

# LA APLICACION DE LOS METODOS ESTADISTICOS A LA INVESTIGACION FORESTAL

Por el Ing. H. Arthur MEYER, en colaboración  
con el Ing. Carlos TREVIÑO SALDAÑA

## I. INTRODUCCION

1. La palabra latina "status", de la cual proviene la palabra "estadística" ha sido usada para designar un estado político. Por tal razón, los primeros trabajos que se hicieron con el objeto de investigar las condiciones peculiares que pueden ser encontradas en el seno de una sociedad humana, se llamaron trabajos estadísticos. Estos trabajos principalmente se refirieron a la población y a la distribución de los individuos que la componen, según su edad, altura, etc. Otras características, tales como color, raza y enfermedades, también sirvieron para hacer esta clase de estudios.

Más tarde, los términos población y estadística, fueron aplicados no solamente a los estudios sobre poblaciones humanas, sino también a otras congregaciones de individuos similares, tales como animales, plantas, o a cualesquiera otras, pertenecientes ya sea al mundo orgánico o al inorgánico, o simplemente a un conjunto de observaciones, como por ejemplo, una serie de repetidas mediciones.

El primer paso en los estudios estadísticos consiste en la recolección de los datos, es decir, en la anotación de las diferentes observaciones hechas en cada uno de los individuos que integran la población o una muestra de la misma. Después, estos individuos deberán ordenarse de acuerdo con una o más de las características observadas. Para dar un ejemplo sencillo, podemos medir las alturas de los soldados de un ejército; la altura

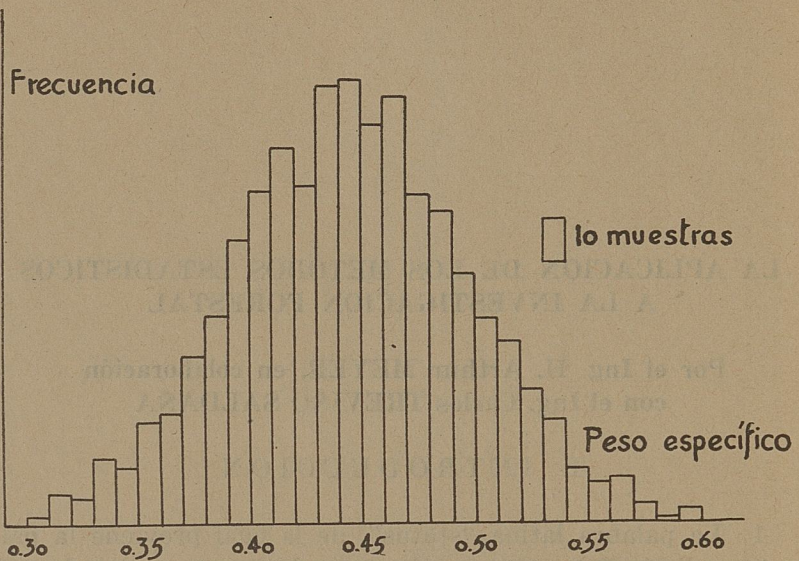


Fig. 1. Distribución del peso específico de 1,240 muestras de Douglas fir (tipo de la costa). Datos de L. J. Markwardt y T. R. C. Wilson.

es, por lo tanto, la variable de acuerdo con la cual la población será clasificada. Una vez hecho esto, y después de haber determinado el número de individuos que hay en las diferentes clases de tamaño, los problemas que se presentan son numerosos y dependen del objeto del estudio.

Es posible que únicamente estemos interesados en el valor medio de las cantidades medidas; y en este caso, para expresar este valor medio, generalmente se toma el promedio aritmético. En otros casos podemos estar interesados en la variación de las mediciones individuales alrededor del valor medio; la medida de tal variación o dispersión es llamada desviación tipo (standard deviation). El cálculo del promedio aritmético y de otras características, prácticamente representan una reducción de los datos observados.

Además de calcular ciertos parámetros que sirvan para caracterizar una población, debemos investigar de un modo más directo la naturaleza de la población. Si representamos gráficamente el número de individuos encontrados en las diferentes clases, en función de la variable independiente, de acuerdo con la cual se ha agrupado la población, podemos fácilmente com-

## Número de árboles

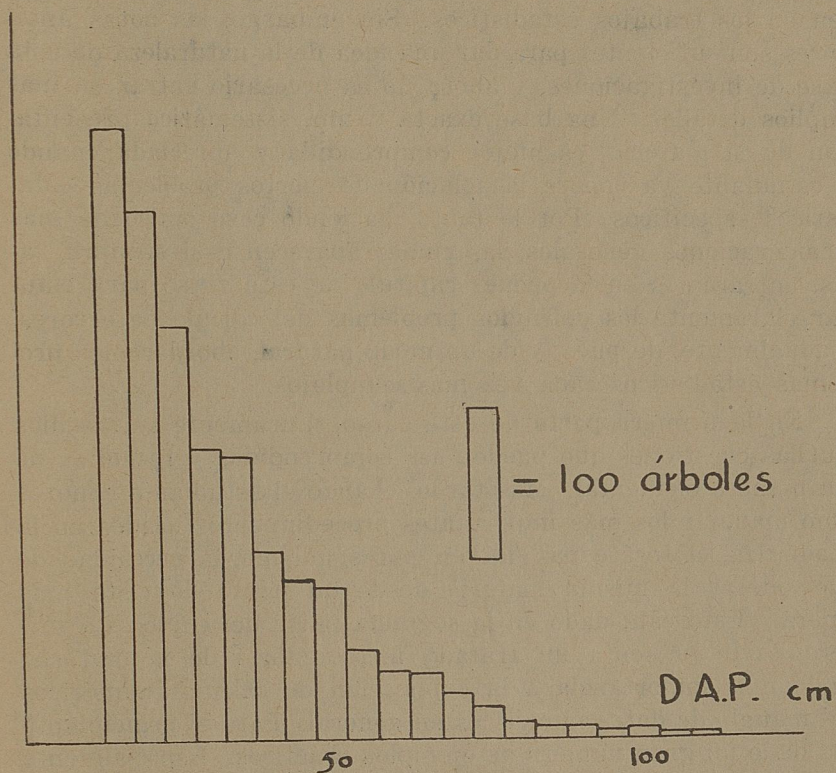


Fig. 2. Distribución del número de árboles en un bosque virgen, en función del diámetro a la altura del pecho. Superficie, 9 Ha.

prender lo que se entiende por distribución de frecuencia. La frecuencia en cada clase es convenientemente representada por la superficie de un rectángulo erigido sobre el intervalo correspondiente a la clase. Las figuras 1 y 2, representan ejemplos diferentes de distribuciones de frecuencia. Las 1,240 muestras de madera de Douglas fir, de las cuales se ha determinado el peso específico, están distribuidas simétricamente alrededor del valor medio del peso específico. En el otro caso, el número de árboles de un bosque virgen, de edades múltiples, representado en función del diámetro a la altura del pecho, muestra una distribución asimétrica. El estudio de la forma de la distribu-

ción de frecuencia juega un papel importante en los estudios estadísticos.

Sería fácil mencionar otros aspectos y problemas que surgen en los trabajos estadísticos. Sin embargo, las notas anteriores son suficientes para dar una idea de la naturaleza de esta clase de investigaciones, y ahora no es necesario entrar en más amplios detalles. Una base exacta y una sistemática presentación de la materia, es mejor comprendida y apreciada cuando el estudiante ya conoce la solución de ciertos problemas estadísticos específicos. Por lo tanto, haciendo caso omiso de más consideraciones generales, las cuales aparecen casi siempre vagas, entraremos en el primer capítulo de este curso para estudiar claramente los definidos problemas del cálculo de errores. Gradualmente, después, y de un modo natural, abordaremos problemas estadísticos cada vez más complejos.

En la primera parte de este curso, únicamente trataremos aquellas cuestiones que pueden ser comprendidas y resueltas de una manera elemental. Más tarde, cuando el estudiante conozca cómo manejar los más importantes procedimientos usados en la estadística, él será capaz de ver, por sí mismo, la necesidad de reconsiderar la misma materia desde un punto de vista más general. Esto está dado en la segunda parte del curso. La primera parte presenta un tratado introductorio de la materia, dando más importancia a la explicación de cálculos numéricos y al manejo de datos numéricos en general. Para el principiante está dado un gran número de ejemplos prácticos. Especialmente se llama la atención al hecho de que únicamente en la segunda parte hablemos en términos de probabilidad, a pesar de que la palabra es usada ocasionalmente en la primera parte.

## II. ELEMENTOS DEL CALCULO DE ERRORES

2. El cálculo de errores puede considerarse ahora como una parte de la estadística, aunque en realidad fué elaborado detalladamente y aplicado sucesivamente desde el principio del siglo pasado, es decir, aproximadamente cien años antes que los métodos modernos de la estadística matemática adquirieran la importancia que hoy tienen en cualquier rama de trabajos experimentales. Empezaremos con una exposición elemental del

cálculo de errores porque consideramos que es una apropiada introducción a los métodos estadísticos en general.

**3. Diferentes clases de errores.** En la medida de una magnitud física hay, invariablemente, un cierto límite de precisión o una cierta cantidad mínima, debajo de la cual la medición es mayor o menor y no puede obtenerse una mayor precisión. Nunca es posible encontrar el verdadero valor de una cantidad observada. Hablar del valor verdadero de una magnitud física, como, por ejemplo, de la longitud de una regla, no tiene sentido real y práctico, así como tampoco lo tendría el decir que la longitud de la regla es racional o irracional. Si no obstante, hablamos del verdadero valor de una cantidad, simplemente introducimos una noción ideal, una hipótesis de trabajo, la cual facilita el lógico tratamiento de ciertos problemas. De una manera razonable podemos interpretar esta expresión, diciendo que el valor verdadero de una cantidad observada, es aquel valor que obtendríamos haciendo un infinito número de observaciones.

Tal definición implica que ocurran errores positivos y negativos de la misma magnitud con la misma probabilidad, de manera que en una serie infinita de observaciones, se compensarían mutuamente los errores positivos y negativos cometidos. Además, es necesario decir que los errores de que hablamos están fuera del control del observador y deben considerarse como errores de observación casuales e inevitables. En la teoría de los errores trataremos únicamente de estos errores accidentales.

No se toman en consideración los errores graves de equivocación que puedan ocurrir, como, por ejemplo, cuando no contamos correctamente el número de cadenas al medir una longitud en el campo. Repitiendo la medición es fácil encontrar y rectificar tales errores.

Tenemos que mencionar otra clase de errores, los llamados errores sistemáticos. Estos errores no pueden disminuirse o eliminarse por repetidas mediciones. La mayor parte de los errores de este tipo son errores instrumentales. Por ejemplo, se comete un error sistemático si al medir la longitud de una distancia se usan cadenas demasiado cortas o demasiado largas. Corrigiendo el instrumento de medida pueden reducirse los errores sistemáticos, y aun eliminarse. Otro método para eliminar los errores sistemáticos es el arreglar la medición de la magni-

tud, de tal manera, que los errores que se cometan queden automáticamente compensados. Un ejemplo de cómo se procede con este método, es el siguiente: al usar un tránsito excéntrico, los errores que se cometen a causa de esta excentricidad, pueden eliminarse usando el aparato en posición directa y luego en posición inversa.

4. **Medición de la exactitud.** Uno de los primeros problemas que se presentan en el cálculo de errores, consiste en establecer una expresión para determinar la exactitud de una medición. Como los errores casuales difieren de una a otra medición, es necesario tomar un cierto promedio de estos simples errores. Por lo pronto, supongamos que conocemos el verdadero valor de la magnitud medida. Si designamos este verdadero valor por  $\mu$ , y los valores observados en N diferentes mediciones con  $X_1, X_2, \dots, X_N$ , los errores verdaderos serían:

$$\begin{aligned}\epsilon_1 &= X_1 - \mu \\ \epsilon_2 &= X_2 - \mu \\ &\dots\dots\dots \\ \epsilon_N &= X_N - \mu\end{aligned}$$

A fin de caracterizar por medio de un solo símbolo la exactitud de estas mediciones, la primera idea sería tomar el promedio aritmético de los valores absolutos de estos errores.

Este promedio

$$\frac{[\epsilon_1] + [\epsilon_2] + [\epsilon_3] + \dots\dots\dots + [\epsilon_N]}{N}$$

es actualmente llamado el error medio, cuando N tiende al infinito.

Como lo ha demostrado la práctica y como ha sido explicado teóricamente, este error medio no es la mejor expresión para determinar o caracterizar la exactitud de una serie de mediciones. En el cálculo de errores y en la estadística en general, es usado un promedio más eficiente, llamado el error medio cuadrático o error tipo (standard error). Tratándose de distribución de valores, esta misma expresión matemática se llama desviación tipo o desviación media cuadrática. El error tipo es igual a la raíz cuadrada del promedio de los errores simples al cuadrado. Los errores individuales, que representan valores positivos y negativos, se convierten en cifras positivas al elevarlas al cuadrado.

Expresando el error medio cuadrático o error tipo por medio de una fórmula, tenemos:

$$\sqrt{\frac{\epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2 + \dots + \epsilon_N^2}{N}} = \sqrt{\frac{1}{N} \sum \epsilon_i^2}$$

Cuando  $N \rightarrow \infty$ , obtenemos el error tipo, que se designa siempre con la letra griega  $\sigma$

En realidad, nunca podemos determinar el verdadero valor de una cantidad medida. Por esto, es también imposible calcular los verdaderos errores  $\epsilon_1, \epsilon_2, \epsilon_3, \dots, \epsilon_n$ . Por lo tanto, el promedio aritmético de las mediciones es considerado como el más probable valor de la magnitud. A primera vista, esta suposición parece razonable y ha dado buenos resultados en la práctica. Designemos el promedio aritmético de  $N$  valores observados  $X_1, X_2, X_3, \dots, X_n$ , con el símbolo  $\bar{X}$ ; entonces, las diferencias

$$\begin{aligned} v_1 &= X_1 - \bar{X} \\ v_2 &= X_2 - \bar{X} \\ &\dots \dots \dots \\ v_N &= X_N - \bar{X} \end{aligned}$$

que son llamadas residuales, representan, por lo tanto, los más probables errores o errores aparentes. Si calculamos ahora el promedio aritmético de las diferencias residuales al cuadrado y luego tomamos la raíz cuadrada de este resultado, de acuerdo con el procedimiento aplicado anteriormente al tratar con los errores verdaderos, obtendremos un valor, el cual algunas veces es demasiado pequeño, porque la suma de las diferencias residuales al cuadrado ( $\sum v_i^2$ ) es más pequeña que la suma de los errores verdaderos al cuadrado ( $\sum \epsilon_i^2$ ). Por lo tanto, se aplica a la desviación tipo observada:

$$s = \sqrt{\frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_N - \bar{X})^2}{N}} =$$

un factor de corrección a fin de obtener una mejor estimación del verdadero error tipo  $\sigma$ . Este factor de corrección será deducido en el inciso 10.

