

**Estadísticas de recursos:** información numérica sobre los medios materiales y humanos que posee el Sistema Nacional de Salud para ejercer su gestión. Por ejemplo: número de hospitales, médicos por habitantes.

**Estadísticas de servicios:** información numérica acerca de la calidad y cantidad de las diferentes acciones que realiza el Sistema Nacional de Salud. Por ejemplo: consultas médicas por habitantes, pruebas citológicas realizadas.

**Estadísticas de vivienda:** información numérica sobre la cantidad y calidad de las viviendas. Por ejemplo: estado constructivo de la vivienda, número de viviendas inhabitables.

**Estadísticas de saneamiento:** información numérica acerca de las condiciones ambientales y sanitarias del entorno donde las comunidades humanas viven y realizan sus actividades. Por ejemplo: procedencia del abastecimiento de agua, número de vertederos de basura.

**Estadísticas económicas:** información numérica relacionada con la actividad económica en salud. Por ejemplo: los costos, gastos, exportaciones.

Más adelante en el tema, abordaremos algunas de ellas, por considerarlas como las más frecuentemente empleadas en el contexto de la Atención Primaria de Salud, pero antes queremos hacer un alto necesario en las técnicas más frecuentemente empleadas en la recolección de la información.

### 5.3 Principios básicos de recolección de la información

En la recolección de la información hay dos elementos característicos y diferentes en esencia que debes conocer, estos son el origen de ésta (fuente) y la manera en que es obtenida la misma (método empleado). Llamaremos fuente de obtención de la información al elemento que la origina y como viste en el curso anterior, puede ser primaria o secundaria.



### Fuentes

**Primaria:** Aquella en la que el investigador obtiene directamente la información del elemento que la origina.

**Secundaria:** Aquella que existe independientemente del estudio y el investigador sólo la utiliza.

Cuando se utiliza una fuente primaria para recolectar la información, pueden emplearse dos procedimientos o métodos fundamentales: la observación y el interrogatorio. Veamos en qué consisten cada uno de ellos y sus principales ventajas y desventajas.

#### 5.3.1 La observación.

La observación constituye el método clásico de obtención de datos, nos permite conocer la realidad objetiva mediante la percepción directa del objeto de estudio. Debemos diferenciar la observación científica, que es a la que hacemos referencia, de la observación espontánea o casual.

**La observación científica** consiste en la percepción directa del objeto de estudio realizada de forma consciente, planificada y objetiva. Es decir, se orienta hacia un fin determinado, debe ser cuidadosamente planificada teniendo en cuenta además de los objetivos, el objeto y el sujeto de la observación; y debe estar despojada lo más posible de todo elemento de subjetividad.

Entre sus ventajas podemos señalar las siguientes:

- Nos permite obtener conocimiento acerca del comportamiento del objeto o fenómeno estudiado tal como se presenta en la realidad, es decir, de manera objetiva.
- Es una forma de acceder a la información de manera directa e inmediata.
- Estimula la curiosidad e impulsa el desarrollo de nuevos hechos de interés científico.
- Puede utilizarse en compañía de otros procedimientos o técnicas, como la entrevista y el cuestionario, lo que permite la comparación de la

información obtenida por las diferentes vías empleadas y alcanzar mayor precisión.

Pero, si bien la observación hace posible investigar el fenómeno directamente, en su manifestación más externa, éste no permite llegar a la esencia del mismo, a sus causas, de ahí que en la práctica, junto con ella suelen emplearse otros métodos como la medición y el experimento. Ésta constituye la desventaja principal de la observación, otras desventajas son las siguientes:

- Requiere personal especializado.
- No es conveniente cuando se estudian grandes masas, pues se encarece mucho su aplicación.
- No puede usarse para recoger información sobre aspectos subjetivos, hechos pasados o inquietudes futuras.

La observación puede ser aplicada de diferentes formas. En el siguiente cuadro se definen las formas más frecuentemente empleadas.

<p><b>Simple:</b> se realiza con cierta espontaneidad, de forma consciente y desprejuiciada.</p>
<p><b>Sistemática:</b> se realiza de forma reiterada y por diferentes observadores, con un control adecuado para garantizar la objetividad y uniformidad de los resultados.</p>
<p><b>Participativa:</b> el observador forma parte del grupo observado y participa con él mientras dura la observación.</p>
<p><b>No participativa:</b> el observador no forma parte del grupo investigado, realiza la observación desde fuera.</p>

**Abierta:** los sujetos a observar conocen que van a ser observados y por quien<sup>(es)</sup>, previo análisis de si este conocimiento por parte de los observados no afectará los resultados del estudio. En caso contrario no se realiza de esta forma.

**Encubierta:** las personas objeto de investigación no conocen que serán observadas. Suele ser más objetiva que la anterior.

Además de estas formas básicas, de acuerdo a los objetivos y pericia del investigador, se puede combinar algunas de ellas.

Es importante que conozcas, que al emplear este método de obtención de información, debes seguir un riguroso proceso de organización de la misma, donde tendrás en cuenta el tipo de objeto a observar, sus propiedades; las características personales del observador, los procedimientos y técnicas que se requieren, los medios con que cuentas, entre otros. Una vez realizado este proceder de análisis, se elabora un plan de observación donde se precisarán las variables a observar, el tiempo de duración de la observación y el resultado esperado, lo cual está determinado por los objetivos del estudio en particular.

Además, debes tener en cuenta que la información a recoger sobre cada uno de los conceptos o variables, deberá ser la definida en la hipótesis de trabajo o en los objetivos del estudio y no otra, lo cual garantiza la *validez* de la observación. Y, que la guía de observación deberá ser lo suficientemente precisa y clara, de forma tal que diferentes observadores al aplicarla, la entiendan y apliquen de la misma manera; esto garantiza la *confiabilidad* de la observación.

### 5.3.2 El interrogatorio

El interrogatorio, como su nombre lo indica, constituye un procedimiento de obtener información mediante la aplicación de preguntas o interrogantes a las personas objeto de estudio. Mediante su uso puede obtenerse información tanto objetiva (edad, sexo), como de la esfera subjetiva (recuerdo de hechos pasados, aspiraciones, sentimientos).

Te habrás dado cuenta que las limitaciones de la observación constituyen ventajas del interrogatorio, podemos considerarlo como el procedimiento idóneo en los casos en que se desea recoger información sobre la esfera subjetiva de las personas y sobre hechos pasados.

Entre sus desventajas generales (más adelante veremos las de cada tipo) podemos citar las siguientes:

- Sus resultados dependen de la memoria del interrogado.
- Sus resultados dependen además, de la voluntad y motivación del interrogado para responder.
- El tipo de preguntas y la manera de formularlas influyen de manera decisiva en las respuestas.

El interrogatorio puede ser de dos tipos: directo e indirecto. Veamos en el cuadro siguiente en qué consiste cada uno de ellos y sus subtipos.

### Tipos de interrogatorio.

<b>Directo:</b> es aquel que se realiza mediante la <b>entrevista</b> .	<b>Estructurada:</b> es aquella entrevista que se realiza empleando una guía previamente elaborada de preguntas (cuestionario).
	<b>No estructurada:</b> es aquella entrevista que se realiza de forma espontánea, las preguntas a formular se dejan al criterio y a la experiencia del entrevistador.
<b>Indirecto:</b> es aquel que se realiza mediante el uso de un <b>cuestionario</b> que debe llenar el interrogado por sí mismo.	<b>Postal:</b> adopta la forma de un formulario impreso, que es enviado por correo a la población o a la muestra a estudiar.
	<b>Grupal:</b> este cuestionario es aplicado a un grupo de personas, previo acuerdo del lugar y hora en presencia del investigador.

**La entrevista:** constituye una técnica de obtención de información por medio del interrogatorio, tiene la particularidad de realizarse mediante un proceso verbal, generalmente, “cara a cara” entre al menos dos individuos. Entre sus ventajas principales, además de las generales del interrogatorio, podemos citar las siguientes:

- Se complementa con la observación. Por lo que el entrevistador puede atender no sólo a lo que se dice, sino a cómo lo dice.
- Puede ser aplicada a todo tipo de personas independientemente de si presenta alguna dificultad para la escritura: analfabetos, niños pequeños, discapacitados.
- Permite que el entrevistado se exprese con libertad.
- Es flexible, es decir, se puede repetir la pregunta o aclarar el significado de una palabra no entendida.
- La información que se recoge puede ser más rica y variada que si usáramos el cuestionario.

Como desventajas podemos citar las siguientes:

- Requieren de un gran entrenamiento por parte del investigador, sobre todo las no estructuradas.
- El entrevistador con su acción puede influir en las respuestas de los sujetos.
- El entrevistado puede no sentirse seguro en cuanto al anonimato y confidencialidad de la información que ofrece.
- Los resultados dependen del nivel de comunicación entre el entrevistador y el entrevistado.

El éxito que se logre en la entrevista depende, entre otras cosas, del nivel de comunicación que alcance el investigador con el entrevistado, de la preparación que tenga el investigador en cuanto al problema de investigación, su entrenamiento y destreza para realizar las preguntas oportunas y de la manera más eficaz; las condiciones psicológicas del entrevistado; la fidelidad a la hora de transcribir las respuestas; de la confianza que tenga el entrevistado en la confidencialidad de la información que está brindando, así como la no influencia del investigador en las respuestas.

Ya vimos que de acuerdo a la forma que adopta, la entrevista puede ser estructurada o no. También puede ser una combinación de ambas, la entrevista semiestructurada: existe una guía con los tópicos principales a

tratar pero, con gran flexibilidad por parte del investigador acerca de las preguntas a realizar.