Desde luego, en el estudio de los métodos estadísticos debe saberse distinguir entre el valor verdadero y el valor estimado del error tipo; esto se hace más evidente según se avanza en el estudio de la estadística.

La letra  $S$  es usada para designar la desviación tipo,  $\sigma$  el error medio cuadrático de una muestra de  $N$  observaciones. A medida que  $N$  crece sin límite, es decir, cuando se aproxima al infinito, el valor de  $S$  difiere cada vez menos del verdadero valor de  $\sigma$ , el cual debe ser considerado como la desviación tipo de una población matriz infinitamente grande de la cual suponemos haber tomado una muestra.  $S$  representa, por lo tanto, una estimación de  $\sigma$ , aunque se obtiene una mejor estimación usando  $N - 1$  en vez de  $N$  (véase inciso 10).

5. **La ley de propagación de errores.** El primer problema que resolveremos, se refiere a encontrar el error tipo de la suma o diferencia de dos cantidades observadas  $X'$  y  $X''$ , cuyos errores tipo verdaderos son  $\sigma'$  y  $\sigma''$  respectivamente. Suponiendo que conocemos los errores verdaderos hechos en una serie de  $N$  observaciones de las dos cantidades  $X'$  y  $X''$ , designemos estos errores por.

$$\begin{array}{l} \epsilon_1', \epsilon_2', \epsilon_3', \dots, \epsilon_N' \\ \epsilon_1'', \epsilon_2'', \epsilon_3'', \dots, \epsilon_N'' \end{array}$$

Si combinamos las correspondientes mediciones hechas en las dos cantidades, en vez de combinar únicamente el promedio de las dos series de mediciones, los errores de cada uno de estos pares, serían:

$$\begin{array}{l} \epsilon_1 = \epsilon_1' + \epsilon_1'' \\ \epsilon_2 = \epsilon_2' + \epsilon_2'' \\ \dots\dots\dots \\ \epsilon_N = \epsilon_N' + \epsilon_N'' \end{array}$$

Con estos errores podemos ahora calcular, de un modo elemental, el error tipo de la suma de las cantidades  $X'$  y  $X''$ :

$$\sigma^2 = \frac{1}{N} \sum \epsilon_i^2 = \frac{1}{N} \sum (\epsilon_i' + \epsilon_i'')^2 = \frac{1}{N} \sum (\epsilon_i'^2 + \epsilon_i''^2 + 2 \epsilon_i' \epsilon_i'')$$

Como los errores  $\epsilon_i'$  y  $\epsilon_i''$  pueden ser igualmente positivos y negativos, los productos  $\epsilon_i' \epsilon_i''$  serán también unas veces positivos y otras veces negativos. Por lo tanto, la suma de estos productos, tenderá hacia cero cuando el número de observaciones tienda al infinito. De cualquier modo,  $2 \sum \epsilon_i' \epsilon_i''$  será definitivamente más pequeño que  $\sum \epsilon_i'^2$  o que  $\sum \epsilon_i''^2$ , ya que la suma de los errores individuales son siempre positivos. Por esta razón, debemos despreciar la doble suma de los productos en cruz  $\epsilon_i' \epsilon_i''$  y escribir simplemente:

$$\sigma^2 = \frac{1}{N} \sum \varepsilon_1'^2 + \frac{1}{N} \sum \varepsilon_1''^2 = \sigma'^2 + \sigma''^2$$

El error tipo  $\sigma$  de  $X = X' + X''$  es, por lo tanto, igual a la raíz cuadrada de los errores tipo al cuadrado  $\sigma'$  y  $\sigma''$ :

$$\sigma = \sqrt{\sigma'^2 + \sigma''^2}$$

De las deducciones anteriores, puede fácilmente comprenderse que se obtiene el mismo resultado cuando se calcula el error tipo de la diferencia de  $X'$  y  $X''$ . La suma de los productos en cruz es en este caso negativa, pero ya que esta suma es después omitida, no cambiará el resultado.

Por razonamientos inductivos fácilmente comprendemos que el error tipo de la suma o de la diferencia de tres o más cantidades  $X', X'', X''', \dots$  será igual a

$$\sigma = \sqrt{\sigma'^2 + \sigma''^2 + \sigma'''^2 + \dots}$$

Primero calculemos el error tipo de la suma o diferencia de  $X'$  y  $X''$ . Después tenemos  $X' \pm X''$  como una cantidad cuyo error tipo es conocido y añadamos una tercera cantidad  $X'''$  y obtendremos, por la misma fórmula, el error tipo de la nueva suma o diferencia, que actualmente es  $(X' \pm X'') \pm X'''$ . A esta nueva suma añadamos la cuarta cantidad, etc., etc. Si los errores  $\sigma', \sigma'', \sigma''', \dots$  son todos iguales, entonces el error tipo de  $X' \pm X'' \pm X''' \pm \dots$  será

$$\sigma = \sqrt{\sigma'^2 + \sigma''^2 + \sigma'''^2 + \dots} = \sigma \sqrt{N}$$

en donde  $N$  es el número de términos, sumados o substraídos.

**Ejemplo 1).**—Haciendo un inventario completo, el volumen de un bosque ha sido determinado hoy y diez años antes. En el primer cálculo se obtuvo un volumen de 4000 m<sup>3</sup> y 10 años después uno de 4800 m<sup>3</sup>. El error tipo del volumen medio se supone que sea de 1% en cada caso. ¿Cuál es el error tipo de la diferencia de volúmenes?

Error tipo del primer volumen:

$$\sigma_1 = 40 \text{ m}^3$$

Error tipo del segundo volumen:

$$\sigma_2 = 48 \text{ m}^3$$

Error tipo de la diferencia de volúmenes:

$$\sigma = \sqrt{40^2 + 48^2} = \sqrt{3904} = \pm 62 \text{ m}^3$$

Expresando este error tipo de la diferencia de volúmenes en por ciento de esta diferencia, obtenemos:

$$\sigma (\%) = \frac{62}{800} 100 = \pm 7.8 \%$$

**Ejemplo 2).**—El error tipo en volumen de 1 Ha. de un bosque más o menos uniforme se supone que sea de 1% cuando se mide una superficie de tal tamaño. ¿Cuál es el error tipo en volumen de una superficie de 25 Ha. de la misma clase, cuando se mide toda esa superficie?

Como todas las 25 Ha. están cubiertas por el mismo tipo de bosque, se supone que el error tipo de medida en volumen en cada una de las 25 Ha. será de 1%.

El error tipo de la superficie total será, por lo tanto, si  $V$  es el volumen por hectárea:

$$\frac{V}{100} \sqrt{25}$$

Si expresamos este error en por ciento del volumen total, obtenemos:

$$\left( \frac{\frac{V}{100} \sqrt{25}}{25 V} \right) 100 = \frac{1}{\sqrt{25}} = \frac{1}{5} \%$$

error que es 5 ( $=\sqrt{25}$ ) veces menor que el porcentaje de error por hectárea.

**6. La ley de propagación de errores. Continuación.** Las últimas fórmulas nos dan el error medio cuadrático de la suma o de la diferencia de varias cantidades observadas cuyos respectivos errores medios cuadráticos  $\sigma'$ ,  $\sigma''$ ,  $\sigma'''$ ,... son conocidos. Ahora debemos investigar cómo se propagan estos errores de observación cuando la cantidad final calculada no es simplemente la suma o la diferencia de dos o más observaciones directas.

El caso más simple es encontrar el error medio cuadrático de una cantidad

$$Y = k \cdot X$$

en donde  $k$  es una constante y  $X$  es una cantidad observada cuyo error tipo es  $\sigma$ . Cuando  $X$  se multiplica por  $k$ , entonces el error de  $X$  es también aumentado  $k$  veces, y es claro y no necesita explicación, que el error tipo de  $Y$  será igual a

$$\sigma_Y = k \cdot \sigma$$

De una manera elemental, también podemos deducir el error tipo de una cantidad

$$Y = X' \cdot X''$$

cuando los errores tipo de  $X'$  y  $X''$  son  $\sigma'$  y  $\sigma''$  respectivamente. Si los valores verdaderos de  $X'$ ,  $X''$  y  $Y$  son llamados respectivamente  $\xi'$ ,  $\xi''$  y  $\eta$ , tenemos.

$$\eta = \xi' \cdot \xi''$$

Además, cuando obtenemos en un par de observaciones los valores  $X' = \xi' + \epsilon'$  y  $X'' = \xi'' + \epsilon''$  podemos escribir  $Y = \eta + \epsilon$  en la forma

$$Y = X' \cdot X'' = (\xi' + \epsilon') (\xi'' + \epsilon'')$$

El error de  $Y$  para este par especial de observaciones es entonces igual a

$$\begin{aligned} \epsilon &= Y - \eta = (\xi' + \epsilon') (\xi'' + \epsilon'') - \xi' \xi'' \\ &= \xi' \xi'' + \xi' \epsilon'' + \xi'' \epsilon' + \epsilon' \epsilon'' - \xi' \xi'' \\ &= \xi' \epsilon'' + \xi'' \epsilon' + \epsilon' \epsilon'' \end{aligned}$$

o, despreciando el último término que es pequeño comparado con los dos primeros, obtenemos:

$$\epsilon = \xi' \epsilon'' + \xi'' \epsilon'$$

Finalmente, el error tipo de  $Y$  será

$$\sigma^2 = \frac{1}{N} \sum (\xi' \epsilon'' + \xi'' \epsilon')^2 = \frac{1}{N} \sum (\xi'^2 \epsilon''^2 + \xi''^2 \epsilon'^2 + 2 \xi' \xi'' \epsilon' \epsilon'')$$

Despreciando el término  $2 \xi' \xi'' \epsilon' \epsilon''$ , el cual es pequeño porque los productos  $\xi' \xi'' \epsilon' \epsilon''$  serán igualmente positivos y negativos, obtenemos:

$$\begin{aligned}\sigma^2 &= \xi'^2 \frac{1}{N} \sum \epsilon''^2 + \xi''^2 \frac{1}{N} \sum \epsilon'^2 \\ &= \xi'^2 \sigma''^2 + \xi''^2 \sigma'^2 \\ \sigma &= \sqrt{\xi'^2 \sigma''^2 + \xi''^2 \sigma'^2}\end{aligned}$$

Como no conocemos los valores verdaderos de  $X'$  y  $X''$  debemos substituirlos por el promedio aritmético de  $X'$  y  $X''$ . Haciendo tal substitución, obtenemos para el error tipo de  $Y$ , la fórmula

$$\sqrt{\bar{X}'^2 \sigma''^2 + \bar{X}''^2 \sigma'^2}$$

**Ejemplo 3).**—Se ha medido la longitud y la anchura de una superficie rectangular, y los valores obtenidos son 25 m. y 6 m. Los errores tipo de estas mediciones se suponen que son 0.1%. ¿Cuál es el error tipo del área del rectángulo?

Ya que los errores absolutos de la longitud y de la anchura son 0.025 m. y 0.006 m., el error tipo del área, será:

$$\begin{aligned}&\sqrt{(25 \cdot 0.006)^2 + (6 \cdot 0.025)^2} \\ &= \sqrt{(0.15)^2 + (0.15)^2} \\ &= \sqrt{0.045} \\ &= 0.212 \text{ m}^2\end{aligned}$$

es decir, 0.14% del área.

**7. Generalización de la ley de propagación de errores.** Habiendo estudiado de un modo elemental algunos casos concernientes a la propagación de errores, podremos ahora encontrar el error tipo de una cantidad  $Y$ , cuando

$$Y = f(x)$$

Las anteriores deducciones resuelven este problema cuando  $Y$  es una función lineal de  $X$ . A fin de determinar el error en  $Y$  para funciones de grado superior, podemos escribir, aplicando el teorema de Taylor,

$$f(X + \Delta X) = f(X) + \Delta X f'(X) + \frac{\Delta X^2}{2!} f''(X) + \frac{\Delta X^3}{3!} f'''(X) + \dots$$

Si  $X$  es el verdadero o el más probable valor de la cantidad observada, y  $\Delta X$  es el error de  $X$ , entonces  $\Delta X$  será pequeño comparado con  $X$ , y los términos siguientes a  $\Delta X f'(X)$ , pueden ser despreciados. Por lo tanto, tenemos

$$Y + \Delta_Y = f(X) + \Delta X f'(X)$$

Substrayendo de esta ecuación la siguiente

$$Y = f(X)$$

obtenemos

$$\Delta_Y = \Delta X f'(x)$$

y

$$(\Delta_Y)^2 = [\Delta X f'(x)]^2$$

Ahora, haciendo el promedio de un gran número de observaciones en cada uno de los miembros de esta ecuación, obtenemos.

$$\sigma_Y^2 = \sigma^2 [f'(x)]^2$$

o sea

$$\sigma_Y = \sigma f'(x) = \sigma \frac{df(x)}{dX}$$

Esta fórmula es una aproximación en muchos casos, ya que hemos despreciado los términos desde la segunda potencia de  $\Delta X$  en adelante; es decir, que hemos supuesto que  $\Delta X$ , el error de  $X$ , es siempre pequeño. También hemos supuesto que  $f''(X)$ , etc., no toman valores grandes.

Sin embargo, podemos convencernos de que las fórmulas dadas en los dos incisos anteriores, están de acuerdo con el presente resultado. Por ejemplo,

$$Y = k \cdot X$$

De aquí,

$$f(x) = kX, f'(x) = k, f''(x) = 0$$

Entonces

$$\sigma_Y = \sigma k$$

valor que no es aproximado, puesto que no hubo supresión de términos, ya que la segunda y siguientes derivadas dan cero.

**Ejemplo 4).**—El error tipo obtenido al medir repetidamente el diámetro  $X$  de un árbol, es  $\sigma$ , ¿cuál es el error tipo de la superficie basal correspondiente al diámetro de este árbol?

Superficie basal:

$$Y = \frac{\pi}{4} X^2$$

$$\frac{dY}{dX} = \frac{\pi}{4} 2x = \frac{\pi}{2} X$$

$$\sigma_Y = \sigma \frac{\pi}{2} X$$

Supongamos que determinamos experimentalmente el error tipo de  $X$  y que hemos obtenido un valor de 0.012 m. para  $X = 0.40$  m. El error será, por lo tanto, de 3%. De acuerdo con esto, el porcentaje de error de la superficie basal, será:

$$\sigma_Y (\%) = \frac{\sigma \frac{\pi}{2} X}{\frac{\pi}{4} X^2} \cdot 100 = \sigma \frac{200}{x}$$

$$= 0.012 \frac{200}{0.40} = 6\%$$

Es decir, que el porcentaje de error del área basimétrica, es dos veces mayor que el porcentaje de error del diámetro.

**8. Generalización de la ley de propagación de errores. Continuación.** Investigamos ahora cómo se propagan los errores cuando  $Y$  es una función de diferentes variables  $X'$ ,  $X''$ ,  $X'''$ , ...:

$$Y = f(X', X'', X''', \dots)$$

Suponiendo que de todas las variables directamente observadas, solamente  $X'$  esté afectada por un error, entonces el error tipo de  $Y$  será:

$$\sigma_{Y'} = \sigma' \frac{dY}{dX'}$$

en donde  $\frac{dY}{dX'}$  representa el coeficiente diferencial de  $Y$  con respecto a  $X'$  y  $\sigma'$  el error tipo de  $X'$ . Si únicamente es la variable  $X''$  la que está afectada por errores, entonces el error de  $Y$  será:

$$\sigma_{Y''} = \sigma'' \frac{dY}{dX''}$$

Las mismas afirmaciones pueden hacerse en relación con las demás variables  $X'''$ ,  $X''''$ ...

Podemos ahora considerar los errores  $\sigma_y, \sigma_y'', \dots$  como errores independientes cuya suma total es el error tipo de Y, causado por los errores en  $X', X'', X''', \dots$ . Así como el error tipo de la suma o diferencia de ciertas cantidades ha sido determinado, podemos calcular también el error tipo de Y:

$$\sigma_Y = \sqrt{\sigma_{Y'}^2 + \sigma_{Y''}^2 + \sigma_{Y'''}^2 + \dots}$$

$$= \sqrt{\sigma'^2 \left(\frac{dY}{dX'}\right)^2 + \sigma''^2 \left(\frac{dY}{dX''}\right)^2 + \sigma'''^2 \left(\frac{dY}{dX'''}\right)^2 + \dots}$$

Como ilustración de la fórmula obtenida, podemos calcular el error tipo de la función

$$Y = X' \cdot X''$$

en donde el error tipo de  $X'$  es igual a  $\sigma'$  el error tipo de  $X''$  es igual a  $\sigma''$ . Los coeficientes diferenciales, parciales, serán:

$$\frac{dY}{dX'} = X''$$

$$\frac{dY}{dX''} = X'$$

por lo que obtenemos para el error de Y:

$$\sigma_Y = \sqrt{\sigma'^2 X''^2 + \sigma''^2 X'^2}$$

que es la fórmula anteriormente obtenida.

**Ejemplo 5).**—Se desea calcular el volumen V de un árbol y el error tipo de este volumen, cuando el diámetro a la altura del pecho D, la altura total H y el coeficiente mórfico F, así como sus respectivos errores tipo, son conocidos.

Siendo

$$D = 0.40 \pm 0.012 \text{ m.}$$

$$H = 30.0 \pm 0.50 \text{ m.}$$

$$F = 0.463 \pm 0.01389$$

primero obtenemos

$$V = \frac{\pi}{4} D^2 H F = 0.7854 \cdot 0.16 \cdot 30 \cdot 0.463 = 1.75 \text{ m}^3$$

y entonces

$$\sigma_V^2 = \sigma_D^2 \left(\frac{dV}{dD}\right)^2 + \sigma_H^2 \left(\frac{dV}{dH}\right)^2 + \sigma_F^2 \left(\frac{dV}{dF}\right)^2$$

Los coeficientes diferenciales parciales, son:

$$\frac{dV}{dD} = \frac{2\pi D H F}{4} = \frac{\pi}{4} 11.42$$

$$\frac{dV}{dH} = \frac{\pi D^2 F}{4} = \frac{\pi}{4} 0.07408$$

$$\frac{dV}{dF} = \frac{\pi D^2 H}{4} = \frac{\pi}{4} 4.8$$

El error tipo del volumen del árbol es finalmente:

$$\begin{aligned} \sigma_v &= \frac{\pi}{4} \sqrt{(0.012)^2 (11.42)^2 + (0.50)^2 (0.07408)^2 + (0.01389)^2 (4.8)^2} \\ &= 0.7854 \sqrt{0.017781 + 0.001305 + 0.004444} = 0.120 \text{ m}^3 \end{aligned}$$

Comparando el error tipo parcial al cuadrado, causado por la inexactitud en la medición de los diámetros, con los demás contenidos bajo el radical, resulta ser el más grande. El error en volumen de  $0.12 \text{ m}^3$  es 6.9% del volumen total del árbol.

Si calculamos con el mismo grado de exactitud el volumen de 100 árboles de más o menos iguales dimensiones, el error tipo del volumen de estos árboles será de:

$$0.12 \sqrt{100} = 1.2 \text{ m}^3$$

y el porcentaje de error en volumen de estos árboles será:

$$\frac{1.2}{175} \cdot 100 = 0.7\%$$

o sea diez veces más chico que el porcentaje de error de un solo árbol.

**9. El error tipo del promedio aritmético.**—Es razonable que se tenga más confianza en el promedio aritmético de una serie de mediciones, que en una medición aislada. Por lo tanto, es lógico que el error tipo del promedio aritmético sea más pequeño que el error tipo de una observación aislada.

A fin de calcular el error tipo del promedio aritmético, tenemos que aplicar la ley de propagación de errores en su forma más sencilla. Calculando la suma de  $N$  observaciones aisladas

$$X_1, X_2, X_3, \dots, X_N$$

sin ningún conocimiento de sus respectivos errores podemos asignar a cada valor observado, el mismo error tipo  $\sigma$ , y entonces calcular el error tipo de la suma de todas las mediciones, las

cuales han sido hechas en la misma magnitud y bajo las mismas condiciones; el error tipo de esta suma tiene un valor de

$$\sigma \sqrt{N}$$

Para obtener el promedio aritmético de las  $N$ . observaciones, debemos dividir la suma de  $X_1, X_2, X_3, \dots X_N$  entre  $N$ . Haciendo esto, también reducimos  $N$  veces el error tipo de la misma suma, obteniendo

$$\frac{\sigma \sqrt{N}}{N} = \frac{\sigma}{\sqrt{N}}$$

El error tipo  $\sigma_{\bar{x}}$  del promedio aritmético de  $N$ . observaciones  $X_1, X_2, X_3, \dots X_N$ , está dado, por lo tanto, por la fórmula

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$

en donde  $\sigma$  es el error tipo de una observación aislada. Esta importante fórmula del cálculo de errores, dice: "El error tipo o error medio cuadrático del promedio aritmético de  $N$  observaciones, se obtiene dividiendo el error tipo de una observación aislada por la raíz cuadrada del número de observaciones".

**Ejemplo 6).**—Al efectuar repetidas mediciones de la altura de un árbol por medio de un hipsómetro, obtenemos algunos resultados diferentes debidos a los errores accidentales de las mediciones. Supongamos que el error tipo de una observación aislada sea 0.5 m. ¿Cuál es el error tipo del promedio aritmético de 10 observaciones? ¿Cuántas observaciones es necesario hacer para que el error tipo del promedio aritmético sea tan pequeño como 0.1 m.?

$$\sigma = 0.5 \text{ m.}$$

El error tipo del promedio aritmético de 10 observaciones, es

$$\sigma_{\bar{x}} = \frac{0.5}{\sqrt{10}} = \frac{0.5}{3.16} = 0.158 = 0.16$$

A fin de resolver la segunda cuestión tenemos que resolver la siguiente ecuación, en la cual el número  $N$  es la incógnita:

$$0.1 = \frac{0.5}{\sqrt{N}} ; N = \left( \frac{0.5}{0.1} \right)^2 = 25$$

Son necesarias 25 observaciones.

**10. Factor de corrección para calcular el error tipo de un limitado número de observaciones.**—En todas las deducciones anteriores, hemos supuesto que conocemos el verdadero valor de la cantidad observada, por lo que ha sido posible calcular el verdadero error tipo de una serie de observaciones. Al fin del párrafo 4, se indicó la necesidad de aplicar un factor de corrección cuando se calcula el error tipo tomando como base los errores aparentes, los cuales se obtienen haciendo la diferencia entre los valores observados y el promedio de las observaciones. En realidad, en la práctica nunca tenemos otra base diferente a ésta. Si fuera posible hacer un infinito número de observaciones, el promedio aritmético sería igual al verdadero valor de la cantidad observada; pero, en la práctica, cada cálculo del error tipo está basado en un limitado número de observaciones. Por lo tanto no es posible calcular el verdadero error tipo  $\sigma$ , sino un valor aproximado, una estimación de  $\sigma$ , la cual se designa con la letra  $S$ . Esta distinción es muy importante en la teoría de la estadística, como será más tarde puesta en evidencia. La cantidad  $S$ , que en general es definida como la desviación tipo de la serie de observaciones, representa, por lo tanto, una estimación de la verdadera desviación tipo. De manera análoga, podemos considerar al promedio aritmético  $\bar{X}$ , como una estimación del verdadero promedio  $\mu$ , el cual se obtiene haciendo un número infinito de observaciones.

Hemos mencionado una buena razón para suponer que  $S$  no sea la mejor estimación de  $\sigma$ ; deduciendo un factor de corrección para  $S$ , podemos obtener una mejor estimación de la verdadera desviación tipo.

Designando los errores aparentes por  $v_1, v_2, \dots, v_N$ , de tal manera que

$$\begin{aligned} v_1 &= X_1 - \bar{X} \\ v_2 &= X_2 - \bar{X} \\ &\dots\dots\dots \\ v_N &= X_N - \bar{X} \end{aligned}$$

podemos escribir

$$v_i = X_i - \bar{X} = (X_i - \mu) - (\bar{X} - \mu)$$

Elevando al cuadrado cada miembro de esta expresión, y sumando todas las observaciones hechas, obtenemos:

$$\Sigma v_i^2 = \Sigma (X_i - \mu)^2 - 2(\bar{X} - \mu) \Sigma (X_i - \mu) + N(\bar{X} - \mu)^2$$

Ya que

$$\Sigma (X_i - \mu) = N(\bar{X} - \mu),$$

tenemos:

$$\Sigma v_i^2 = \Sigma (X_i - \mu)^2 - N(\bar{X} - \mu)^2$$

Reemplacemos ahora cada término por su valor promedio, el cual se obtendría haciendo un gran número de diferentes series de  $N$  observaciones:

$$\text{Valor promedio de } \Sigma v_i^2 = N\sigma^2 - \sigma^2 = \sigma^2(N-1)$$

(El valor promedio de  $[\bar{X} - \mu]^2$  es, evidentemente, igual al cuadrado del error medio cuadrático del promedio aritmético, para el cual hemos encontrado la fórmula  $\frac{\sigma^2}{N}$ ). Despejando a  $\sigma^2$ , obtenemos:

$$\sigma^2 = \frac{\text{valor promedio de } \Sigma v_i^2}{N-1}$$

A pesar de que  $\frac{\Sigma v_i^2}{N-1}$  varía de una a otra muestra de  $N$  observaciones, se desprende de lo anterior que  $\frac{\Sigma v_i^2}{N-1}$  será una estimación imparcial de  $\sigma^2$ , ya que este valor promedio es teóricamente igual a  $\sigma^2$ . Por lo tanto, podemos escribir

$$S'^2 = \frac{\Sigma v_i^2}{N-1}$$

para una estimación imparcial de  $\sigma^2$ .

Si determinamos el error tipo de una serie de observaciones, tomando como base los errores aparentes  $v_1, v_2, \dots, v_N$ , obtenemos la mejor estimación del error tipo, dividiendo la suma de los errores aparentes al cuadrado, o cantidades residuales, entre  $N - 1$  en vez de dividirlos entre  $N$  y sacando raíz cuadrada al cociente obtenido.

Por lo tanto, el factor de corrección, es:

$$\sqrt{\frac{N}{N-1}}$$

una cantidad que es más grande que uno, lo que ya se había previsto. La importancia de este factor de corrección viene a ser cada vez más pequeña a medida que aumenta el número de observaciones. Para un corto número de observaciones,  $S'$  es una mejor estimación de  $\sigma$  que la desviación tipo  $S$ .

**Ejemplo 7).**—La altura de un árbol es medida por medio de dos hipsómetros: el de Christen y el de Büchi, 20 veces con cada uno. Se desea comparar la exactitud de los instrumentos. Las dos series de mediciones, así como el cálculo de los promedios aritméticos y el del error tipo, se encuentran en la tabla 1. El error medio cuadrático de las mediciones hechas con el hipsómetro de Christen, es igual a 0.71 m., mientras que el error medio cuadrático de las mediciones hechas con el de Büchi, es sólo de 0.44 m. Dividiendo estos valores por la raíz cuadrada del número de observaciones hechas en cada serie, obtenemos:

	Promedio aritmético y su error tipo	
Christen. . . . .	26.4	$\pm 0.16$ m.
Büchi. . . . .	26.6	$\pm 0.10$ m.
	0.2	m.