Las entrevistas no estructuradas y semiestructuradas son muy útiles en los estudios descriptivos, y en las etapas previas a la investigación (diseño); son adaptables y susceptibles de aplicarse a toda clase de sujetos y situaciones; permiten profundizar en el tema de interés, pero precisan de tiempo y personal calificado para realizarlas y procesar la información obtenida.

Por su parte, la entrevista estructurada a partir de un formulario (cuestionario) permite obtener información fácil de procesar, no se necesita de un entrevistador muy diestro para realizarla y existe uniformidad en el tipo de información obtenida, sin embargo no permiten profundizar en los aspectos que surjan durante la misma.

Al entrevistar a una persona debes seguir ciertas reglas que te ayudarán a obtener mejores resultados, a continuación te mencionamos las más importantes:

1. Presentación: debes presentarte de manera cordial, revelar el nombre de la institución que te avala así como explicar de manera sencilla los objetivos de la investigación, haciendo énfasis en la importancia de la información que te brindará así como la confidencialidad de la misma.
2. No expresar actitudes dominantes respecto al entrevistado, escucharás paciente y amistosamente.
3. No debes manifestar opiniones propias ni darás valoraciones morales de ningún tipo.
4. No debes **nunca** dar lugar a discusiones con el entrevistado.
5. Debes actuar con naturalidad y mantener los patrones de educación requeridos, tu apariencia personal lo más adecuada posible.
6. No interrumpirás ni harás preguntas extras a menos que sea necesario para ayudar a la persona a expresarse, para librar al entrevistado de tensiones, para volver a la entrevista, a elementos de información olvidados o abandonados; o para insistir o profundizar en alguna pregunta o aspecto de la entrevista.
7. Te recomendamos registrar la información lo más rápido y fiel posible, lo óptimo es usar grabadora.
8. Debes tener en cuenta las características del entrevistado: edad, sexo, escolaridad, etc.
9. Te despedirás de manera amable y cortés, agradeciéndole la información brindada y el tiempo consumido.



**El cuestionario:** es un instrumento básico de la encuesta y la entrevista estructurada, consiste en un documento concebido para ayudar en el proceso de obtener información a través de una secuencia de preguntas o aspectos a indagar.

**La encuesta:** es un método de adquisición de información, mediante el empleo de un cuestionario previamente elaborado, a través del cual se puede conocer la opinión o valoración de las personas encuestadas acerca del fenómeno de estudio.

A diferencia de la entrevista, en la encuesta las personas responden por escrito el cuestionario, sin la intervención directa de los investigadores. Una vez confeccionado el cuestionario, no se requiere de personal calificado a la hora de hacerlo llegar a los encuestados. La encuesta cuenta con una estructura lógica rígida, que permanece inalterable a lo largo de todo el proceso investigativo. Las preguntas deben ser escogidas de modo especial, del mismo modo deben ser seleccionadas las posibles variantes de respuestas, lo que facilita el procesamiento y análisis de los resultados.

La planificación de una encuesta es un proceso complejo que debe ser llevado a cabo por personal calificado. La construcción del cuestionario a aplicar es una tarea que exige del investigador un alto grado de organización y conocimiento del problema de estudio, la mayor parte de los errores que pueden producirse durante la recogida de la información se deben a la construcción inadecuada de estos.

Un elemento a tener en cuenta es el fenómeno de la **no respuesta**, no todas las personas encuestadas responden las preguntas del cuestionario, lo cual trae consigo una disminución del número real de resultados obtenidos, que si es grande (30% o más de los individuos seleccionados) afecta considerablemente la validez del estudio, esto debe tenerse en cuenta a la hora de estimar el tamaño muestral.

Otro elemento de vital importancia lo constituye la calidad del cuestionario en sí. En el cuestionario se formulan una serie de preguntas que nos permiten medir una o más variables de interés, posibilita además, observar los hechos a través de la valoración que hace de los mismos el encuestado, limitándose la investigación a las valoraciones subjetivas de éste. La estructura y el carácter del cuestionario lo definen el contenido y la forma de las preguntas que se formulan.

Las preguntas en el cuestionario, pueden clasificarse desde diferentes puntos de vista: según su contenido, según el grado de libertad de las respuestas, según su función y según su naturaleza. En el siguiente cuadro se resumen las características fundamentales de los tipos de preguntas, de acuerdo a los ejes expuestos.

### Tipos de preguntas en un cuestionario.

<b>Según el contenido</b>	<b>Directa</b>	Coincide el contenido de la pregunta con el objetivo del investigador. Por ejemplo: ¿Le gusta a usted practicar gimnasia?
	<b>Indirecta</b>	El objetivo del investigador está implícito pero no coincide el contenido de la pregunta con éste. Por ejemplo: ¿Quisiera usted que sus hijos practiquen la gimnasia?
<b>Según el grado de libertad de la respuesta</b>	<b>Abiertas</b>	Dan completa libertad al interrogado para que responda. No permite medir con exactitud la propiedad solo se alcanza a obtener una opinión. Por ejemplo: en su opinión ¿cuáles son las ventajas de la medicina general integral con respecto al modelo anterior de atención primaria?
	<b>Cerradas</b>	Las posibles respuestas son previstas por el investigador, el encuestado solo debe seleccionar la que mejor se ajuste a su situación. Por ejemplo: ¿considera usted ventajoso el modelo de atención del médico de la familia? Si <input type="checkbox"/> No <input type="checkbox"/> No tengo opinión <input type="checkbox"/>

	<b>Mixtas</b>	<p>Son una combinación de las anteriores. Por ejemplo: ¿considera usted ventajoso el modelo de atención del médico de la familia?</p> <p>Si <input type="checkbox"/> No <input type="checkbox"/> No tengo opinión <input type="checkbox"/></p> <p>¿Por qué? _____</p>
--	---------------	---

<b>Según la función</b>	<b>De contenido</b>	<p>Todas aquellas directamente relacionadas con las variables de interés. Por ejemplo: ¿cuántos hijos tiene?</p>
	<b>De filtro</b>	<p>Son aquellas que permiten saber si el encuestado conoce la materia sobre la que se quiere recoger su opinión. Por ejemplo: en lugar de preguntar ¿cuál es su opinión sobre el programa Vale la pena?, Se le puede preguntar primero ¿Ha visto usted el programa Vale la pena? Si <input type="checkbox"/> No <input type="checkbox"/></p>
	<b>De control</b>	<p>Son aquellas que se realizan para comprobar la consistencia o veracidad de las respuestas sobre una pregunta del cuestionario. Se hace una pregunta sobre un tema y más adelante se hace otra con igual tema y contenido pero redactada de diferente forma. Por ejemplo: le preguntamos sobre la edad y más adelante podemos pedirle la fecha de nacimiento. Estas preguntas deben ir separadas en el cuestionario.</p>

Según la naturaleza de lo preguntado, las preguntas pueden agruparse en: subjetivas y objetivas. Las preguntas **subjetivas** son las relacionadas con

las opiniones, motivaciones y sentimientos de las personas encuestadas. Por su parte las **objetivas**, son aquellas encaminadas a buscar datos de naturaleza objetiva, que existen independientemente de las motivaciones, opiniones, voluntad o deseos de los encuestados.

Existen otras formas de clasificar las preguntas pero creo te hemos dado bastantes ejemplos, además recuerda que no son excluyentes, sino que se pueden mezclar entre sí, por ejemplo una misma pregunta puede ser objetiva, cerrada, de contenido y directa. Por ejemplo: ¿Cuántos hijos usted tiene?

Ninguno       1       2       3       Más de 3

A continuación te brindamos algunas indicaciones generales o reglas que puedes seguir en la construcción de cuestionarios, pensamos te serán de mucha utilidad a la hora de realizar una investigación.

1. Las preguntas que se realizan deben derivarse exclusivamente de las que se formula el investigador.
2. Los cuestionarios serán tan breves como sea posible.
3. Debes establecer la necesidad de cooperación del encuestado, lo valioso de la información solicitada, la garantía de anonimato y confidencialidad de la información, lo cual deberá contribuir favorablemente en la investigación.
4. Las preguntas deben ser claras y autoexplicativas, de manera tal que no sea necesario el uso de instructivos.
5. Debe ser confeccionado de manera que se agilice y facilite el modo de contestar.
6. Las preguntas no deben construirse negativamente, pues esto puede traer confusiones a la hora de responder.
7. Las preguntas no deben exigir mucho esfuerzo de la memoria y estar redactadas teniendo en cuenta las características particulares de los encuestados: sexo, edad, escolaridad, ocupación, etc.
8. El lenguaje debe ser sencillo, evitar el uso de preguntas tendenciosas, que induzcan a responder de una manera determinada.
9. El orden de las preguntas debe disponerse de acuerdo a las características psicológicas de las mismas.
10. No se deben plantear dos preguntas en una. Se debe evaluar cuidadosamente la sintaxis de las preguntas.
11. Las preguntas deben tener sentido para todos los encuestados (interrogados).

12. Cuando halla que abordar aspectos controvertidos o embarazosos, las preguntas deben ser construidas de forma tal que no constituyan un conflicto para el yo del sujeto.
13. Salvo situaciones excepcionales debes procurar emplear preguntas cerradas.
14. Debes evitar repetir las mismas alternativas de respuestas en sucesivas y diferentes preguntas, de esta forma contrarrestas el efecto de monotonía<sup>20</sup>.
15. En primer lugar debes preguntar los datos sociodemográficos, a continuación preguntas generales simples que lo van llevando hasta preguntas más complejas.
16. Se debe incluir una pregunta final que recoja la opinión del interrogado respecto al cuestionario.
17. Los cuestionarios deben ser probados en la práctica (encuesta piloto) antes de ser aplicados definitivamente. Esta prueba piloto sirve básicamente para corregir las preguntas del cuestionario que puedan resultar conflictivas.