Según la fórmula que nos da el error medio cuadrático de la suma o diferencia de dos cantidades, obtenemos para el error medio cuadrático de la diferencia:

$$1/\sqrt{(0.16)^2 + (0.10)^2} = 1/\sqrt{0.0356} = \pm 0.19$$

valor que aproximadamente es igual que la diferencia misma. Por lo tanto, esta diferencia puede ser causada por errores accidentales de observación y no hay razón para creer que se deba a errores sistemáticos de los instrumentos.

Se puede ahora calcular cuántas mediciones hay que hacer con el hipsómetro de Christen para obtener la misma precisión que se obtiene con el de Büchi. Aplicando la fórmula del error medio cuadrático del promedio aritmético, obtenemos:

$$0.44 = \frac{0.71}{\sqrt{N}}$$
$$0.1936 = \frac{0.5041}{N}$$
$$N = 2.6$$

Por lo tanto, se necesita hacer tres veces más observaciones (arredondando 2.6) con el hipsómetro de Christen para obtener la misma exactitud que con el otro instrumento.

TABLA 1.—Mediciones de la altura de un árbol

Hipsómetro de Christen			Hipsómetro de Büchi				
Medición			Medición				
No.	$x_i$	$v_i$	No.	$x_i$	$v_i$	$v_i^2$	
	m	m		m	m		
1	27.0	0.6	0.36	1	26.2	-0.4	0.16
2	26.5	0.1	0.01	2	26.1	-0.5	0.25
3	25.5	-0.9	0.81	3	25.9	-0.7	0.49
4	28.0	1.6	2.56	4	25.9	-0.7	0.49
5	27.0	0.6	0.36	5	25.9	-0.7	0.49
6	26.5	0.1	0.01	6	26.9	0.3	0.09
7	26.0	-0.4	0.16	7	26.4	-0.2	0.04
8	26.4	0.0	0.00	8	27.0	0.4	0.16
9	28.0	1.6	2.56	9	26.9	0.3	0.09
10	25.5	-0.9	0.81	10	26.6	0.0	0.00
11	26.3	-0.1	0.01	11	26.8	0.2	0.04
12	26.0	-0.4	0.16	12	26.9	0.3	0.09
13	26.6	0.2	0.04	13	27.4	0.8	0.64
14	25.8	-0.6	0.36	14	26.9	0.3	0.09
15	26.3	-0.1	0.01	15	27.1	0.5	0.25
16	26.5	0.1	0.01	16	26.9	0.3	0.09
17	25.8	-0.6	0.36	17	26.9	0.3	0.09
18	25.5	-0.9	0.81	18	26.4	-0.2	0.04
19	26.0	-0.4	0.16	19	26.9	0.3	0.09
20	26.5	0.1	0.01	20	26.4	-0.2	0.04
527.7			9.57	532.4			3.72
26.4				26.6			

$$S^2 = \frac{9.57}{19} = 0.540$$

$$S^2 = \frac{3.72}{19} = 0.196$$

$$S' = \pm 0.71$$

$$S' = \pm 0.44$$

$$S'_x = \frac{0.71}{\sqrt{20}} = \pm 0.16$$

$$S'_x = \frac{0.44}{\sqrt{20}} = \pm 0.10$$

11. **El promedio aritmético ponderado.**—No siempre podemos considerar que en una serie de mediciones cada una de ellas esté hecha con la misma exactitud. Por lo tanto, el simple promedio aritmético de estas observaciones no sería el valor más probable de la cantidad observada. La mejor manera de proceder, en tal caso, puede explicarse con un ejemplo sencillo: supongamos que las mediciones  $X_1, X_2, X_3, X_4, X_5$ , fueron hechas en las mismas condiciones y con el mismo cuidado. El valor más probable de la cantidad observada sería el promedio aritmético sencillo:

$$\bar{X} = \frac{1}{5} (X_1 + X_2 + X_3 + X_4 + X_5)$$

Consideremos ahora los promedios parciales:

$$\bar{X}_1 = \frac{1}{2} (X_1 + X_2) \quad \text{y} \quad \bar{X}_2 = \frac{1}{3} (X_1 + X_2 + X_3)$$

como dos diferentes observaciones de diferente exactitud, o como se dice, de diferente peso. Sabiendo, en nuestro caso, cómo fueron obtenidas estas dos observaciones, también podemos conocer el mejor promedio de las dos:

$$\bar{X} = \frac{2\bar{X}_1 + 3\bar{X}_2}{2 + 3}$$

promedio que evidentemente es igual a  $\bar{X}$  calculado anteriormente. Los coeficientes 2 y 3 de las dos observaciones  $X_1$ , y  $X_2$ , son llamados los pesos; ellos indican la relativa importancia atribuida a las dos observaciones. Si desde un principio se asignan pesos definidos a las diferentes observaciones, entonces, de acuerdo con las anteriores explicaciones, una observación de peso  $N$  tiene que considerarse como equivalente al promedio aritmético de  $N$  observaciones de peso 1.

Por consiguiente, el mejor promedio de  $N$  observaciones  $X_1, X_2, X_3, \dots, X_N$ , cuyos pesos son  $w_1, w_2, w_3, \dots, w_N$ , está dado por la fórmula:

$$\bar{X} = \frac{w_1 X_1 + w_2 X_2 + \dots + w_N X_N}{w_1 + w_2 + \dots + w_N}$$

o escribiéndola más simplificada:

$$\bar{X} = \frac{\sum w_i X_i}{\sum w_i}$$

**12. El error tipo del promedio aritmético ponderado.**—Es fácil de demostrar, de acuerdo con la definición de  $\bar{X}$ , que

$$\sum w_i v_i = 0$$

escribiendo

$$\begin{aligned} \frac{1}{\sum w_i} \sum w_i v_i &= \frac{1}{\sum w_i} (\sum w_i (X_i - \bar{X})) \\ &= \frac{1}{\sum w_i} (\sum w_i X_i - X \sum w_i) \end{aligned}$$

obtenemos,

$$\frac{1}{\sum w_i} \sum w_i v_i = \bar{X} - \bar{X} = 0$$

Consideremos ahora el error tipo  $\sigma$ , correspondiente a una observación de peso uno; no importa que tal observación se haya hecho o no. Los errores tipo de las observaciones cuyos pesos son  $w_1, w_2, \dots, w_N$ , serán:

$$\frac{\sigma}{\sqrt{w_1}}, \frac{\sigma}{\sqrt{w_2}}, \dots, \frac{\sigma}{\sqrt{w_N}}$$

ya que una observación de peso  $w_i$  (de acuerdo con la definición de peso) debe considerarse como de la misma precisión que el promedio aritmético de  $w_i$  observaciones de peso uno.

El promedio aritmético ponderado de las anteriores observaciones,

$$\bar{X} = \frac{\sum w_i X_i}{\sum w_i}$$

puede escribirse en la forma

$$\bar{X} = \frac{w_1 X_1}{\sum w_i} + \frac{w_2 X_2}{\sum w_i} + \dots + \frac{w_N X_N}{\sum w_i}$$

Esta expresión es una función, en la cual el valor de  $X$  depende de las variables independientes  $X_1, X_2, \dots, X_N$ , cuyos errores tipo son  $\frac{\sigma}{\sqrt{w_i}}$

De acuerdo con la fórmula de la propagación de errores

$$\sigma_{\bar{X}}^2 = \sigma^2 \left( \frac{dY}{dX^1} \right)^2 + \sigma'^2 \left( \frac{dY}{dX^2} \right)^2 + \dots$$

tenemos:

$$Y = \bar{X} = \frac{w_1}{\sum w_i} X_1 + \frac{w_2}{\sum w_i} X_2 + \dots ;$$

$$X' = X_1, \quad X'' = X_2, \text{ etc.}$$

por lo tanto,

$$\frac{d Y}{d X'} = \frac{d \bar{X}}{d X_1} = \frac{w_1}{\sum w_i}$$

$$\frac{d Y}{d X''} = \frac{d \bar{X}}{d X_2} = \frac{w_2}{\sum w_i}, \text{ etc.}$$

y, finalmente,

$$\sigma_{\bar{x}}^2 = \left( \frac{\sigma}{\sqrt{w_1}} \right)^2 \left( \frac{w_1}{\sum w_i} \right)^2 + \left( \frac{\sigma}{\sqrt{w_2}} \right)^2 \left( \frac{w_2}{\sum w_i} \right)^2 + \dots$$

$$= \frac{\sigma^2}{(\sum w_i)^2} (w_1 + w_2 + \dots + w_n)$$

$$= \frac{\sigma^2}{(\sum w_i)^2} \sum w_i = \frac{\sigma^2}{\sum w}$$

De aquí, que el error tipo del promedio aritmético ponderado es

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{\sum w_i}}$$

es decir, que el error tipo del promedio aritmético ponderado es igual al error tipo de una observación de peso uno, dividida entre la raíz cuadrada de la suma de los pesos de todas las diferentes observaciones, o, en otras palabras, el peso del promedio aritmético es igual a la suma de los pesos de una observación aislada.

Para determinar el error tipo  $\sigma$  de una observación de peso uno, elevemos al cuadrado la expresión

$$e_i = v_i + (\bar{X} - \mu)$$

la cual proviene de:

$$e_i = X_i - \mu = (X_i - \bar{X}) + (\bar{X} - \mu)$$

y multiplicando la expresión resultante por el peso  $w_i$  de la observación  $X_i$ , y después sumando las expresiones para todos los valores de  $i$ , obtenemos:

$$\sum w_i e_i^2 = \sum w_i v_i^2 + (\bar{X} - \mu)^2 \sum w_i + 2 (\bar{X} - \mu) \sum w_i v_i$$

Pero como

$$\sum w_i v_i = 0$$

el último término de la ecuación anterior se suprime, quedando solamente:

$$\sum w_i E_i^2 = \sum w_i v_i^2 + (\bar{X} - u)^2 \sum w_i$$

El error tipo  $\sigma$ , de una observación de peso  $w_i$ , es, como ya se dijo:

$$\sigma_i = \frac{\sigma}{\sqrt{w_i}}$$

por lo que tenemos:

$$w \sigma_i^2 = \sigma^2$$

Si ahora reemplazamos cada término de esta ecuación por su valor promedio (comparar con el inciso 10), resulta:

$$N \sigma^2 = \text{valor promedio de } \sum w_i v_i^2 + \sigma^2$$

El último término resulta de que el error tipo del promedio aritmético  $\bar{X}$  es  $\frac{\sigma}{\sqrt{\sum w_i}}$ . Por lo tanto, tenemos:

$$\sigma^2 (N - 1) = \text{valor promedio de } \sum w_i v_i^2$$

y

$$S^2 = \frac{\sum w_i v_i^2}{N - 1}$$

representa una estimación imparcial de  $\sigma^2$ , ya que el valor promedio de  $S^2$  es teóricamente igual a  $\sigma^2$ . Naturalmente, que  $S^2$  fluctúa de una serie a otra de observaciones.

Finalmente, podemos escribir:

$$S = \sqrt{\frac{1}{N - 1} \sum w_i v_i^2}$$

**Ejemplo 8).** Se han hecho, con el hipsómetro de Büchi y con el de Christen, 2 y 5 observaciones respectivamente. Asignando el peso de 1 a las observaciones hechas con el hipsómetro de Christen y el peso de 2.6 a las observaciones hechas con el de Büchi (véase ejemplo 7, del inciso 10), calculemos el promedio aritmético ponderado, el error tipo de una observación de peso uno y el error tipo del promedio aritmético ponderado. Todos los cálculos se encuentran en la Tabla 2.

TABLA 2.—Cálculo del promedio aritmético ponderado  
y de su error tipo

Observación número	Instru- mento	$X_i$	$w_i$	$w_i X_i$	$v_i$	$w_i v_i$	$w_i v_i^2$
1	Büch	26.4	2.6	68.6	-0.3	-0.8	0.23
2		27.0	2.6	70.2	0.3	0.8	0.23
3		27.0	1.0	27.0	0.3	0.3	0.09
4	Christen	26.5	1.0	26.5	-0.2	-0.2	0.04
5		25.5	1.0	25.5	-1.2	-1.2	1.44
6		28.0	1.0	28.0	1.3	1.3	1.69
7		27.0	1.0	27.0	0.3	0.3	0.09
$\Sigma$			10.2	272.8			3.81
$\bar{X}$	$\left( = \frac{\Sigma w_i X_i}{\Sigma w_i} \right)$			26.7			

error tipo de una observación de peso uno:  $S'^2 = \frac{3.81}{6} = 0.635$

$$S' = 0.79 \text{ m}$$

error tipo del promedio aritmético ponderado:  $S'_x = \frac{0.79}{\sqrt{10.2}} = 0.25 \text{ m}$

$$\bar{X} = 26.7 \pm 0.25 \text{ m.}$$

13. **Sobre la distribución de errores.**—Aun no hemos dicho de un modo definitivo cómo aprovechar o interpretar prácticamente el error tipo calculado; pero esto solamente es posible si conocemos cómo se distribuyen los errores individuales de una serie de observaciones. En el inciso 3, se ha dicho que ocurren con la misma frecuencia los errores positivos y los errores negativos. La experiencia nos muestra también, que los errores pequeños cometidos en una gran serie de observaciones, ocurren con más frecuencia que los errores grandes. Cuando agrupamos los errores en diferentes clases, de acuerdo con su magnitud, y luego determinamos la frecuencia con que aparecen en cada clase, es decir, la distribución de la frecuencia de aparición en cada grupo o clase de tamaño, encontramos una distribución simétrica, de forma acampanada, semejante a la figura 1, de la introducción de este trabajo. A reserva de verificar una minuciosa investigación sobre esta distribución de los errores, podemos

decir desde ahora que, en una serie de observaciones, no ocurren errores más grandes que tres veces la magnitud del error tipo.

En la figura 1, el número de observaciones en una clase está representado por un rectángulo erigido sobre el intervalo ocupado por dicha clase. El mismo tipo de diagrama es usado para representar gráficamente la frecuencia de los errores de diferente magnitud. Si tenemos a nuestra disposición una serie muy grande de observaciones, podríamos hacer muy pequeños los intervalos de las clases de tamaño y obtendríamos aún, en cada clase, un gran número de observaciones. En esta forma, los rectángulos que representan la frecuencia en las clases de tamaño, se harían cada vez más angostos, y la línea quebrada que forma el contorno de la figura 1, se tornaría en una curva suave. Esta curva representa la distribución de los errores de acuerdo con su tamaño, o, en otras palabras, indica el número de errores que se espera aparezcan en cada clase de tamaño, lo que es de suma importancia en el cálculo de errores. Esta curva recibe el nombre de curva de error, o también el de curva normal. (Figura 3.) Ya hemos dicho que la máxima frecuencia de errores se encuentra alrededor del valor cero, haciéndose cada vez más pequeña a medida que aumenta el valor absoluto de los errores. Por lo tanto, la acampanada curva de error decrece gradualmente en ambas direcciones.

Las anteriores discusiones muestran que todas estas informaciones pueden lograrse a partir de una base empírica solamente; sin embargo, la curva de error fué realmente encontrada y descrita exactamente por los matemáticos, valiéndose de consideraciones puramente matemáticas. Nosotros no entraremos ahora en estas consideraciones que permiten establecer una función de la curva de error; este aspecto del problema será discutido en la segunda parte de este trabajo.

Esto no supone que todos los errores sigan la ley normal, pero aun cuando la curva normal sea obedecida aproximadamente, la ecuación que derivaremos para ésta, dará excelentes resultados en la práctica. Aun cuando en la distribución de los errores haya desviaciones considerables de la normalidad, esto no afectaría seriamente el resultado que se obtiene considerando a la curva de error como perfectamente normal.

**14. Algunas propiedades de la curva de error.**—Un estudio preliminar de las propiedades de la curva de error será muy pro-

vechoso para cuando se llegue a la última parte de este trabajo. Estas propiedades pueden ser desarrolladas a partir de la fórmula de la curva normal del error, la cual puede escribirse en la forma siguiente:

$$Y^* dX = \frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} dX$$

En esta expresión,  $\mu$  es el verdadero valor de la cantidad observada,  $N$  es el número total de observaciones,  $\sigma$  el error medio cuadrático,  $\pi$  es igual a 3.14159... y  $e$ , es la base de los logaritmos neperianos, y cuyo valor es 2.7182...;  $X$  es el valor de una observación, y la diferencia  $X - \mu$  es, por lo tanto, el error verdadero,  $\epsilon$ , de esta observación. En la notación diferencial dada arriba,  $Y^*dX$  representa el número de observaciones en la clase cuyo intervalo es  $X \pm \frac{1}{2} dX$ , representado en la figura 3 por el área de la faja cuya anchura es  $dX$ , y la cual está erigida precisamente sobre la clase cuyo intervalo es  $X \pm \frac{1}{2} dX$ . Ya que el contorno superior de la distribución de frecuencia afecta la forma de una curva suave, la superficie de esta faja no es exactamente un rectángulo y, cuando la anchura del intervalo de la clase  $dX$ , tiende a cero, la diferencia entre el área del rectángulo y el área de la faja que están bajo la curva, se hace cada vez más pequeña. La notación diferencial está, por lo tanto, de acuerdo con lo que se ha dicho anteriormente acerca de las representaciones gráficas de las frecuencias.

Ahora, si deseamos calcular el número de observaciones que, de acuerdo con la curva de error, deberá haber dentro de la clase cuyos límites son  $a$  y  $b$  (figura 3), tenemos que resolver el integral definido

$$\int_a^b Y^* dX = \int_a^b \frac{N}{\sigma \sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} dX$$

Para la evaluación de este integral entre cualesquier límites de clases, se han elaborado ya tablas extensas que nos dan las frecuencias relativas; estas frecuencias relativas se obtienen dividiendo las frecuencias absolutas entre el número total de observaciones. De aquí que la suma de todas las frecuencias relativas calculadas para una serie de observaciones, sea siempre igual a

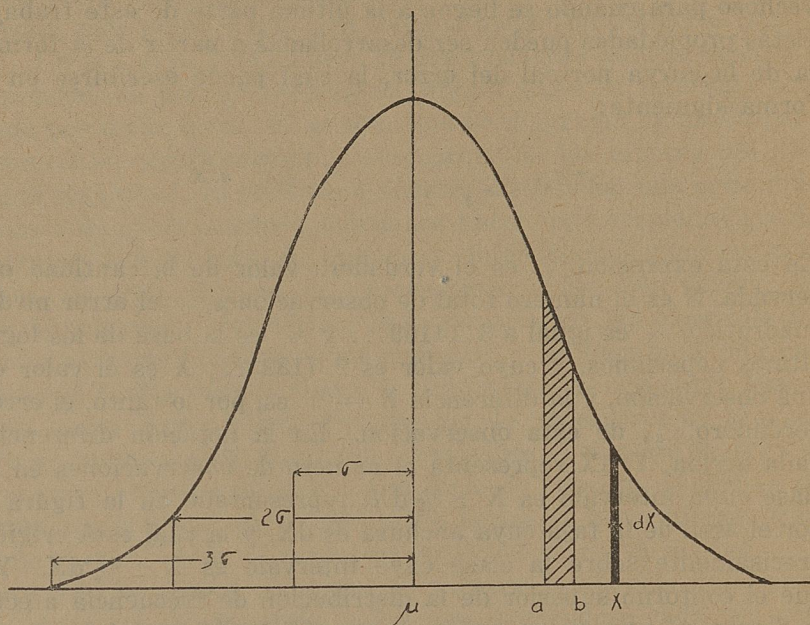


Figura 3. La curva de error o curva normal.

uno, y, del mismo modo, el área total que se encuentra debajo de la curva normal, la cual representa estas relativas frecuencias, deba también ser igual a la unidad. En este caso, la fórmula de la curva de error debe leerse ( $N=1$ ):

$$Y dX = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} dX$$

Ya que  $X - \mu$  es igual al error verdadero,  $\epsilon$ , de una observación, podemos substituir  $\epsilon$  por  $X - \mu$ ; esto es lo mismo que considerar al origen de las coordenadas donde se representa la curva de error, como el verdadero valor de la cantidad observada:

$$Y d\epsilon = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{\epsilon^2}{2\sigma^2}}$$

Esta fórmula puede también ser simplificada, expresando los errores verdaderos  $\epsilon$  en unidades de la desviación tipo. Haciendo tal substitución

$$X = \frac{\varepsilon}{\sigma}$$

y designando ahora la variable dependiente con  $y$ , obtenemos la fórmula de la curva de error más comúnmente usada:

$$y \, dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

En esta forma diferencial  $ydx$  representa la frecuencia relativa de los errores comprendidos en la clase cuyo intervalo es  $X \pm \frac{1}{2} dX$ . La frecuencia relativa en cualquier clase cuyo intervalo esté comprendido entre  $a$  y  $b$ , está dada por el integral definido:

$$\int_a^b y \, dx = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

La evaluación de esta integral, sobre la cual se basa la formación de la tabla de las áreas comprendidas bajo la curva de error (véase inciso 15), es un problema puramente matemático. Fácilmente podría demostrarse que el integral de la función desde menos infinito hasta más infinito, es igual a uno; esto será demostrado en la segunda parte de este trabajo. Ahora sólo discutiremos brevemente la forma de la curva de error, definida por la función

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Diferenciando con respecto a  $X$ , obtenemos:

$$\frac{d y}{d x} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} (-X)$$

Si hacemos este resultado igual a cero, encontramos las siguientes soluciones para el máximo y el mínimo de la curva de error:

$$(i) \quad X = 0$$

$$(ii) \quad e^{-\frac{X^2}{2}} = \frac{1}{e^{\frac{X^2}{2}}} = 0$$

$$\therefore X \pm \infty$$

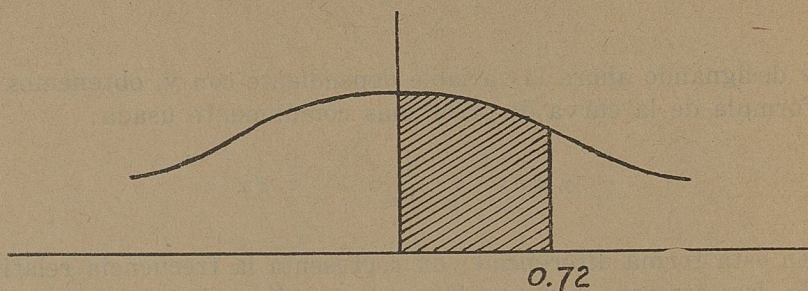


Fig. 4.

La posición de los dos puntos de inflexión de la curva, se encuentran haciendo la segunda derivada de  $y$  con respecto a  $X$  igual a cero; el resultado es:

$$\frac{d}{dx} \left[ e^{-\frac{X^2}{2}} (-X) \right] = e^{-\frac{X^2}{2}} (-X) (-X) \times e^{-\frac{X^2}{2}} (-1) = 0$$

$$e^{-\frac{X^2}{2}} (X^2 - 1) = 0$$

$$\therefore X = \pm 1$$

Es decir, que los puntos de inflexión de la curva se encuentran comprendidos en una distancia de más o menos una vez el error tipo, medido desde el origen.

15. **Tabla de las áreas comprendidas bajo la curva normal (curva de error).**—La siguiente tabla (Tabla 3) nos da las áreas comprendidas bajo la curva de error

$$y dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{X^2}{2}} dx,$$

como lo indica la figura 4, en donde se ha usado la misma escala para graduar el eje de las abscisas y el eje de las ordenadas. La variable  $X$  está expresada en unidades de la desviación tipo. El área sombreada que se extiende desde el centro de la curva hasta la abscisa 0.72, se ha encontrado que es igual a 0.2642, es decir, la frecuencia relativa de los errores que se encuentran comprendidos entre  $\pm 0.72$ , es igual a

$$2 (0.2642) = 0.5284$$

Hablando en términos de probabilidad, podemos decir que hay una probabilidad de 0.5284 de que en una serie de observaciones

TABLA 3.—Áreas comprendidas bajo la curva normal.

	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0586	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0909	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4428	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4706	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4818	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4953	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4965	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974



el valor absoluto de un error, seleccionado al azar, sea menor que el valor absoluto de  $0.72 \sigma$ . Por otro lado, la probabilidad de que un error aislado ocurra fuera de la clase  $0 \pm 0.72 \sigma$  es igual a

$$1 - 0.5284 = 0.4716$$

Calculando del mismo modo la probabilidad de que un error se encuentre fuera de la clase  $0 \pm k$ , obtenemos los resultados siguientes:

k (múltiplo de $\sigma$ )	Probabilidad de que un error ocurra fuera de la clase
0.5	0.6170
1.0	0.3174
1.5	0.1336
2.0	0.0456
2.5	0.0124
3.0	0.0026
3.5	0.0004
4.0	0.0000

De aquí, vemos que la probabilidad de que un error se encuentre fuera de un cierto espacio expresado en múltiplos del error tipo, decrece rápidamente tan pronto como  $k$  llega a ser mayor que 2. En otras palabras, hay una pequeña probabilidad de que el error de una observación aislada sea tan grande como 2 ó 3 veces el error tipo. Si se encuentra un error de tal tamaño, es de presumir que sean otras causas, además de las accidentales, las responsables de tan grande error.

**16. Límites convencionales de significancia.**—En la práctica es necesario demarcar algunos límites arbitrarios después de los cuales no puede considerarse a un error como accidental. Mientras que es muy razonable esperar que los errores que ocurren en una serie de observaciones, oscilen entre más o menos dos veces el error tipo, es muy improbable que un error tan grande como tres veces el error tipo, sea un error accidental; porque la probabilidad de obtener tal error por pura casualidad es únicamente de 0.0026, o sea alrededor de 3 veces en mil experimentos. Por lo tanto, un error de tal tamaño se considera como significativamente diferente de lo que nosotros esperamos.

Los niveles convencionales de significancia frecuentemente aceptados corresponden a las probabilidades de 0.05 (1 : 20) y 0.01 (1 : 100). Un resultado es llamado "significativo" (significativamente diferente de lo que se espera, tomando como base una hipótesis o una suposición razonable) cuando la probabilidad de que ocurra es más pequeña que 0.05 y un resultado es considerado como "altamente significativo" cuando la probabilidad de que ocurra es más pequeña que 0.01. Para propósitos especiales pueden ser establecidos niveles de significancia más conservadores.

Un error que excede 1.96 veces el error, ocurre con una probabilidad de 0.05; el error correspondiente a la probabilidad de 0.01 es de 2.58 veces el error tipo. Por lo tanto, también puede decirse que los niveles de significancia están comprendidos entre 2 y  $2\frac{1}{2}$  veces el error tipo. Conviene decir que usando niveles de significancia más y más conservadores, disminuye el riesgo de rechazar una buena hipótesis, pero también aumenta la probabilidad de admitir una mala hipótesis. Al aplicar la prueba de significancia, puede ocurrir que desechemos una hipótesis que en realidad es verdadera; esto es una falta del primer tipo. Si ocurre que, como resultado de una prueba, aceptamos una hipótesis que en realidad es falsa, cometemos una falta del segundo tipo. Estas dos posibles faltas deben balancearse una con otra y las consecuencias de un posible error compararse con las consecuencias de otro error. Por estas razones no debe confiarse mucho en algunos niveles de significancia convencionales.

17. **La importancia de la distribución normal no sólo en el cálculo de errores: la desviación tipo.**—No solamente los errores accidentales de observación se distribuyen aproximadamente de acuerdo con la curva de error; sino que también aparece esta distribución, llamada distribución normal, en otros datos estadísticos, tomados ya sea del campo de la medicina, de la biología, del comercio, de agricultura o de asuntos forestales. La distribución de frecuencia del peso específico de 1240 muestras de Douglas fir, es un ejemplo de esto, como puede verse en la figura 1. El número de árboles de un rodal de edades iguales, agrupados en diferentes clases diamétricas, aparecen distribuidos más o menos en tal forma, aunque, en este caso, la distribución es

algunas veces inclinada. En muchos otros casos donde la distribución de las observaciones aisladas no es estrictamente normal, se ha encontrado, sin embargo, que la distribución de los promedios de un número de observaciones aisladas tiende a ser normal. Esto explica un poco mejor la gran importancia de la curva normal en toda clase de trabajos estadísticos.