En muchas ocasiones la información que necesitamos se encuentra registrada, por tanto solo nos corresponde buscarla y hacer un buen uso de ella, es este registro una fuente secundaria. Ej. La historia clínica, los registros de natalidad. Si existen varias fuentes con la misma información, la escogencia de una u otra dependerá de la calidad y de la accesibilidad de la misma.

Es conveniente que conozcas, además, que según la frecuencia de la recogida de la información se definen tres métodos o procedimientos, estos son: La encuesta (vista anteriormente), el censo y los registros continuos.

---

<sup>20</sup> El interrogado tiende a responder la misma alternativa de respuesta para las diferentes preguntas.

**Encuesta**

*Método de recolección de la información estructurada según criterios previos de sistematización, que se efectúa con propósito específico, generalmente con alcance restringido a un sector de la población y de manera **ocasional**.*

**Censo**

Proceso de recolectar, completar y publicar datos demográficos, económicos y sociales pertenecientes a un tiempo específico, de todas las personas de un país o territorio determinado. Suele realizarse **periódicamente**

**Registro Continuo**

Método de recogida de la información de forma **ininterrumpida** y sistemática. El evento se registra en la medida que acontece.

El método de registro continuo es el más empleado en las estadísticas de salud. Este método nos permite registrar la información, como su nombre lo indica, continuamente, es decir, en la medida en que ésta se origina, siendo los datos suministrados por esta vía muy dinámicos. La mayoría de la información que tú notificas diariamente forma parte de estos registros.

La información así generada se refiere unas veces al total de una población o país, por ejemplo el registro de nacimientos y defunciones, y otras veces a fenómenos que suceden en determinada institución, como es el caso del registro de pacientes atendidos en determinado hospital o policlínico. Puede abarcar el universo de población o una muestra de ella.

Podrás suponer, que este método requiere de una infraestructura confiable, el uso de la tecnología de información adecuada que permita la captación, flujo, análisis y procesamiento de esos datos, lo que hace que sea un proceso costoso pero muy útil.

Gran cantidad de experiencias han demostrado claramente los adelantos que se pueden lograr, en materia de eficacia y eficiencia, con la utilización de un sistema de recolección continua de datos que se haya elaborado en forma adecuada y establecido con propiedad, con el objetivo de producir información administrativa y clínica orientada a la gestión para el apoyo operativo y la toma de decisiones.

#### **5.4 Estadísticas de población**

El término **población** es muy familiar para ti, y se refiere a un conjunto o colección de objetos, elementos o individuos que presentan al menos una característica en común.

De ahí que es un término muy relativo, pues de acuerdo a los objetivos de determinado estudio o acción a realizar, así será definida la misma. Por ejemplo, la población de adolescentes de un área de salud, aquí la característica distintiva es el hecho de ser adolescente y residir en dicha área de salud. Como este encontrarás disímiles ejemplos en la práctica diaria.

La ciencia que estudia las poblaciones humanas, su tamaño, composición, distribución geográfica, sus variaciones y causas de estas variaciones, es la **Demografía**. Cuyo nombre procede de los vocablos griegos “Demos” y “Graphein”, que significan población y descripción respectivamente.

De ahí que las Estadísticas demográficas constituyen la aplicación de la Estadística al estudio de las poblaciones humanas, con énfasis en las variables demográficas.

##### **5.4.1 Estructura poblacional.**

A la clasificación de los miembros de una población, según las categorías de una o más variables, es a lo que se denomina distribución o estructura poblacional. Por ejemplo, según la ocupación, los efectivos de una determinada población pueden ser clasificados como trabajadores, estudiantes, jubilados o desvinculados. De la misma forma se procedería con otras variables, siendo la edad y el sexo las más comúnmente empleadas, constituyéndose así la composición o estructura por edad y sexo de la población.

La estructura por edad y sexo de la población nos brinda una descripción detallada de cualquier fenómeno que acontece en el marco poblacional.

Por ejemplo, se puede a través de la observación y análisis de las proporciones de jóvenes y (o) de ancianos, inferir el grado de desarrollo socioeconómico de un país, pues se sabe que altas concentraciones de jóvenes indican elevadas tasas de natalidad, característica de países subdesarrollados, mientras que una alta proporción de ancianos nos habla de una reducción mantenida de la natalidad, acompañada de un incremento en la expectativa de vida, propia de países desarrollados.

La estructura por edad y sexo de una población es muy importante, además, porque de ella depende el comportamiento de la morbi-mortalidad. No resulta nuevo para ti, que existen marcadas diferencias entre las enfermedades que padecen las personas jóvenes y los ancianos, así mismo las causas y frecuencia de la muerte, son bien diferentes entre unos grupos de edades y otros. Toda esta información es muy necesaria, para la planificación y gestión de los servicios de salud a brindar en una población.

La representación gráfica, frecuentemente usada, de la estructura por edad y sexo de una población, es la llamada **pirámide poblacional**, la que seguramente has tenido que construir para el análisis de la situación de salud de tu comunidad. Se trata de un histograma<sup>21</sup> doble, uno para cada sexo, a ambos lados del eje de las ordenadas, donde se representan las proporciones de hombres y mujeres dentro de cada grupo de edad. Suele colocarse el sexo masculino a la izquierda. Su nombre se debe a la similitud de su forma con esta figura geométrica.

Su construcción es muy simple, debes realizar una distribución de frecuencias relativas<sup>22</sup> de la población según grupos de edades, para cada sexo por separado. El número de clases y la amplitud de cada intervalo lo decides de acuerdo a las características de tu población y a los objetivos que persigues, pero suelen usarse grupos quinquenales de edad.

Una vez construida la distribución de frecuencias relativas, confeccionas el histograma horizontal para cada sexo, representando los grupos de edades de manera ascendente desde la base hacia la cúspide de la pirámide, recuerda ubicar el sexo masculino a la izquierda. Debes colocar una leyenda lateral con la información referente a las edades de cada grupo, puedes usar las edades del intervalo o los años de nacimiento.

---

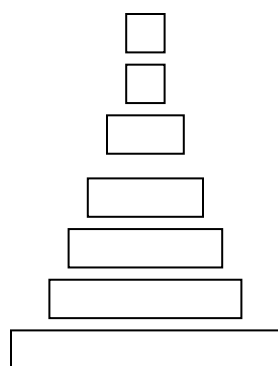
<sup>21</sup> Debes revisar el tema 2 del curso anterior para refrescar los conocimientos.

<sup>22</sup> Revisa el tema 1 del curso anterior.

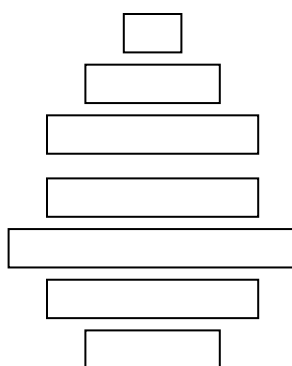


El análisis de este gráfico te permite conocer detalles de la población, tanto del presente como de eventos pasados hace varios años, podemos observar por ejemplo, cuando ocurrió un descenso en la natalidad o las huellas de una epidemia fatal.

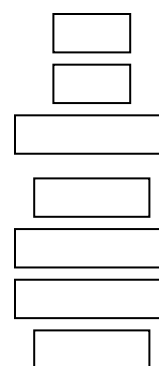
Existen tres perfiles clásicos de pirámide poblacional de acuerdo a su forma: expansiva, constrictiva y estacionaria. A continuación puedes ver sus características gráficas.



**Expansiva**



**Constrictiva**



**Estacionaria**

Entre estas formas existen matices, claro está.

La pirámide expansiva es característica de poblaciones jóvenes, con elevadas tasas de natalidad, la constrictiva, por su parte, se caracteriza por una tendencia a la disminución de la natalidad y al envejecimiento poblacional, similar a la de nuestro país; mientras que la estacionaria representa una población muy envejecida, con muy baja natalidad y crecimiento poblacional prácticamente nulo.

#### 5.4.2 Crecimiento poblacional.

El crecimiento de una población depende de tres variables fundamentales, estas son: la fecundidad, la mortalidad y las migraciones.

La **Fecundidad** se refiere a la capacidad real de reproducirse la población, es decir, la capacidad real de producir nacidos vivos, a diferencia de la **fertilidad** que es la capacidad potencial de reproducirse. Por lo tanto, la natalidad es consecuencia de la fecundidad. Más adelante, en este tema

abordaremos los indicadores de fecundidad dentro de las estadísticas vitales.

Por su parte, la **mortalidad** se refiere a las defunciones como componente del cambio poblacional, es decir, la totalidad de las muertes que se producen en una población. De manera similar se debe ver la natalidad. Se conoce que el comportamiento de estas variables es diferente de acuerdo a las características de la población, por ejemplo, se observan diferencias según el sexo, la edad, el nivel socioeconómico, el nivel educacional, etc.

Más adelante veremos específicamente como se comporta la mortalidad de acuerdo a la edad y el sexo.

En el caso de las **migraciones**, como tercera fuente de variación de la población, vemos que se refiere al movimiento de personas a través de una división política (frontera), para establecer una nueva residencia permanente. Pueden realizarse entre regiones dentro de un país o entre países, denominándosele interna e internacional, respectivamente.

A las personas que migran se les llama inmigrantes respecto al país de destino y emigrantes, respecto al país de origen, es decir, los inmigrantes llegan y los emigrantes se van.