Al conjunto de individuos que forman una distribución de frecuencia, se le llama población. De acuerdo con utilísimas hipótesis de trabajo frecuentemente usadas en estadística, suponemos que una población actualmente observada representa una muestra de una población infinitamente grande, llamada población matriz. Si los individuos de una muestra de la población, aparecen distribuidos normalmente, es muy razonable suponer que la hipotética población matriz esté, asimismo, normalmente distribuida. Sin embargo, generalmente se arguye del modo opuesto: si suponemos una distribución normal específica, se investiga si una cierta serie de observaciones puede considerarse razonablemente como una muestra de población matriz normalmente distribuida.

A fin de caracterizar una población normal, lo primero que debemos conocer es el promedio aritmético de esta población. El promedio aritmético de la población matriz corresponde a lo que hemos llamado anteriormente el valor verdadero de la cantidad observada y al cual designamos con la letra  $\mu$ . A pesar de que dos poblaciones normales de cualquier tipo tengan el mismo promedio, puede haber diferencias entre una y otra: los individuos de la primera pueden estar estrechamente agrupados alrededor del promedio, mientras que los individuos de la segunda pueden estar dispersos en un espacio más ancho. Así como el error tipo es una medida de la dispersión de los errores de observación, así la llamada desviación tipo, que está definida por la misma fórmula que el error tipo, puede ser usada para medir la dispersión de los individuos de cualquier población. Las diferentes observaciones  $X_1, X_2, X_3, \dots, X_n$ , no son hechas solamente en una cantidad, sino sobre diferentes individuos, pero pertenecientes a la misma población.

**18. Significado de las diferencias.**—En toda clase de trabajos prácticos y científicos, sucede frecuentemente que tenemos que comparar los resultados de dos o más observaciones; necesi-

La primera, la segunda y la cuarta diferencia, no son, evidentemente, significativas; las otras, sin embargo, son altamente significativas.

Los resultados de este experimento nos muestran que, mientras mayor espacio se deje entre árbol y árbol en las plantaciones de Red pine, el incremento en diámetro se verá distintamente favorecido y menos definitivamente el incremento en altura.

**19. Método corto para calcular el promedio aritmético y la desviación tipo, para datos clasificados.**—Cuando se trabaja con gran distribución, de 50 ó más observaciones, los individuos son generalmente agrupados en clases, ya sea después o al mismo tiempo que se hacen las observaciones. Al hacer un inventario forestal, nosotros determinamos, por ejemplo, directamente el número de árboles para cada categoría diamétrica. A fin de determinar el promedio aritmético y la desviación tipo para datos clasificados, es conveniente trabajar con un promedio supuesto  $A$ , que nos simplificará mucho el trabajo de cálculo. Con la introducción de este promedio supuesto, las fórmulas originales para  $\bar{X}$  y  $S$ , se transforman del modo siguiente:

$$(i) \quad \bar{X} = \frac{\sum n_i X_i}{N} = \frac{\sum n_i X_i - N A}{N} + A$$

$$= \frac{\sum n_i (X_i - A)}{N} + A$$

$$(ii) \quad S^2 = \frac{\sum n (X_i - \bar{X})^2}{N} = \frac{\sum n [(X_i - A) - (\bar{X} - A)]^2}{N}$$

$$= \frac{\sum n_i (X_i - A)^2}{N} - 2(\bar{X} - A) \frac{\sum n_i (X_i - A)}{N} + (\bar{X} - A)^2 \frac{\sum n}{N}$$

$$S^2 = \frac{\sum n_i (X_i - A)^2}{N} - 2(\bar{X} - A) (\bar{X} - A) + (\bar{X} - A)^2$$

$$= \frac{\sum n_i (X_i - A)^2}{N} - (\bar{X} - A)^2$$

$$S = \sqrt{\frac{\sum n_i (X_i - A)^2}{N} - (\bar{X} - A)^2}$$

A fin de simplificar los cálculos, es claro que el promedio supuesto  $A$  deba ser escogido de tal modo que la diferencia entre la clase y dicho promedio  $(X_i - A)$  sea lo más corta posible. El

uso de estas fórmulas es fácilmente comprendido con el siguiente ejemplo.

**Ejemplo 11).**—En el laboratorio de productos forestales de Madison, Wisconsin, L. J. Markwardt y T. R. C. Wilson, determinaron el peso específico de 1,240 muestras de Douglas fir (tipo de la costa). El peso específico de cada muestra fué calculado hasta tres cifras decimales. Las muestras cuyo peso específico son de:

0.420, 0.421, 0.422..... 0.429 inclusive;  
0.430, 0.431, 0.432..... 0.439 inclusive, etc.,

fueron combinadas en clases cada serie. Suponiendo que las cifras calculadas fueron arredondadas hasta milésimos, los límites de clases de las dos series anteriores, serán:

0.4195    y    0.4295  
0.4295    „    0.4395;

y los correspondientes puntos medios de las clases, serán:

0.4245  
0.4345

En la siguiente tabla se consignan: el número de muestras que tienen un peso específico dentro de las clases indicadas, y los cálculos necesarios para obtener el promedio aritmético  $\bar{X}$  y la desviación tipo S. El promedio supuesto para estos cálculos es

0.4345;

las diferencias

$X_i - A$

serán, por lo tanto, con sólo dos cifras decimales. Como vamos a calcular al mismo tiempo el promedio aritmético  $\bar{X}$  y la desviación tipo S, es más fácil multiplicar  $n_i (X_i - A)$  por  $(X_i - A)$ , a fin de obtener  $n_i (X_i - A)^2$  en vez de calcular primero  $(X_i - A)^2$  y luego multiplicar por  $n_i$ .

TABLA 4.—Cálculos numéricos del promedio aritmético y de su desviación tipo, del peso específico de 1,240 muestras de Douglas Fir.

Promedio supuesto A = 0.4345

Peso específico límite de clases	Mitad de clases $X_i$	Número de muestras en las clases (frecuencia) $n_i$	$(X_i-A)$	$(X_i-A)n_i$	$(X_i-A)^2n_i$
0.2995—0.3095	0.3045	2	—0.13	—0.26	0.0338
0.3095—0.3195	0.3145	7	—0.12	—0.84	0.1008
0.3195—0.3295	0.3245	6	—0.11	—0.66	0.0726
0.3295—0.3395	0.3345	15	—0.10	—1.50	0.1500
0.3395—0.3495	0.3445	13	—0.09	—1.17	0.1053
0.3495—0.3595	0.3545	23	—0.08	—1.84	0.1472
0.3595—0.3695	0.3645	25	—0.07	—1.75	0.1225
0.3695—0.3795	0.3745	38	—0.06	—2.28	0.1368
0.3795—0.3895	0.3845	47	—0.05	—2.35	0.1175
0.3895—0.3995	0.3945	64	—0.04	—2.56	0.1024
0.3995—0.4095	0.4045	75	—0.03	—2.25	0.0675
0.4095—0.4195	0.4145	85	—0.02	—1.70	0.0340
0.4195—0.4295	0.4245	76	—0.01	—0.76	0.0076
0.4295—0.4395	0.4345	99	0.00	0.00	0.0000
0.4395—0.4495	0.4445	100	0.01	1.00	0.0100
0.4495—0.4595	0.4545	90	0.02	1.80	0.0360
0.4595—0.4695	0.4645	96	0.03	2.88	0.0864
0.4695—0.4795	0.4745	74	0.04	2.96	0.1184
0.4795—0.4895	0.4845	70	0.05	3.50	0.1750
0.4895—0.4995	0.4945	56	0.06	3.36	0.2016
0.4995—0.5095	0.5045	46	0.07	3.22	0.2254
0.5095—0.5195	0.5145	41	0.08	3.28	0.2624
0.5195—0.5295	0.5245	30	0.09	2.70	0.2430
0.5295—0.5395	0.5345	23	0.10	2.30	0.2300
0.5395—0.5495	0.5445	12	0.11	1.32	0.1452
0.5495—0.5595	0.5545	9	0.12	1.08	0.1296
0.5595—0.5695	0.5645	10	0.13	1.30	0.1690
0.5695—0.5795	0.5745	4	0.14	0.56	0.0784
0.5795—0.5895	0.5845	1	0.15	0.15	0.0225
0.5895—0.5995	0.5945	3	0.16	0.48	0.0768
Sumas		1240		11.97	3.4077

$$\bar{X} = \frac{\sum m_i (X_i - A)}{N} + A = \frac{11.97}{1240} + 0.4345 = 0.0097 + 0.4345 = 0.4442$$

$$S = \frac{\sum m_i (X_i - A)^2}{N} - (\bar{X} - A)^2 = \sqrt{\frac{3.4077}{1240} - (0.4442 - 0.4345)^2}$$

$$= \sqrt{0.002748 - 0.000094} = \sqrt{0.002654}$$

$$= 0.0515$$

**20. La aplicación de la curva de error a una distribución observada.**—Aunque este problema no concierne directamente al cálculo de errores, discutiremos ahora esto porque los datos usados en el inciso anterior (ejemplo 11), se prestan muy bien para ilustrar el procedimiento.

Los únicos parámetros que aparecen en la expresión matemática de la curva normal, son  $\mu$  y  $\sigma$ . De aquí que primero debemos calcular el promedio aritmético  $\bar{X}$ , y la desviación tipo  $S$ , que representan una estimación de  $\mu$  y de  $\sigma$ . En el ejemplo antes mencionado se encontraron para  $\bar{X}$  y  $S$  los valores 0.4442 y 0.0515, respectivamente. El problema de ajustar la curva normal a la distribución de frecuencia observada, consiste en determinar las frecuencias esperadas en las diferentes clases calculadas, tomando como base la curva normal caracterizada por el promedio aritmético 0.4442 y la desviación tipo 0.0515. Primero hay que calcular la desviación de los diferentes límites de clase con respecto al promedio aritmético; estas desviaciones se expresan después en unidades de la desviación tipo. De la tabla de las áreas comprendidas bajo la curva normal, podemos obtener por interpolación, las áreas comprendidas desde el promedio hasta los límites de clase y luego, haciendo las diferencias sucesivas de estas áreas, obtenemos la frecuencia relativa para cada clase. Obtenidas éstas, debemos multiplicarlas por 1,240, número total de observaciones, a fin de obtener las frecuencias absolutas esperadas en las diferentes clases. Estas pueden compararse con las frecuencias absolutas observadas.

**21. Algunas consideraciones sobre la corrección de Shepard.**—

Si nosotros calculamos la desviación tipo para datos agrupados, como hemos hecho en el ejemplo anterior, algunas veces es necesario introducir un factor de corrección. Por el proceso de agrupar los datos, cada observación es, en efecto, transferida y registrada en el centro de la clase en donde ella aparece, siendo que en realidad, la magnitud de una observación puede aparecer en cualquier punto del intervalo de la clase dada. Una consecuencia de esta clasificación, es una cierta variación en los momentos de una distribución de frecuencia, que ha sido descrita por Shepard. Se ha encontrado que el valor de la desviación tipo (momento segundo) de una distribución, la cual decrece gradualmente en ambas direcciones hasta más y menos infinito, como

TABLA 5.—Cálculos numéricos para aplicar la curva normal.

Límite de clases	Límite de clase $\bar{X}$ = $\Delta$	$\frac{\Delta}{s}$	Area bajo la curva normal desde el promedio hasta $\frac{\Delta}{s}$	Diferencias de áreas entre dos límites de clase	Frecuencia absoluta esperada
		— $\infty$	0.5000		
0.2995	—0.1447	—2.810	0.4975	0.0025	3.1
0.3095	—0.1347	—2.616	0.4956	0.0019	2.4
0.3195	—0.1247	—2.421	0.4922	0.0034	4.2
0.3295	—0.1147	—2.227	0.4870	0.0052	6.4
0.3395	—0.1047	—2.033	0.4790	0.0080	9.9
0.3495	—0.0947	—1.839	0.4670	0.0120	14.9
0.3595	—0.0847	—1.645	0.4500	0.0170	21.1
0.3695	—0.0747	—1.451	0.4266	0.0234	29.0
0.3795	—0.0647	—1.256	0.3955	0.0311	38.6
0.3895	—0.0547	—1.062	0.3559	0.0396	49.1
0.3995	—0.0447	—0.868	0.3073	0.0486	60.3
0.4095	—0.0347	—0.674	0.2498	0.0575	71.3
0.4195	—0.0247	—0.480	0.1844	0.0654	81.1
0.4295	—0.0147	—0.285	0.1122	0.0722	89.5
0.4395	—0.0047	—0.091	0.0363	0.0759	94.1
0.4495	0.0053	0.103	0.0410	0.0773	95.9
0.4595	0.0153	0.297	0.1168	0.0758	94.0
0.4695	0.0253	0.491	0.1883	0.0715	88.7
0.4795	0.0353	0.685	0.2533	0.0650	80.6
0.4895	0.0453	0.880	0.3106	0.0573	71.0
0.4995	0.0553	1.074	0.3586	0.0480	59.5
0.5095	0.0653	1.268	0.3976	0.0390	48.4
0.5195	0.0753	1.462	0.4282	0.0316	37.9
0.5295	0.0853	1.656	0.4511	0.0229	28.4
0.5395	0.0953	1.851	0.4679	0.0168	20.8
0.5495	0.1053	2.045	0.4796	0.0117	14.5
0.5595	0.1153	2.239	0.4875	0.0077	9.8
0.5695	0.1253	2.433	0.4926	0.0051	6.3
0.5795	0.1353	2.627	0.4957	0.0031	3.8
0.5895	0.1453	2.821	0.4976	0.0019	2.4
0.5995	0.1553	3.016	0.4987	0.0011	1.4
			0.5000	0.0013	1.6
Suma				1.0000	1240.0

sucede por ejemplo con la distribución normal, llega a ser demasiado elevado cuando proviene de una serie de datos agrupados y se hace necesario aplicar un factor de corrección igual a

$$\frac{a^2}{12}$$

donde  $a$  es igual al ancho de la clase. Esta cantidad debe ser substraída del cuadrado de la desviación tipo calculada a base de datos agrupados; esta cantidad es el promedio de lo que aumenta el momento segundo al agrupar los datos.