La relación que se establece entre estas variables y el crecimiento de la población (CP) es bastante intuitiva, veamos, una población crecerá cuando los nacimientos superan en número a las defunciones y (o) las inmigraciones (entradas) superan las emigraciones (salidas). Observemos esta relación en la siguiente expresión matemática.

$$\text{CP} = \text{Nacimientos} - \text{Muertes} + \text{inmigraciones} - \text{emigraciones}$$

Si tenemos una población cerrada a las migraciones, entonces, su crecimiento dependerá solamente del número de nacimientos y defunciones que ocurran en ella.

Entre la población de dos momentos se establece una relación interesante. Si partimos del monto poblacional de un momento inicial, digamos de un censo, y le agregamos los nacidos vivos y las inmigraciones de un período de tiempo determinado, además, le restamos las defunciones y emigraciones de ese período, obtendremos así el monto poblacional a final del período. Esto resulta muy útil cuando se realizan estimaciones de

población para períodos de tiempo, sin necesidad de realizar un censo de población cada cortos períodos, lo que resultaría muy costoso.

A la ecuación matemática que expresa esta relación se le llama ecuación compensadora:

$$N^t = N^0 + B^{(0,t)} - D^{(0,t)} + I^{(0,t)} - E^{(0,t)}$$

Donde 0 y t representan los momentos inicial y final respectivamente.

N = Población

B = Nacimientos ( del inglés Birth)

D = Defunciones

I = Inmigraciones

E = Emigraciones.

Esta ecuación nos permite además, realizar el cálculo retrospectivo del monto poblacional para un momento en el pasado, conociendo la población del presente y los registros vitales y migratorios; o calcular el número de nacidos vivos o las defunciones ocurridas entre dos censos, conociendo el registro migratorio.

Por ejemplo, un médico de familia es ubicado para trabajar en un consultorio. Al llegar se encuentra con que los registros de la población están desactualizados. Es presionado por la dirección del policlínico a que entregue cuanto antes un informe sobre la población que atiende, está desesperado, no tiene tiempo de hacer un censo en ese momento por la gran carga de trabajo asistencial. ¿Qué hacer?, fácil, tomar los datos del último diagnóstico de salud, la información sobre natalidad y mortalidad del área<sup>23</sup>, información esta última fácil de obtener, y realizar una entrevista con los líderes de la comunidad, para averiguar sobre las migraciones. Con esto ya puede estimar la población actual y cumplir con su trabajo.

Veamos con cifras:

Población hace dos años atrás = 746 habitantes

Nacidos vivos del período (2 años) = 16

Defunciones del período = 5

Migraciones del período = -- Entradas = 8

-- Salidas = 14

<sup>23</sup> Me refiero al área de atención de ese consultorio.

Entonces la población actual ( $N^t$ ) será:

$$N^t = 746 + 16 - 5 + 8 - 14 = 751$$

### 5.4.3 Patrón de mortalidad según edad y sexo.

Como dijimos antes, el comportamiento de la mortalidad es diferente en dependencia de la edad de las personas. Los riesgos de morir son más elevados en las edades extremas de la vida. El gráfico que representa este comportamiento, tiene la forma de una letra J alargada, caracterizado por altos riesgos en los primeros años de la vida, riesgos mínimos entre los 5 y 15 años y a partir de aquí comienzan a incrementarse hasta el final de la vida.

Existe, también un comportamiento diferencial de la mortalidad según el sexo, caracterizado por mayores riesgos de morir para el sexo masculino, durante toda la vida. A este patrón se le conoce con el nombre de sobremortalidad masculina. Abundaremos sobre el tema en la sección de estadísticas vitales.

Esto es importante que lo tengas en cuenta al hacer cualquier análisis de la mortalidad, pues de no hacerlo cometerás graves errores al interpretar resultados de estudios o las estadísticas de una población. Más adelante verás, como puede ser controlado el efecto confusor de éstas y otras variables mediante la estandarización, al abordar la Investigación Epidemiológica.

### 5.4.4 Esperanza de vida al nacer.

Entre los indicadores demográficos destaca la esperanza de vida al nacer, siendo muy usada, junto a la tasa de mortalidad infantil, para relacionarlo con el grado de desarrollo de una región o país.

La esperanza de vida al nacer o a la edad exacta 0 año, al decir de los demógrafos, no significa la cantidad de años que vivirán las personas nacidas en ese año, se refiere a la cantidad de años que en promedio vivirán las personas nacidas en ese año, para el cual es calculada, si viven sus vidas sometidos a los mismos riesgos de morir de ese momento.

Esto es importante que lo entiendas, pues este indicador se calcula usando como referencia las tasas específicas de mortalidad de un momento y como supondrás, los riesgos de morir no se mantienen estables durante la vida de una persona. De hecho la tendencia esperada

es a la mejoría, siempre y cuando las condiciones socioeconómicas y de otra índole no empeoren, por lo tanto, el tiempo de vida promedio de esas personas debe ser mayor. Lo cual no le resta validez al mencionado indicador.

## 5.5 Estadísticas vitales

Las estadísticas vitales, como ya se ha dicho, constituyen toda la información numérica sobre los hechos vitales que se producen en una población. A continuación te ofrecemos una definición de hecho vital.



**Hechos vitales:** Aquellos eventos relacionados con el comienzo y fin de la vida del individuo y con los cambios de su estado civil que pueden ocurrir durante su existencia. Estos son el nacimiento, la muerte, matrimonios, divorcios, adopción, anulación, y otros.

Los hechos vitales de mayor interés para la salud pública son los nacimientos y las defunciones. Las Estadísticas vitales son necesarias para la planificación, ejecución y control de programas y acciones de salud, sirven para la realización de investigaciones y para la docencia. Se puede decir además que tienen una utilidad colectiva por lo antes expuesto y también utilidad individual, porque en muchas ocasiones hay que hacer uso de los registros para la solución de gestiones personales.

El Sistema de Estadísticas vitales es el encargado del registro oficial de los hechos vitales, que podrás suponer son básicos en la gestión de salud, puesto que de estos registros surgen indicadores tan importantes como la tasa de mortalidad infantil, la distribución de la mortalidad según causas, la tasa de natalidad y de fecundidad, entre otros, que además brindan información del grado de desarrollo alcanzado por un país.

No pretendo hacerte una relación detallada de dicho sistema ni de todos sus registros, solo quiero llamar tu atención hacia aquellos aspectos que puedan ser de tu interés.

Concentrémonos en las estadísticas vitales de mayor interés para la salud pública, o sea en las estadísticas de natalidad y de mortalidad. Veamos cada una de ellas por separado.

### 5.5.1 Estadísticas de natalidad

Las Estadísticas de Natalidad se refieren a la información numérica acerca de los nacimientos que ocurren en determinada colectividad, en un período de tiempo específico, y a su distribución de acuerdo a ciertas características o variables propias del nacimiento o de los padres. Por ejemplo: edad materna, peso al nacer, antecedentes obstétricos. Algunas definiciones básicas en las estadísticas de natalidad son las siguientes:

**Nacido Vivo:** Es el producto de la concepción que cualquiera que sea la duración del embarazo, sea expulsado o extraído completamente del seno materno, siempre que después de esa expulsión o extracción manifieste cualquier signo de vida.

**Parto Institucional:** Parto que ocurre en una institución del Sistema Nacional de Salud.

El modelo de inscripción de nacimiento constituye el documento básico para la recolección del dato primario, el cual es llenado por parte de las estructuras estadísticas de los hospitales donde ocurren los nacimientos<sup>24</sup> o por el personal de las Oficinas del Registro del Estado Civil para los nacimientos no institucionales. Esta información también es registrada por el personal de salud en los libros de partos.

Las estadísticas de natalidad son de gran utilidad para la administración científica, para los estudios demográficos, como vimos anteriormente; en las investigaciones en salud reproductiva, así como para otras actividades de diferentes ramas socio económicas del país.

#### Tendencia de la Natalidad

La natalidad a nivel mundial muestra una tendencia general descendente, a pesar de existir un comportamiento diferencial entre los países. Pudiera decirse que existe una relación inversa entre el nivel socio económico y el nivel de natalidad. Los niveles de natalidad suelen ser elevados en los países en desarrollo y bajos en los países desarrollados.

---

<sup>24</sup> En nuestro país el 99.9% de los nacimientos son institucionales.

Entre las causas que han determinado este descenso de la natalidad podemos citar las siguientes:

- Declinación real de la capacidad reproductiva de las personas.
- Factores culturales que rigen las costumbres matrimoniales. Incremento del divorcio y la soltería.
- Limitación voluntaria del tamaño de la familia, programas de planificación familiar y patrones culturales de fecundidad.
- Incremento de la Infertilidad dada entre otras causas, por hábitos y enfermedades del mundo moderno: alcoholismo, drogadicción, enfermedades de transmisión sexual y otros.
- Disminución de la mortalidad infantil y preescolar que otorga seguridad a la decisión de tener pocos hijos ya que con alta probabilidad sobrevivirán.

Para medir la natalidad se utilizan diferentes tipos de Indicadores: números absolutos, proporciones, tasas y razones. A continuación te muestro las fórmulas de cálculo de los indicadores<sup>25</sup> más utilizados, además de su interpretación y algunos ejemplos.

#### ◆ Tasa bruta de Natalidad (TBN)

$$TBN = \frac{\text{Número de Nacidos vivos}}{\text{Población total}} \times 1000$$

Expresa la probabilidad de que ocurran nacimientos vivos en una población. Así, la TBN de Cuba para el año 1999 fue:

$$TBN = \frac{150785}{11142691} \times 1000 = 13,5$$

Significa que en Cuba nacieron como promedio 14 niños por cada 1000 habitantes, durante el año 1999.

#### ◆ Tasa de fecundidad general (TFG)

$$TFG = \frac{\text{Número de Nacidos vivos}}{\text{Población femenina de 15 a 49 años}} \times 1000$$

Expresa la capacidad real de reproducirse de una población. Veamos la tasa para el país durante el año 1999.

$$TFG = \frac{150785}{3063266} \times 1000 = 49.2$$

<sup>25</sup> Muchos de estos indicadores los viste en el curso anterior, pero por su importancia te los mostramos nuevamente.

Se interpreta: En Cuba nacieron aproximadamente 50 niños por cada 1000 mujeres en edad fértil, durante el año 1999.

◆ Tasas específicas de fecundidad por edades (TEF)

$$TEF_{\text{Edad}} = \frac{\text{No de Nacidos vivos de madres de un grupo de edad}}{\text{Población femenina de ese grupo de edad}} \times 1000$$

Expresan la capacidad real de reproducción de grupos de edades específicos. La representación gráfica a través del polígono de frecuencia de estas tasas, dibuja la Curva de Fecundidad. Ésta puede ser Temprana cuando la fecundidad más alta está en las edades entre 20 y 24 años, Tardía cuando la fecundidad más elevada está en el grupo de edad de 25 a 29 años y Dilatada cuando la fecundidad más elevada está entre 20 y 29 años. La fecundidad temprana es propia de regiones de condiciones socioeconómicas más desfavorables mientras que la tardía es característica de regiones de más desarrollo. En nuestro país la fecundidad más alta se encuentra en el grupo de edad entre 20 a 24 años con una tendencia a trasladarse al grupo de 25 a 29 años. La TEF para el grupo de 25 a 29 años, durante el año 1999, en Cuba fue de 87.4 nacimientos por cada 1000 mujeres de 25 a 29 años de edad.

◆ Tasa Global de Fecundidad (TGF)

$$TGF = 5 \sum_{x=15}^{49} \text{Tasas específicas de fecundidad por edad}$$

Expresa el número de hijos que en promedio tuviera cada mujer de una cohorte ficticia al terminar su vida fértil, si tuvieran sus hijos de acuerdo a los niveles de fecundidad de un momento y que no están expuestas a los riesgos de morir durante ese período.

◆ Tasa bruta de reproducción (TBR)

$$TBR = 5K \sum_{x=15}^{49} \text{Tasas específicas de fecundidad por edad}$$

Donde  $K = 0.4878$ . Se refiere a la proporción de nacimientos femeninos. La constante **5** se refiere a la amplitud de los intervalos de clases para la edad; generalmente se emplean grupos quinquenales.



Esta tasa expresa la cantidad de hijas (hembras) que en promedio tuvieran las mujeres de una cohorte ficticia al terminar su edad fértil, si se cumplen los mismos supuestos que para la TGF.

◆ Tasa neta de reproducción (TNR): su fórmula de cálculo no la veremos por ser bastante compleja, pero te diré que ésta expresa la cantidad de hijas que en promedio tendrían las mujeres de una cohorte ficticia al terminar su edad fértil, si tienen sus hijos de acuerdo a los niveles de fecundidad de un momento dado. Incorpora en su cálculo el riesgo de morir de las mujeres en esas edades.

La magnitud de esta tasa mide las condiciones de reemplazo de una población, si es mayor o igual que 1, la población tiene garantizado el reemplazo poblacional porque cada mujer tiene una hija o más.

**ES MUY IMPORTANTE QUE TENGAS EN CUENTA LA CORRESPONDENCIA ENTRE EL NUMERADOR Y DENOMINADOR DE LAS TASAS, EN CUANTO A TIEMPO Y LUGAR.**

### **5.5.2 Estadísticas de mortalidad**

Las Estadísticas de Mortalidad forman parte de las Estadísticas Vitales, como ya vimos, se refieren al número de defunciones ocurridas en determinada colectividad, durante un período de tiempo definido, y su distribución de acuerdo a diferentes variables, entre las que se destaca, por su importancia en salud pública, las causas de la muerte.

Éstas, al igual que las de Natalidad, son de gran utilidad para la planificación, ejecución y control de programas y acciones de salud. Además de ser empleadas sistemáticamente en los estudios demográficos, en las investigaciones y para la docencia de pregrado y postgrado

El Sistema Nacional de Información Estadística de Defunciones y Defunciones Perinatales, es el conjunto de procedimientos encaminados a recolectar, procesar, validar y emitir los datos relacionados con la mortalidad. Su registro básico es el certificado médico de defunción, del cual te hablaremos más adelante.

### **Tendencia de la mortalidad**

De forma similar al comportamiento de la Natalidad, la Mortalidad a nivel mundial exhibe una tendencia general descendente, desde hace alrededor de 200 años; debido fundamentalmente, al desarrollo científico técnico alcanzado por la humanidad, el control y prevención de muchas

enfermedades principalmente transmisibles, así como el desarrollo socio económico. Las mejoras de las condiciones higiénicas y sanitarias unido al alto desarrollo científico técnico, traen consigo una elevación de las condiciones de vida de las personas y ello ha influido en el descenso de la mortalidad.

Desde luego que este proceso no ha ocurrido de forma balanceada en todo el mundo, no es un secreto para nadie que existen diferencias abismales en cuanto a estos aspectos entre las diferentes regiones e incluso entre regiones de un mismo país, lo que sitúa a países y regiones en condiciones de franca desventaja con relación a otros.

Muchos son los factores que intervienen en estas variaciones de la mortalidad, entre ellos podemos citar los siguientes:

◆ **Factores Biológicos.** Por ejemplo: la edad y el sexo, como vimos en estadísticas de población.

- **Sexo:** hemos dicho que a lo largo de toda la vida existe una sobremortalidad masculina, a pesar de que ocurren más nacimientos masculinos que femeninos, la razón de masculinidad es aproximadamente de 105%. Sin embargo, el efecto de la mortalidad hace que haya más mujeres que hombres en edades avanzadas de la vida, ¿No te ha llamado la atención que suelen existir más viudas que viudos?.

Entre los factores que favorecen esta sobremortalidad masculina pueden encontrarse factores biológicos, como la mayor reserva genética y la resistencia que poseen las mujeres a la enfermedad y la muerte; además de factores sociales, históricamente los patrones de conducta y realización de actividades de cada sexo fueron diferentes, generalmente los hombres desempeñan los trabajos más riesgosos físicamente; por otro lado, es conocido que las mujeres suelen poseer una percepción más desarrollada que el hombre de los riesgos y de tener una mayor responsabilidad por su salud.

- **Edad:** también planteamos, que la estructura de la mortalidad por causas es diferente en los distintos grupos de edad. Se sabe que en las edades tempranas prevalecen las causas de muerte por afecciones congénitas y neoformativas, en las edades jóvenes los accidentes, y en edades avanzadas las enfermedades degenerativas. Claro está, que en esto juega un papel preponderante las características socioeconómicas de cada región en particular, por

lo que no es posible ver este fenómeno de manera aislada, es necesario verlo desde un punto de vista multifactorial. También es diferente la frecuencia y (o) el riesgo de morir en las diferentes edades. De forma general suele ser mayor el riesgo en edades tempranas y tardías de la vida y bajo en edades jóvenes, pero, esto también se comporta diferente de acuerdo al grado de desarrollo de cada región o país.

◆ **Organización social:** la Organización Social establece el modo de producción de la sociedad y por lo tanto determina la división de la sociedad en clases y la clase dominante. Esto a su vez, establece la forma de distribución de los productos, los alimentos, las medicinas y servicios médicos en general. Todo ello es diferente para cada país o región, como todos sabemos, lo cual determina y condiciona distintos perfiles de mortalidad.

◆ **Medio ambiente:** el Medio Ambiente incluye el clima, el aire, los procesos ecológicos y otros. Cuando estos componentes son desfavorables influyen desfavorablemente en la mortalidad.

Entre los indicadores<sup>26</sup> más empleados para medir la mortalidad se encuentran las tasas, a pesar de que pueden usarse los números absolutos, razones y proporciones, a continuación veremos algunos de los más usados, con sus fórmulas de cálculo y algunos ejemplos:

◆ **Mortalidad absoluta:** se refiere al número absoluto de defunciones para un lugar y tiempo dados. Por ejemplo, en nuestro país, durante el año 1999, ocurrieron un total de 79486 defunciones.

◆ **Mortalidad proporcional:** como su nombre lo indica, se refiere a la proporción de defunciones de acuerdo a cierta característica, con relación al total de defunciones ocurridas en un lugar y tiempo dados. Suele calcularse por edad, sexo, causa de muerte y otras variables de interés y se expresa generalmente por 100.

◆

$$\text{Mortalidad Proporcional} = \frac{\text{Nb. de defunciones según la característica X}}{\text{Número total de defunciones}} \times 100$$

La Mortalidad proporcional por causas es útil cuando se requiere conocer la importancia relativa o el peso de algunas causas de muerte, sobre la

<sup>26</sup> Al igual que los de Natalidad, algunos de ellos los viste en el curso anterior pero preferimos mostrártelos nuevamente para refrescar los conocimientos.

mortalidad general. Veamos el peso que tuvieron las Enfermedades Cerebrovasculares (ECV) sobre la mortalidad general, en nuestro país durante el año 1999.

$$TMP_{ECV} = \frac{8402}{79486} \times 100 = 10.57\%$$

Se interpreta: en Cuba durante el año 1999, aproximadamente el 11% de las defunciones fueron causadas por ECV, la cual ocupa la tercera causa de muerte para todas las edades.