Como la deducción exacta de este factor de corrección es algo complicada, la vamos a omitir; sin embargo, indicaremos de qué modo el denominador 12 aparece en los cálculos; de una manera somera, argüimos lo siguiente: El error  $\epsilon$ , cometido al suponer una observación como de la magnitud del centro de la clase, fluctúa entre  $-\frac{a}{2}$  y  $+\frac{a}{2}$ . Suponiendo que las observaciones están igualmente repetidas en toda la clase de tamaño, el error tipo de la agrupación de las observaciones se obtiene fácilmente evaluando la expresión:

$$\frac{1}{a} \int_{-\frac{a}{2}}^{+\frac{a}{2}} \epsilon^2 d\epsilon$$

De aquí obtenemos:

$$\frac{1}{a} \int_{-\frac{a}{2}}^{+\frac{a}{2}} \epsilon^2 d\epsilon = \frac{1}{a} \left[ \frac{\epsilon^3}{3} \right]_{-\frac{a}{2}}^{+\frac{a}{2}} = \frac{1}{3} \left[ \left(\frac{a}{2}\right)^3 + \left(\frac{a}{2}\right)^3 \right] = \frac{1}{3} \cdot \frac{a^3}{4} = \frac{a^2}{12}$$

Es necesario hacer hincapié en que ésta no es la deducción de la corrección de Sheppard, sino solamente una indicación del modo cómo se origina el valor.

En el ejemplo 11 el ancho de las clases fué

$$a = 0.01$$

el valor de la corrección de Sheppard es, por lo tanto, de

$$\frac{a^2}{12} = \frac{0.0001}{12} = 0.000008$$

Substrayendo esta cantidad del cuadrado de la desviación tipo previamente encontrada obtenemos:

$$S^2 = (0.0515)^2 - 0.000008$$

$$= 0.002652 - 0.000008$$

$$S = 0.0514$$

valor que es ligeramente más pequeño que 0.0515. En otros casos, la corrección de Sheppard puede ser de gran importancia.

Es indispensable decir que la corrección de Sheppard sólo se aplica a distribuciones continuas y no a las distribuciones discontinuas, donde la variable toma valores discretos, como por ejemplo, números enteros.

**22. Exactitud de la determinación del incremento y del volumen, como una aplicación del cálculo de errores.**—El problema de determinar la exactitud de los inventarios forestales y del cálculo de los incrementos, sirve muy bien como un ejemplo ilustrativo de los varios métodos usados en el cálculo de errores, explicados en los incisos anteriores. Muestra, al mismo tiempo, la complejidad con que un problema puede presentarse en este campo de la estadística.

Es conveniente señalar que aquí no vamos a tratar con los errores de muestreo, los cuales ocurren cuando se recorre el bosque por fajas o por sitios de ensayo; solamente consideraremos los errores de medición que ocurren cuando se hace un inventario de una determinada área de bosque. Después investigaremos la influencia de estos errores de medida sobre la determinación del incremento, basada en las medidas sucesivas del bosque.

A fin de simplificar algo el problema, consideraremos solamente los errores cometidos al forcipular los árboles de un rodal, suponiendo que no ocurren errores adicionales cuando se calcula el volumen del rodal por el uso de las curvas de alturas y de las tablas de volúmenes. En otras palabras, solamente nos interesamos en la exactitud relativa de las determinaciones del volumen por árbol, como una función estable de su diámetro a la altura del pecho; incidentalmente este concepto es de gran importancia práctica para el control periódico del volumen y del incremento.

a). Relación entre el porcentaje de error en área basimétrica y el porcentaje de error en volumen.

Si el error tipo de los diámetros medidos  $X$  es  $\sigma_x$ , entonces el error tipo del área basimétrica  $\sigma_A$  se ha encontrado que es igual a (inciso 7, ejemplo 4):

$$\sigma_A = \frac{d}{dX} \left( \frac{\pi}{4} X^2 \right) \sigma_x = \frac{\pi}{2} X \sigma_x \quad (1)$$

y el porcentaje de error de la superficie basal, será de:

$$\sigma_A \% = \frac{\frac{\pi}{2} X \sigma_x}{\frac{\pi}{4} X^2} \cdot 100 = \frac{200 \sigma_x}{X} \quad (2)$$

Por otro lado, el volumen de un árbol puede ser expresado como una función de su diámetro a la altura del pecho, por medio de la fórmula

$$V = c X^p$$

en donde la constante  $p$  tiene un valor algo menor que 2.5, pero siempre mayor que 2.0, dependiendo de la especie, sitio, etc. (Si a  $p$  se le da un valor de 2.0 y a  $c$  el de  $\frac{\pi}{4}$ , obtenemos la fórmula del área basimétrica). El error tipo del volumen,  $\sigma_v$ , es igual a

$$\sigma_v = \frac{d}{dX} (c X^p) \sigma_x = c p X^{p-1} \sigma_x$$

De donde fácilmente obtenemos el porcentaje del error en volumen

$$\sigma_v \% = \frac{c p X^{p-1} \sigma_x}{c X^p} \cdot 100 = \frac{100 p \sigma_x}{X}$$

Sacando de la fórmula (2)

$$\frac{\sigma_x}{X} = \frac{\sigma_A \%}{200}$$

podemos expresar el porcentaje de error en volumen en función del porcentaje de error en área basimétrica:

$$\sigma_v \% = 100 p \frac{\sigma_A}{200} = \frac{P}{2} \sigma_A \%$$

Este resultado significa que solamente tenemos que multiplicar el porcentaje de error en área basimétrica por el factor  $\frac{P}{2}$  a fin de obtener el porcentaje de error en volumen. Esta relación entre el porcentaje del error tipo del área basimétrica y el

del volumen se aplica, evidentemente, sólo a las mediciones de un árbol aislado, o a las mediciones de un número de árboles del mismo diámetro y del mismo error. Sin embargo, con datos experimentales fué posible demostrar que la misma fórmula da muy buenos resultados también cuando se trata de determinar el error de las mediciones en volumen de un bosque entero, compuesto de árboles de diferentes diámetros. De todos modos, tomando para su máximo valor de 2.5, siempre se tiene la seguridad de decir que el porcentaje de error en volumen no es mayor que 1.25 el porcentaje de error en área basimétrica. Para encontrar el porcentaje de error en volumen de un bosque, basta con determinar solamente el porcentaje de error en área basimétrica y multiplicar este valor por 1.25.

Ahora discutiremos separadamente los diferentes tipos de errores que se cometen cuando se mide el diámetro de los árboles por medio de la forcípula; después será investigado el efecto combinado de estos errores. Finalmente, discutiremos el error del incremento, determinado por inventarios sucesivos del bosque.

b). El error sistemático de la forcípula.

Al medir el diámetro de los árboles con la forcípula, observaremos siempre que los dos brazos de ella no están exactamente paralelos cuando se aplican al árbol; esto produce un error sistemático en las mediciones, error que se designa como error de la forcípula. A fin de expresar este error numéricamente, lo mejor es dividir la desviación del brazo móvil (medida en el extremo) de la posición paralela, entre la longitud del brazo, eliminando así la influencia de dicha longitud sobre la desviación arriba mencionada. Esta relación  $\frac{d}{l}$  será designada con la letra

h. Como la forcípula se ajusta con precisión alrededor del árbol, la desviación del brazo móvil de su posición paralela, con respecto al brazo fijo, es generalmente en la dirección externa. Cuando se mide un árbol de diámetro X, los brazos de la forcípula tocan al árbol a una distancia de la regla igual a  $\frac{X}{2}$ .

El error en la medición del diámetro será, por lo tanto, igual a  $\frac{hX}{2}$ , y si la desviación es en la dirección antes indicada, el diámetro medido será  $\frac{hX}{2}$  más chico. El error del área basimétrica correspondiente a este diámetro será igual a

$$\frac{d}{dX} \left( \frac{\pi}{4} X^2 \right) \frac{hX}{2} = \frac{\pi}{4} X^2 h$$

y el porcentaje de error del área basimétrica

$$\frac{\frac{\pi}{4} X^2 h}{\frac{\pi}{4} X^2} = 100 = 100 h.$$

Cualquier desviación del brazo movable de la forcípula de su posición paralela, da por resultado un constante porciento de error del área basimétrica para todas las clases diamétricas (el porcentaje de error depende solamente del valor de  $h$ ) y, por lo tanto, para un bosque entero. La forcípula debe ajustarse y corregirse siempre lo mejor que se pueda para hacer que este error sistemático sea lo más pequeño posible. Las forcípulas usadas en trabajos prácticos frecuentemente no son hechas con mucho cuidado, sino de una manera burda, lo que hace que el error en área basimétrica sea del 1% o mayor.

c). El error accidental de la medición.

Este error tiene por origen las irregularidades de la sección transversal de los árboles, a la altura del pecho. Por eso, cuando se mide con forcípulas el diámetro de un árbol en diferentes lugares y especialmente en diferentes direcciones, se obtienen resultados diferentes. Es posible demostrar experimentalmente que estos errores son directamente proporcionales al tamaño del árbol. El error tipo  $\sigma_x$  de la medición del diámetro  $X$  puede, por lo tanto, ser expresado por la función

$$\sigma_x = c X$$

en donde  $c$  es una constante que debe ser determinada empíricamente. El mejor procedimiento para esto consiste en forcipular varias veces uno o más rodales, exactamente de la misma manera como se hace un inventario forestal en la práctica. Para evitar que un mismo árbol sea medido dos veces en el mismo inventario o que se omita uno, es necesario marcar los árboles de un modo conveniente, después de cada medición.

En una investigación de este tipo, se midieron cinco rodales, designados con A, B, C, D y E, aplicando categorías diamétricas de 1 centímetro. En la tabla 6 están consignadas las mediciones hechas en el sitio E, de 1.7 Ha. A fin de que la tabla no

resulte demasiado grande, se agruparon los árboles en clases diamétricas de 4 centímetros. El inventario fué repetido 12 veces en cada rodal. Se ve que el número de árboles, en las categorías diamétricas de 4 centímetros, varía constantemente de una medición a otra. Además, el número total de árboles no es siempre el mismo; lo que demuestra que, a pesar de todas las precauciones tomadas para evitar cometer errores, algunas veces se omitió o se midió dos veces un mismo árbol. En la tabla 7 están calculadas las áreas basimétricas correspondientes al número de árboles en las diferentes categorías diamétricas. Basándose en los doce totales, podemos calcular el área basimétrica promedio y el error tipo del inventario:

$$\bar{A} = \frac{100.98 + 101.42 + \dots}{12} = 100.97$$

$$S_A^2 = \frac{(100.98 - 100.97)^2 + (101.42 - 100.97)^2 + \dots}{12 - 1} = 0.0813$$

$$S_A = 0.29 \text{ m}^2$$

Suponiendo ahora que conocemos la constante  $c$ , que aparece en la ecuación  $\sigma_x = cX$ , la cual da el error tipo de las mediciones del diámetro, podemos calcular el correspondiente error tipo del área basimétrica; de acuerdo con las explicaciones anteriores, tenemos:

$$\sigma_a = \frac{\pi}{2} X \sigma_x$$

$$\sigma_a = \frac{\pi}{2} \bar{X} c X = K X^2$$

en donde

$$K = \frac{\pi}{2} c$$

El símbolo  $\sigma$  subíndice  $a$ , es usado para designar el error tipo del área basimétrica de un árbol aislado;  $\sigma$  subíndice  $A$ , para designar el error en área basimétrica de un bosque entero; y con las

TABLA NUMERO 6.—Inventario de un sitio de ensayo de 1 7 Ha., medido 12 veces

Diámetro a la altura del pecho	MEDICION NUMERO											
	1	2	3	4	5	6	7	8	9	10	11	12
em.	NUMERO DE ARBOLES											
15-19	30	33	30	29	30	31	31	31	30	34	31	31
19-23	168	163	168	176	172	170	169	175	172	172	170	172
23-27	192	200	199	202	199	202	198	200	202	199	211	196
27-31	224	213	218	213	219	215	220	215	220	213	212	216
31-35	202	212	206	210	206	210	210	215	213	210	219	203
35-39	166	157	161	154	158	160	152	151	148	153	152	163
39-43	106	106	106	100	102	101	108	102	108	102	99	105
43-47	52	55	53	56	57	54	54	55	51	55	54	56
47-51	34	33	30	36	34	32	31	32	34	34	32	31
51-55	7	14	14	9	9	13	13	14	11	10	10	13
55-59	9	7	8	7	8	7	6	8	7	7	9	6
59-63	4	2	2	3	3	3	4	2	3	4	4	4
63-67	4	4	5	5	5	5	5	5	5	4	4	4
67-71	1	1	...	...	...	...	...	...	...	1	1	1
Total. . . . .	1198	1200	1200	1200	1202	1203	1201	1205	1204	1198	1208	1201

TABLA NUMERO 7.—Inventario de un sitio de ensayo de 1 7 Ha., medido 12 veces

Diámetro a la altura del pecho	MEDICION NUMERO												
	1	2	3	4	5	6	7	8	9	10	11	12	
cm.	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>	m <sup>2</sup>
15—19	0,68	0,75	0,68	0,66	0,68	0,70	0,70	0,70	0,68	0,77	0,70	0,70	0,70
19—23	5,82	5,65	5,82	6,09	5,96	5,89	5,85	6,06	5,96	5,96	5,89	5,89	5,96
23—27	9,42	9,82	9,77	9,91	9,77	9,92	9,72	9,82	9,91	9,77	10,36	10,36	9,62
27—31	14,80	14,07	14,40	14,07	14,46	14,20	14,53	14,20	14,53	14,07	14,00	14,00	14,28
31—35	17,28	18,13	17,62	17,96	17,62	17,96	17,96	18,39	18,22	17,99	18,73	18,73	17,36
35—39	17,85	16,88	17,31	16,56	16,99	17,20	16,34	16,24	15,91	16,45	16,34	16,34	17,52
39—43	14,00	13,99	13,99	13,20	13,47	13,34	14,26	13,47	14,26	13,47	13,07	13,07	13,86
43—47	8,27	8,75	8,43	8,90	9,06	8,59	8,59	8,75	8,11	8,75	8,59	8,59	8,90
47—51	6,41	6,22	5,66	6,79	6,41	6,03	5,85	6,03	6,41	6,41	6,03	6,03	5,85
51—55	1,54	3,09	3,09	1,99	1,99	2,87	2,87	3,09	2,43	2,20	2,21	2,21	2,87
55—59	2,04	1,79	2,04	1,79	2,04	1,78	1,53	2,04	1,79	1,79	2,30	2,30	1,53
59—63	1,17	0,58	0,58	0,88	0,88	0,88	1,17	0,58	0,88	1,17	1,17	1,17	1,17
63—67	1,33	1,33	1,66	1,66	1,66	1,66	1,66	1,66	1,66	1,33	1,33	1,33	1,33
67—71	0,37	0,37	...	...	...	...	...	...	...	0,37	0,37	0,37	0,37
Total...	100,98	101,42	101,05	100,46	100,99	101,02	101,03	101,03	100,75	100,47	101,09	101,09	101,32