◆ Tasa bruta<sup>27</sup> de mortalidad (TBM)

$$TBM = \frac{\text{No. total de defunciones}}{\text{Población total}} \times 1000$$

Mide el riesgo absoluto de morir de los miembros de dicha población, para ese momento. Es ampliamente utilizada para establecer comparaciones entre regiones o entre diferentes momentos de una misma región, sin olvidar que en ocasiones es necesario estandarizarla, para evitar el efecto confusor de la estructura por edad y sexo de la población u otras variables, como verás más adelante en temas posteriores.

La tasa bruta de mortalidad para Cuba durante el año 1999<sup>28</sup> fue de 7.1 defunciones por cada 1000 habitantes.

$$TBM = \frac{79486}{11142691} \times 1000 = 7.13$$

◆ Tasa de mortalidad específica<sup>29</sup> por edad (TME<sub>edad</sub>)

$$TME_{Edad} = \frac{\text{No. de defunciones del grupo de edad X}}{\text{Población de ese grupo de edad.}} \times 1000$$

Mide el riesgo específico de morir por todas las causas en ese grupo de edad. Veamos el cálculo de la TME para el grupo de 65 años y más, en Cuba durante el año 1999.

$$TME_{65y+} = \frac{54751}{1049000} \times 1000 = 52.2$$

Se interpreta: en Cuba, durante el año 1999, fallecieron como promedio 52 personas mayores de 64 años, por cada 1000 habitantes de ese grupo de edad.

<sup>27</sup> También se le llama tasa cruda o general de mortalidad.

<sup>28</sup> Los datos sobre el año 1999 aún son provisionales.

<sup>29</sup> Del curso anterior conoces que una tasa es específica cuando se calcula en base a una parte de la población y no al total de ella.

◆ Tasa de mortalidad específica por sexo ( $TME_{\text{sexo}}$ )

$$TME_{\text{Sexo}} = \frac{\text{No. de defunciones del sexo X}}{\text{Población total sexo X}} \times 1000$$

Mide el riesgo de morir por todas las causas<sup>30</sup> para las personas de ese sexo. Por ejemplo, en Cuba durante el año 1999, fallecieron como promedio 8 personas del sexo masculino, por cada 1000 habitantes de dicho sexo.

$$TME_{\text{Masc.}} = \frac{43794}{5579257} \times 1000 = 7.8$$

◆ Tasa de mortalidad infantil (TMI) y sus componentes.

La mortalidad infantil se refiere a los fallecidos menores de un año. Su fórmula de cálculo se muestra a continuación. Recuerda que la viste en el curso anterior.

$$TMI = \frac{\text{Fallecidos menores de un año}}{\text{Total de Nacidos vivos}} \times 1000$$

Te aclaro que no mide exactamente el riesgo de morir en la población infantil menor de un año, ya que se calcula en base al total de nacidos vivos del período y no respecto a la población menor de un año en ese mismo período y lugar. No obstante, es un indicador de gran trascendencia tanto dentro del sector salud como para otros sectores. En Cuba, durante el año 1999, fallecieron como promedio 6 niños menores de un año por cada 1000 nacidos vivos.

$$TMI = \frac{966}{150785} \times 1000 = 6.4$$

La TMI puede descomponerse en tres componentes a saber: mortalidad neonatal precoz (defunciones de los recién nacidos menores de 7 días) neonatal tardía (defunciones de los recién nacidos entre 7 y 27 días) y posneonatal (defunciones ocurridas entre los 28 días y 11 meses). Esta propiedad es importante ya que el patrón causal de la mortalidad es diferente para cada componente, por lo que el predominio de uno u otro componente ofrece información adicional sobre las posibles causas de las defunciones y la factibilidad de prevenirlas.

Para cada componente se calcula una tasa con denominador común: total de Nacidos Vivos, con el numerador relativo a los fallecidos de las edades

<sup>30</sup> Esta tasa y la anterior se pueden hacer todo lo específica que queramos, por ejemplo para una causa de muerte dada.

definidas para cada componente. La suma de estas tres tasas es igual a la TMI de esa región y período.

◆ Tasa de Mortalidad Perinatal

Esta tasa mide el riesgo de morir en los momentos cercanos al parto. Puede ser de dos tipos: perinatal I y perinatal II. Se diferencian básicamente en el período que abarca cada una.

$$TMP_{\text{Perinatal I}} = \frac{\text{Defunciones fetales tardías} + \text{Defunciones neonatales precoces}}{\text{Total de Nacidos Vivos} + \text{Defunciones fetales tardías}} \times 1000$$

Defunciones fetales tardías: se considera como tal al fallecimiento del producto de la concepción, antes de su expulsión o extracción completa del cuerpo de la madre, con peso mayor o igual a 1000 gramos.

$$TMP_{\text{Perinatal II}} = \frac{\text{Def. fetales intermedias y tardías} + \text{Def. Neonatales precoces y tardías}}{\text{Nacidos Vivos} + \text{Defunciones fetales intermedias y tardías}} \times 1000$$

Defunciones fetales intermedias: fallecimiento del producto de la concepción, antes de su expulsión o extracción completa del cuerpo de la madre, con peso mayor o igual a los 500 gramos.

◆ Tasa de Mortalidad Materna Directa: esta tasa relaciona el número de defunciones maternas provocadas por causas directas del embarazo, parto y puerperio (hasta los 42 días siguientes al parto) con el total de nacidos vivos ocurridos en esa región y período dado.

$$TMMD = \frac{\text{Defunciones maternas directas}}{\text{Total de Nacidos Vivos}} \times 100000$$

En Cuba se produjeron durante el año 1999, un promedio de 29 muertes maternas, por causas directas del embarazo, parto y puerperio, por cada 100000 nacidos vivos.

$$TMMD = \frac{44}{150785} \times 100000 = 29.2$$

◆ Tasa de Mortalidad Materna Indirecta: relaciona las defunciones maternas provocadas por causas indirectas, es decir, aquellas que se desencadenan o se agudizan durante el embarazo parto y puerperio (HTA, Asma, etc.), con los nacidos vivos ocurridos en la misma región y período.

$$TMMI = \frac{\text{Defunciones maternas indirectas}}{\text{Total de Nacidos vivos}} \times 100000$$

La Tasa de mortalidad materna indirecta de nuestro país, durante el año 1999, fue aproximadamente de 14 defunciones maternas por dichas causas, por cada 100000 nacidos vivos

$$TMMI = \frac{21}{150785} \times 100000 = 13,9$$

◆ Tasa de Mortalidad Materna Total: relaciona el total de muertes maternas por causas directas, indirectas y otras causas, con el total de nacidos vivos del mismo período y región. La suma de las dos anteriores no es igual a ésta, ya que aquí se incluyen defunciones por otras causas.

◆ Tasa de Letalidad: expresa la gravedad de una determinada enfermedad o daño. Relaciona las defunciones causadas por esa enfermedad con el total de personas que la padecen. Puede ser considerada además, como una medida de morbilidad por algunos autores, pues me da una estimación acerca del pronóstico y severidad de la enfermedad. Suele expresarse en forma de porcentaje.

Existen otros indicadores de mortalidad como son los años de vida potencialmente perdidos que serán abordados en cursos posteriores.

### 5.5.3 El Certificado Médico de Defunción.

Como dijimos anteriormente, el Certificado Médico de defunción constituye el registro básico de las Estadísticas de Mortalidad. Además, el personal autorizado como registrador de la mortalidad en este modelo, es única y exclusivamente el personal médico. Por lo tanto creímos necesario dedicarle unas líneas a este aspecto.

En Cuba están vigentes tres certificados médicos de defunción, estos son:

- Certificado médico de defunción para las defunciones de 28 días y más de edad.
- Certificado médico de defunción neonatal para defunciones de nacidos vivos hasta los 27 días de vida.
- Certificado médico de defunción fetal para mortinatos o defunciones fetales de 20 semanas o más de gestación o de 500 gramos de peso o más.

En ellos, el médico debe registrar información general del fallecido, y establecer las causas de la muerte, lo cual constituye un proceder complejo que exige del médico certificante un conocimiento completo de

los estados morbosos que afectaban al fallecido, además de una sólida preparación profesional.

Las causas de la muerte son codificadas de acuerdo a la décima revisión de la Clasificación Internacional de Enfermedades (CIE), vigente en nuestro país desde el primero de Enero del año 2000. Este proceso de codificación es llevado a cabo por personal técnico y profesional entrenado que labora en las estructuras de estadísticas sanitarias, y su realización correcta depende tanto de la competencia del codificador, como de la calidad con que fueron registradas las causas de la muerte por el médico certificante.

La Causa Básica de la Defunción es la enfermedad o lesión que inició la cadena de acontecimientos patológicos que condujeron directamente a la muerte, o las circunstancias del accidente o violencia que produjo la lesión fatal. Es importante definirla con precisión y claridad, además de ser lo más fidedigna posible, puesto que de ello depende el éxito de los programas preventivos que pudieran planificarse con el análisis de la mortalidad.

Por otra parte, no es nuevo para ti que el objetivo fundamental de los programas de salud es romper la cadena causal de las enfermedades e impedir la muerte, para ello lo óptimo es prevenir la causa primaria que da origen a todos los demás trastornos o afecciones que conducen a la enfermedad, secuela o la muerte.

Ahora bien, ¿Cómo llenar correctamente un certificado médico de defunción en su variable “causas de muerte”? A continuación te brindamos algunas instrucciones que te serán muy útiles.

Llenado de causas de muerte en certificados de fallecidos de 28 días y más:

En la Parte I del certificado, se anotan las enfermedades relacionadas con la cadena de acontecimientos que condujeron directamente a la muerte y en la Parte II se indican otras entidades morbosas que hubieran contribuido, pero que no están relacionadas directamente con la causa directa de muerte.

La afección registrada en la última línea de la Parte I es la causa básica y será la utilizada para las tabulaciones. Las causas consecuentes a ésta se sitúan en las líneas superiores y en el orden que se presentaron en la



historia de enfermedad y muerte de la persona. Veamos un ejemplo tomado de Gran (2000):

Mujer de 65 años.

Parte I

- Embolia pulmonar.
- Fractura patológica.
- Carcinoma secundario del fémur.
- Carcinoma de la mama.

Parte II:

Se anotaría cualquier estado patológico importante que contribuyó a la muerte, pero que no estuvo relacionado con la enfermedad que condujo directamente a ella.

La causa básica en este ejemplo es: Carcinoma de mama. Esta será la causa a tabular (codificar) en las estadísticas rutinarias de mortalidad.

En el caso de los certificados médicos de defunción para menores de 28 días, se llena bajo el enfoque de “causa principal”. En la primera línea, de arriba hacia abajo, se coloca la causa que a juicio del médico certificador fue la principal o de mayor importancia para provocar la muerte. En la segunda línea, una o dos causas más (de existir) que consideren tuvieron también importancia. Se codifica en estos casos la causa principal.

## **5.6 Estadísticas de morbilidad.**

Recordarás del curso anterior, que la Morbilidad se refiere al patrón de enfermedades y afecciones que sufren los habitantes de una región, de ahí que las Estadísticas de Morbilidad pueden definirse como la información numérica acerca de las enfermedades, traumatismos y sus secuelas, incapacidades y otras alteraciones de la salud diagnosticadas o detectadas en la población durante un período de tiempo.

No solo la información sobre la mortalidad expone logros a la salud, si bien es una fuente de información bien conocida, esos logros de la salud pueden estar dados por la disminución de la morbilidad, por lo que estas estadísticas tienen gran importancia en la administración científica de la salud tanto en la planificación, ejecución y evaluación de programas e intervenciones de salud.

Mediante ellas se puede conocer la demanda de la población sobre servicios y programas de acuerdo a las afecciones que le aquejan, el

riesgo de enfermar, la gravedad de estas afecciones, la demanda que hacen esas enfermedades sobre los recursos médicos, todo lo que es muy útil para planificar programas y servicios, cuyo objetivo final operacional es la prevención de enfermedades.

Además, la recogida activa de esta información nos permite detectar los eventos que aparecen y tomar las medidas pertinentes para frenar la historia de las enfermedades y prevenir a los que están a su alrededor, a la vez que puedo evaluar el cumplimiento de esos programas y las ganancias en salud.

Otra utilidad, es en la investigación, sobre todo la epidemiológica, para determinar el modo de transmisión de las enfermedades, el periodo de incubación, la infecciosidad y los aspectos inmunológicos. En la docencia también son útiles las estadísticas de morbilidad.

**COMO VES, TODOS ESOS REGISTROS QUE HACES A DIARIO SOBRE LAS ALTERACIONES DEL ESTADO DE SALUD DE LA POBLACIÓN QUE ATIENDES FORMAN PARTE DE ESTAS ESTADÍSTICAS TAN ÚTILES Y NECESARIAS. ENTRE ELLAS TENEMOS LA HOJA DE CARGO, LA TARJETA DE ENFERMEDADES DE DECLARACIÓN OBLIGATORIA, LA DISPENSARIZACIÓN POR ENFERMEDADES CRÓNICAS NO TRANSMISIBLES, EL CONTROL DE LAS MORDEDURAS DE PERROS, LAS REACCIONES ADVERSAS A LAS VACUNAS, ETC.**

Debes saber, que las estadísticas de morbilidad suelen ser difíciles de registrar dada la complejidad de la morbilidad por sí misma. El proceso salud enfermedad es un fenómeno dinámico y con una carga importante de subjetividad. No existe una línea precisa que separe la salud de la enfermedad, esto se debe entre otras cosas al amplio espectro de presentación de las enfermedades; la presencia de formas asintomáticas e inaparentes de enfermedad trae consigo grandes dificultades para establecer el diagnóstico.

Por otro lado tenemos el aspecto subjetivo de la salud, pues una persona puede sentirse muy bien y ser portador de una neoplasia u otra afección asintomática, además, puede estar enferma de varias patologías a la vez o poseer un diagnóstico presuntivo que cambia o se corrobora al realizarse estudios específicos y he ahí una gran dificultad para la medición del evento.

Al igual que la Mortalidad, la Morbilidad puede medirse a través del uso de varios indicadores: números absolutos, razones, tasas y proporciones.

Entre los indicadores más usados para medir la Morbilidad, se encuentran la incidencia y la prevalencia, que ya conociste en estadística descriptiva y de las que te hablaremos con más detalle cuando abordemos la Investigación Epidemiológica en la parte II del curso. No obstante, te presentamos dos conceptos básicos para la medición de estas estadísticas, los cuales no son nuevos para ti.

- ◆ La **incidencia** es el número de casos nuevos de una enfermedad en un periodo y lugar dado.
- ◆ La **prevalencia** es el número total de casos de una enfermedad en un periodo y lugar dado. Incluye tanto los casos nuevos como los que ya existían.

Estos conceptos se expresan preferentemente a través de las tasas de incidencia y prevalencia ya vistas en el curso anterior. Te repito que las abordaremos en detalle más adelante.

Actualmente las estadísticas de morbilidad toman en consideración los factores de riesgo incorporando así los enfoques actuales de la epidemiología y la importancia del trabajo preventivo.

## 5.7 Sistemas de Información Estadística.

Ya conoces de la importancia de las estadísticas sanitarias, en qué consisten y cuales son las más usadas en la atención primaria de salud, ahora pasaremos a mostrarte como fluye esa información hacia los diferentes niveles.

Dada la gran cantidad y diversidad de información que se requiere en el sector salud, se hace necesario organizarla en diferentes sistemas de información (esto es válido para cualquier sector). Cuando la información está estructurada en sistemas bien definidos, ella se recoge, procesa, almacena, recupera y distribuye más eficientemente y las personas y las organizaciones la utilizan con mayor eficacia.

De ahí que, un Sistema de Información es aquel que se conforma con el conocimiento adquirido sobre un hecho dado, así como de elementos de análisis, evaluación, comparación y control que permite la toma de decisiones sobre un conjunto de actividades donde este opera. Comprende tanto informes cuantitativos como cualitativos. De la parte

cuantitativa se ocupan precisamente los sistemas de información estadística y se pueden definir de la siguiente forma:

**Sistema de Información Estadística (SIE):**

Es un subsistema de información, mediante el cual se realiza la recolección, flujo, procesamiento y presentación de los datos numéricos en los diferentes niveles en que opera. Pueden ser nacionales, complementarios y locales.

Los sistemas de información en salud no son novedades; durante décadas se han implantado y puesto en funcionamiento una variedad de sistemas manuales de recopilación y notificación de datos. Más recientemente, algunos de esos sistemas han pasado a convertirse en sistemas automatizados que utilizan los recursos modernos del procesamiento y análisis de datos. Los adelantos en tecnología de computación y telecomunicaciones han creado oportunidades eficaces en función del costo para el despliegue de sistemas que contribuyan a mejorar las condiciones de salud personal y comunitaria en entornos locales, nacionales e internacionales.

Estos sistemas se diseñan a través de un minucioso trabajo científico multidisciplinario, donde intervienen los funcionarios que demandan la información, los estadísticos y los especialistas en sistema.

Un requisito par la introducción adecuada de sistemas de información es el enunciado claro de los objetivos del mismo, su escala cronológica, las limitaciones y los posibles efectos secundarios de la introducción de la recopilación de datos y otros procedimientos relacionados con estos.

Antes de abordar los diferentes tipos de sistemas quisiera que conozcas someramente la estructura Estadística de Salud de Cuba.

El país cuenta con la Oficina Nacional de Estadística (ONE) que es el órgano central del estado encargado de dirigir, ejecutar y controlar la política nacional en relación con las estadísticas oficiales del país.

Los diferentes Ministerios y niveles nacionales del Estado cuentan con sus direcciones de estadísticas que satisfacen las demandas de información

numérica de acuerdo a los requerimientos de cada uno y a su vez se rigen metodológicamente por la ONE.

El Ministerio de Salud Pública de acuerdo con la división territorial del país posee cuatro niveles que desarrollan determinadas funciones de la manera siguiente:

<b>Niveles</b>	<b>Funciones</b>
Nacional	Normativa
Provincial	Dirección y coordinación
Municipal	Dirección y coordinación
Instituciones de salud	Ejecutiva

En cada uno de los cuales existen oficinas de estadística con la finalidad de producir información propia para la gestión, de esta forma se tiene: Dirección nacional de estadísticas (DNE), departamentos provinciales y municipales de estadísticas, departamentos de estadísticas de áreas de salud y departamentos de registros médicos en los hospitales. Un técnico de estadísticas de salud de un Policlínico atiende en promedio las estadísticas que generan 15 médicos y enfermeras de la familia

Por todos estos niveles fluye de manera coordinada y organizada la información relacionada con la salud, desde las áreas hasta el nivel central (DNE). En cada nivel se emite información organizada y / o analizada destinada a los directivos de salud de cada nivel para su trabajo gerencial.