## AREA BASIMETRICA

Promedio aritmético: 100,97 m<sup>2</sup>.

letras  $n_1, n_2, \dots$  el número de árboles en las clases diamétricas, cuyos puntos medios son  $X_1, X_2, \dots$ . Aplicando la ley de propagación de errores, obtenemos:

$$\begin{aligned} \sigma_A &= \sqrt{n_1 \sigma_{a1}^2 + n_2 \sigma_{a2}^2 + \dots} \\ &= \sqrt{n_1 (k_1 X_1^2)^2 + n_2 (k_2 X_2^2)^2 + \dots} \\ &= K \sqrt{n_1 X_1^4 + n_2 X_2^4 + \dots} \\ &= K \sqrt{\sum n_i X_i^4} \end{aligned}$$

La expresión subradical es fácilmente calculada a partir de nuestra tabla de existencia (tabla 6); en nuestro ejemplo obtenemos:

TABLA 8

Dímetros, mitad de categoría $X_i$ m	$n_i$	$X_i^4$	$n_i X_i^4$
0.17	31	0.000835	0.026
0.21	171	0.001945	0.332
0.25	200	0.003906	0.781
0.29	217	0.007073	1.535
0.33	210	0.011859	2.490
0.37	156	0.018742	2.924
0.41	104	0.028258	2.939
0.45	54	0.041006	2.214
0.49	33	0.057648	1.902
0.53	11	0.078565	0.868
0.57	7	0.106560	0.739
0.61	3	0.138458	0.415
0.65	5	0.178506	0.893

$$\begin{aligned} \sum n_i X_i^4 &= 18.058 \\ \sqrt{\sum n_i X_i^4} &= 4.25 \end{aligned}$$

Como el valor de  $\sigma_A$  ha sido determinado experimentalmente y se le ha encontrado un valor de  $0.29 \text{ m}^2$ , substituimos en la ecuación sus valores y obtenemos, para K:

$$K = \frac{0.29}{4.25} = 0.068$$

De este modo se han calculado diferentes valores de K, para diferentes rodales y usando clases diamétricas de diferente tamaño. Estos valores se encuentran en la tabla 9.

TABLA 9.—Valores de K obtenidos en diferentes rodales y para diferentes clases diamétricas

Rodal	Area del rodal Ha.	Núm. de árboles medidos	Tamaño de las clases diamétricas						Promedio de K
			1 cm.	2cm.	3 cm.	4 cm.	5 cm.	6 cm.	
A	2.0	687	0.038	0.043	0.056	0.046	0.033	0.083	0.050
B	1.5	446	57	55	64	53	60	56	58
C	2.0	1096	55	59	60	58	73	67	62
D	1.6	289	64	65	68	68	60	69	66
D	1.7	1202	60	68	63	68	85	91	72
Promedio			0.055	0.058	0.062	0.059	0.062	0.073	0.062

Comparando los valores medios de K en las diferentes clases diamétricas, observamos que el uso de clases diamétricas cuyos intervalos sean hasta de 5 cm. de tamaño, no hace que disminuya apreciablemente la exactitud del inventario. Esto es posible confirmarlo teóricamente, pero ya no entraremos en más consideraciones sobre este punto. (1)

d). El error de posición.

Cuando se mide el diámetro de un árbol a la altura del pecho, nosotros no registramos su valor verdadero, sino que solamente determinamos la clase diamétrica dentro de la cual se encuentra dicho árbol, anotando a éste como si su diámetro fuera de la magnitud de la mitad de la categoría diamétrica a que corresponde. Si formamos, por ejemplo, clases diamétricas de 5 cm. de ancho, con los límites a 17.5, 22.5, 27.5, el punto medio de estas clases será 20 y 25 centímetros, respectivamente; por lo tanto, cuando se mide un árbol de 21.2 centímetros y se registra solamente como perteneciente a la clase diamétrica de 20 centímetros, se comete un error de 1.2 centímetros; a este error se le llama error de posición. A mayor tamaño de las clases diamétricas, mayor es el error de posición cometido. Este error debe

(1) H. Arthur Meyer y W. Edwards Deming. On the influence of clasification on the determination of a measure characteristic. Journal of the American Statiscal Association. Vol. 30, pags. 671-677, 1935.

considerarse evidentemente como error accidental, ya que al medir un árbol de cierto diámetro, el error de posición puede diferir cada vez que se cambien los límites de las clases diamétricas. Por otro lado, si consideramos los diámetros límites dados, entonces, el diámetro de cualquier árbol que esté dentro de cierta clase diamétrica puede diferir del valor del punto medio de la clase, de acuerdo con la "posición" del árbol; la diferencia  $E$  podrá variar desde 0, como mínimo, hasta  $\frac{a}{2}$  como máximo, en donde  $a$  es el ancho de la clase diamétrica. Exactamente como se ha obtenido el valor de  $\frac{a^2}{12}$  de la corrección Sheppard se puede calcular la desviación tipo del error de posición, obteniendo el valor de  $\frac{a}{\sqrt{12}}$

A fin de calcular el error de registro del área basimétrica, aplicamos la ley de propagación de errores, y obtenemos:

$$\sigma_a = \frac{\pi}{2} \times \sigma_x = \frac{\pi}{2} \times \frac{a}{\sqrt{12}}$$

Midiendo un rodal de  $n_i$  árboles cuyos diámetros sean  $X_i$ , el error tipo de posición será igual a:

$$\sigma_A = \sqrt{n_1 \left( \frac{\pi}{2} X_1 \frac{a}{\sqrt{12}} \right)^2 + n_2 \left( \frac{\pi}{2} X_2 \frac{a}{\sqrt{12}} \right)^2 + \dots}$$

$$\sigma_A = \frac{\pi}{2} \frac{a}{\sqrt{12}} \sqrt{n_1 X_1^2 + n_2 X_2^2 + \dots}$$

Cuando se mide el mismo rodal varias veces y se usan las mismas clases diamétricas en cada medición, el error de posición de los diferentes árboles, permanecerá siendo el mismo. Esta es la razón por la cual no es necesario tomarlo en cuenta cuando se determinan los errores accidentales de medición. Por lo tanto, vemos que los errores de observación hechos cuando se hace un inventario forestal, constan de dos errores independientes: el error accidental de medición y el error accidental de posición.

El error tipo de posición del mismo rodal usado previamente, está calculado en la tabla 10.

TABLA 10

Diámetros, mitad de categoría	$X_i$ m	$n_i$	$X_i^2$	$n_i X_i^2$
	0.17	31	0.0289	0.90
	0.21	171	0.0441	7.54
	0.25	200	0.0625	12.50
	0.29	217	0.0841	18.25
	0.33	210	0.1089	22.86
	0.37	156	0.1369	21.34
	0.41	104	0.1681	17.48
	0.45	54	0.2025	10.94
	0.49	33	0.2401	7.92
	0.53	11	0.2809	3.09
	0.57	7	0.3249	2.27
	0.61	3	0.3721	1.02
	0.65	5	0.4225	2.11
$\Sigma n_i X_i^2 = 128.22$				
$\sqrt{\Sigma n_i X_i^2} = 11.3$				

El error tipo de posición en área basimétrica para clases diamétricas de 4 cm., es finalmente:

$$\frac{\pi}{2} \frac{00.4}{12} 11.3 = 0.20 \text{ m}^2$$

e). El error total de medición de inventario forestal.

De los dos errores accidentales independientes: el error de medición  $\sigma_M$  y el error tipo de posición  $\sigma_P$ , encontramos para el error tipo total de un inventario forestal  $\sigma_{MP}$ , combinando los dos primeros errores de acuerdo con la fórmula

$$\sigma_{MP} = \sqrt{\sigma_M^2 + \sigma_P^2}$$

Substituyendo en esta expresión los valores obtenidos en nuestro ejemplo, obtenemos:

$$\sigma_{MP} = \sqrt{(0.20)^2 + (0.20)^2} = 0.35 \text{ m}^2$$

para el error tipo total en área basimétrica. Siendo el área basimétrica del rodal mismo igual a

$$100.97 \text{ m}^2$$

encontramos que el porcentaje de error en área basimétrica es igual a 0.35%. El porcentaje de error en volumen es alrededor de 1.25 veces mayor, es decir, 0.44%.

Este error accidental no incluye el error sistemático que puede ocurrir al usar una forcípula imprecisa. Hay que mencionar, también, el uso de clases diamétricas que introduce aún otro error sistemático; si los árboles no están igualmente distribuidos dentro de la clase diamétrica, sino que aparecen, por ejemplo más frecuentemente cerca del límite inferior de la clase diamétrica entonces la suma de todos los diámetros de estos árboles aparecerá demasiado grande, puesto que al registrarlos los consideramos como si fueran del tamaño de la mitad de la categoría diamétrica. Si se parte de una distribución de árboles claramente definida, es posible evaluar este error no solamente en función del diámetro, sino también en función del área basimétrica. El resultado de estos cálculos muestra que este error sistemático es muy pequeño y su importancia práctica lo mismo, por esta razón no será discutido aquí. (1)

f). La exactitud de las determinaciones de los incrementos.

Cuando se determina el crecimiento corriente de un bosque tomando como base los inventarios sucesivos, es importante conocer el error tipo del resultado de este cálculo. Este punto ya ha sido discutido brevemente en el ejemplo 1 del inciso 5; se encontró que el porcentaje de error del incremento es mucho más grande y, por lo tanto, mucho más importante que el porcentaje de error en volumen del mismo bosque.

Ahora discutiremos un punto interesante.

Mientras que el error de la forcípula debe ser considerado como un error sistemático cuando influye sobre la exactitud de la determinación del volumen o del área basimétrica de un bosque, cuando se calculan incrementos, este mismo error debe considerarse como un error accidental de observación. Si el error de la forcípula fué más grande o más chico cuando se hizo el primero de los dos inventarios sucesivos, el incremento calculado

---

(1) Véase H. Arthur Meyer. Die rechnerischen Grundlagen der Kontrollmethoden. (Genauigkeit der stehenden Vorrats- und Zuwachskontrolle.) Beiheft N° 13 zuden Zeitschriften des Schweizerischen Forstvereine, Zürich, 1934. En esta publicación ha sido calculado el error sistemático para las diferentes distribuciones. Todos los datos numéricos usados en este inciso, son tomados de este mismo trabajo.

será demasiado chico. Si los errores de las forcípulas fueron del mismo tamaño en ambos casos, estos errores se anularán. Por lo tanto, debemos calcular el error tipo del incremento debido a los errores instrumentales de las forcípulas.

Designemos el porcentaje de error en área basimétrica, causado por el uso de forcípulas defectuosas en el primero y segundo inventarios por  $x$  y  $y$ , respectivamente y supongamos que estos errores toman cualquier valor entre 0 y  $r\%$  con igual probabilidad. Los porcentajes están referidos al mismo promedio de área basimétrica en el bosque. El error tipo del incremento  $\sigma_1$  puede ser ahora calculado por la siguiente expresión:

$$[\sigma_1 \text{ \%}]^2 = \frac{1}{r^2} \int_0^r \int_0^r (y-x)^2 dx dy$$

los errores aislados  $(y-x)$  están elevados al cuadrado y multiplicados por sus respectivas probabilidades  $\frac{dx}{r} \frac{dy}{r}$ , y después sumadas para todo el intervalo desde 0 hasta  $r$ . Evaluando esta expresión para un valor máximo de  $r=1\%$ , obtenemos:

$$\begin{aligned} [\sigma_1 \text{ \%}]^2 &= \int_0^1 \int_0^1 (y^2 - 2xy + x^2) dx dy \\ &= \left| x \frac{y^3}{3} - 2 \frac{x^2}{2} \frac{y^2}{2} + \frac{x^3}{3} y \right|_0^1 \\ &= \frac{1}{3} - \frac{2}{4} + \frac{1}{3} = \frac{1}{6} \\ \sigma_1 \text{ \%} &= \frac{1}{\sqrt{6}} = 0.4 \text{ \%} \end{aligned}$$

Este error debe ser combinado con el error accidental de medición y con el error accidental de posición. Podemos expresar todos estos errores en por ciento del área basimétrica promedio del bosque y después, como fácilmente se ve, calcular directamente con estos porcentos. Si por ejemplo, el error tipo de ambas mediciones en el primero y segundo inventarios, es igual a 0.3% del área basimétrica del bosque (algunas veces relacionado entre el promedio del primero y del segundo inventario), el error tipo de posición, tanto del primero como del segundo inventario

es de 0.2% y el error tipo del incremento debido a la imprecisión de las forcíparas usadas es igual a 0.4%, entonces el error total del incremento es igual a

$$\sqrt{(0.3^2 + 0.2^2) + (0.3^2 + 0.2^2) + 0.4^2} = 0.65 \%$$

Si el incremento periódico fué, digamos, del 20% de la mencionada área basimétrica promedio, entonces el error del incremento expresado en porciento del mismo, será igual a

$$0.65 \frac{100}{20} = 3.25 \%$$

o sea, mucho más grande que el porciento de error en área basimétrica del mismo bosque.