Los subsistemas del Sistema de Información Estadístico Nacional (SIEN) son jerarquizados por la Oficina Nacional de Estadística y los subsistemas del Sistema de Información Estadístico Complementario (SIEC) de salud en nuestro caso, son jerarquizados por el MINSAP. Los demás SIEC pertenecen a cada órgano central del estado porque ofrecen información propia de cada uno de ellos. En Salud Pública se pueden mencionar algunos como: Sistemas de Información Estadístico de Movimiento Hospitalario, Enfermedades de Declaración Obligatoria, Consulta Externa, Registro Nacional de Cáncer, Dispensarizados. Estos son diseñados, ejecutados y evaluados por el departamento de análisis y sistema de la DNE.

Los SIEN y los SIEC tienen que aprobarse para su puesta en marcha por la Oficina Nacional de Estadística.

Por su parte, los Sistemas de Información Estadísticos Locales (SIEL) son diseñados, ejecutados y controlados por las provincias o por direcciones territoriales más pequeñas, en función de las necesidades de información del territorio y tienen que ser aprobados por las Oficinas Provinciales de Estadística.

Toda la información que registras diariamente fluye por estos sistemas hasta el nivel superior y es importante que posea determinados atributos de calidad: debe ser útil, oportuna, veraz, confiable, suficiente, flexible, factible entre otros.

Como comprenderás, garantizar la calidad de la mayor parte de las estadísticas de salud del país depende en gran medida de ti, quizás te parezca exagerado, pero como te dije, tú eres su registrador por excelencia, por lo tanto, en este terreno y en general cuando hay falta de rigor en la toma del dato primario, vale recordar el viejo refrán de los estadísticos “Garbage in, garbage out”(basura entra, basura sale), es decir, es preferible no tener información alguna que poseer datos que conduzcan a conclusiones equivocadas.

## **Resumen**

En este tema estudiaste que:

1. Las estadísticas de salud incluyen toda información numérica relacionada de modo directo con los problemas de salud, concebidos en una escala social. Constituyen un recurso clave para la gestión de salud, pues son un ingrediente esencial para los servicios de salud y para la planificación, funcionamiento, control y supervisión de programas; además constituyen una herramienta para la evaluación de las actividades de salud y los resultados de las intervenciones clínicas y epidemiológicas. Éstas pueden clasificarse en: vitales, de población, de morbilidad, de recursos, de servicios, de vivienda, de saneamiento y económicas.
2. La mayor parte de las estadísticas sanitarias del país se generan en la Atención Primaria, donde tú como médico de familia eres su registrador por excelencia. De aquí la importancia de tener una actitud seria y responsable a la hora de notificar el dato primario, pues de ello depende la construcción de indicadores que miden el estado de salud de la población

e incluso el grado de desarrollo alcanzado por la nación en materia de salud.

3. Una fuente primaria de recolección de la información es aquella en que el dato es registrado directamente del elemento que la origina, puede obtenerse mediante dos métodos básicos: La observación y el interrogatorio. Siendo el primero el método por excelencia de obtención de información en la investigación empírica, a pesar de tener algunas desventajas: requiere personal especializado, es cara y no es útil para recoger información relacionada con el aspecto subjetivo; éstas desventajas constituyen ventajas del interrogatorio. Las fuentes secundarias son aquellas en que los datos no se obtienen directamente del elemento que la origina, sino que se encuentra registrada y solo nos resta hacer un buen uso de ellas.

4. La estructura de la población por edad y sexo, no es más que la distribución de frecuencias de los elementos de dicha población de acuerdo a estas dos variables demográficas. La forma gráfica comúnmente empleada para representarla es la pirámide poblacional, la cual nos ofrece una descripción detallada de los fenómenos que suceden en la misma en relación a su crecimiento, tanto pasados como recientes.

5. Las estadísticas vitales son las más viejas y consolidadas, tiene sus orígenes en los libros de registro de las iglesias. Constituyen toda la información numérica relacionada con los hechos vitales: aquellos eventos relacionados con el comienzo y fin de la vida de un individuo y con los cambios en su estado civil, tales como nacimiento, defunción, divorcio, matrimonio, adopción, etc. El sistema de estadísticas vitales es el encargado del registro oficial de estos hechos y tiene un alcance nacional.

6. Las estadísticas de morbilidad comprenden toda la información numérica relacionada con las afecciones o desviaciones de la salud de la población. Son las más jóvenes y más difíciles de registrar debido, entre otras cosas, a las dificultades para delimitar la salud de la enfermedad, el aspecto subjetivo del proceso salud-enfermedad, la existencia de la morbilidad oculta, la existencia de formas subclínicas e inaparentes de enfermedad, etc. Sus indicadores básicos son la incidencia y la prevalencia. La primera se refiere al número de casos nuevos de la enfermedad y la segunda al total de casos existentes de la enfermedad para un período de tiempo y una población dada.

7. Cuando la información está estructurada en sistemas bien definidos, ella se recoge, procesa, almacena, recupera y distribuye más eficientemente y las personas y las organizaciones la utilizan con mayor eficacia. La creación de sistemas de información en salud, concebidos y utilizados de manera adecuada, tienen el potencial de producir valiosa información clínica y administrativa orientada a la gestión para el funcionamiento de servicios y programas y para la atención al paciente. Son básicos para la evaluación y monitoreo de la situación de salud de las poblaciones humanas, en el mejoramiento y apoyo de las actividades de promoción de la salud, en un mayor acceso al conocimiento y como herramienta en la educación de proveedores y usuarios de la información.

8. La intervención adecuada del médico en el acopio de la información referente al enfermo y a la población en general, constituye la base fundamental de los registros médicos y por tanto de las estadísticas sanitarias.

## Ejercitación

1. A continuación te ofrecemos varias definiciones de Estadísticas de salud. Escoge la que te resulte más apropiada. *Estadísticas de salud:*

a - Es la información numérica imprescindible y utilizable para planificar, ejecutar y evaluar programas de salud.

b - Son aquellas que se utilizan para conocer el estado de salud de la población y que posibilitan confeccionar y controlar los programas de salud.

c - Incluye toda la información numérica utilizable en la actividad de salud fundamentalmente para la evaluación de los programas de salud.

2. De acuerdo a la clasificación de las estadísticas sanitarias, diga, ¿a cuál de ellas corresponde cada uno de los siguientes indicadores?.

a - Tasa de mortalidad infantil.

b - Número de defunciones por enfermedad meningocócica.

c - Número de enfermeras por habitante.

d - Consultas realizadas por el médico de familia

e - Proporción de mujeres entre 20 y 35 años

f - Tasa de fecundidad general.



g - Número de consultorios de un área.

3. Clasifique las siguientes fuentes de información según su origen.

- a) Entrevista realizada a responsable de un programa de salud.
- b) Cuestionario postal sobre preferencias alimentarias.
- c) Registro nacional de cáncer.
- d) Informe final de una investigación científica.
- e) Anotaciones del médico en la historia clínica del paciente.

4. A continuación te brindamos la información referente a la población cubana, para el año 1999, según edad y sexo. Construye la pirámide poblacional correspondiente.

Tabla 1.1 Población cubana según edad y sexo. 1999.

Grupos de edad (años)	Sexo	
	Masculino	Femenino
0 - 4	385573	361824
5 - 14	882412	839334
15 - 49	3065649	3063266
50 - 59	544772	557573
60 y más	700851	741437
Total	5579257	5563434

Fuente: Anuario estadístico de Cuba. 1999.

## Autoevaluación

1. Enumera al menos cinco interrogantes, a las que puedes dar respuesta con el uso de las estadísticas de salud, en el contexto de tu área de trabajo.

2. ¿ Consideras importante las estadísticas de salud?. Fundamenta tu respuesta.

3. Se conoce por un censo, que la población de un área de salud, cerrada a las migraciones, era de 35467 habitantes el 1 de enero del año 1998. Los registros vitales refieren que el área tuvo una tasa bruta de natalidad de 10.3 y 12.5 por 1000 habitantes, para los años 1998 y 99 respectivamente. Además las tasas de mortalidad, para dicha área, fueron

de 5,6 y 4,8 por 1000 habitantes para los mencionados años. Estima la población para el 1 de Enero del año 2000.

4. Menciona tres indicadores de estadísticas vitales. Fundamenta la utilidad de uno de ellos.

5. ¿En qué consiste la utilidad de los sistemas de información estadística?

## **Bibliografía**

1. MINSAP. **Anuario estadístico de Cuba**. DNE; 1999.
2. Alvarez C, Sierra V. **La Investigación Científica en la Sociedad del Conocimiento**. MES. En preparación; 1998.
3. Silva LC. **Muestreo para la investigación en salud**. Madrid: Díaz de Santos; 1993.
4. Gran M, Castañeda I. **Temas de Estadísticas Sanitarias**. DNE, ENSAP. En proceso de revisión para publicar. La Habana; 2000.
5. Camel F. **Estadísticas médicas y de salud pública**. Edición Revolucionaria; 1987.
6. Castañeda I. ; Gran M. **Monografía: Generalidades de Estadísticas de Salud**. ENSAP. Bioestadística.
7. ISCM, FSP. **Selección de artículos de estadística**. La Habana, 1987.
8. ISCM, FSP. **Indicadores de Salud Pública**. Selección de artículos. La Habana, 1987.
9. MINSAP. **Estadísticas de salud**. Editorial Pueblo y Educación; 1985.
10. MINSAP. **Bioestadística y Computación**. Editorial Pueblo y Educación. La Habana, 1987.
11. OPS. **Sistemas de Información y Tecnología de Información en Salud. Desafíos y soluciones para América Latina y el Caribe**. Abril, 1998.
12. Herrera León, L. **Breve introducción al análisis demográfico**. ENSAP. Bioestadística. Ciudad de la Habana; 1997.
13. OMS. **Cooperación de la OMS para el fortalecimiento de los sistemas nacionales de información en salud**. Washington; 1997