

Universidad Austral de Chile
Facultad de Ciencias Forestales

Apuntes de
MENSURA FORESTAL
I. Estática

Oscar García

Septiembre 1995

Instituto de Manejo
Cátedra de Mensura Forestal

Capítulo 1

Introducción

La Mensura Forestal, Dasometría o Dendrometría, trata de la cuantificación de bosques, árboles y productos forestales. Podemos distinguir en ella técnicas de medición directa o indirecta, procedimientos de estimación usando relaciones estadísticas, y métodos de predicción donde interviene la variable tiempo. Los tópicos principales incluyen:

	MEDICION (directa, indirecta)	ESTIMACION (estadística)	PREDICION (en el tiempo)
TROZAS (productos)	Longitud, diámetro Cubicación (volumen) Reglas madereras Defectos, calidad	Funciones de volumen Conv. madera aserrada Conv. peso-volumen Madera arrumada	
ARBOLES	DAP, altura, corteza Cubicación Análisis fustal	Funciones de volumen Fun. de ahusamiento Fun. de corteza Distribución de productos	
RODALES	Tablas de rodal	Altura-DAP Distribuciones de DAP Funciones de volumen Inventarios	Calidad de sitio Crecimiento Mortalidad

Cubriremos estos temas aproximadamente en el orden indicado, excepto por Inventarios que será tratado en un curso aparte. En el tiempo disponible solo será posible introducir los conceptos y métodos principales en forma bastante superficial. Se repasarán y/o explicarán algunos temas de matemáticas y estadística en la medida y oportunidades en que sea necesario. Enfatizaremos principios, fundamentos y métodos de aplicabilidad general, indicando fuentes de información a las que se pueda recurrir cuando se enfrenten problemas determinados.

Un objetivo importante es el desarrollar habilidades de razonamiento cuantitativo y de aplicación de conocimientos matemáticos y estadísticos a problemas reales y situaciones nuevas. Se espera que muchos detalles prácticos sobre métodos específicos y toma de mediciones sean aprendidos a través de la lectura de la bibliografía, ejercicios, y cursos posteriores. Las clases prácticas y los temas tratados en ellas serán una parte importante e integral del curso.

Tópicos menos importantes o algo periféricos al tema central serán indicados con letra más pequeña y con un ♡. Items marcados con ♡♡ se incluyen más bien a título de curiosidad.

Referencias generales:

INSTITUTO DE MANEJO FORESTAL, Cátedra de Manejo Forestal. Apuntes de años anteriores. (Mucho material relevante y complementario, en castellano).

HUSCH, B., MILLER, C.I., and BEERS, T.W. Forest Mensuration. 3rd Edition. Wiley. 1982. (Buen texto básico, con énfasis en los principios de medición).

AVERY, T.E. and BURKHART, H.E. Forest Measurements. Fourth edition. McGraw-Hill. 1994. (Otro texto básico, explicaciones breves y énfasis en prácticas usadas en Norteamérica).

AVERY, T.E. Natural Resources Measurements. Second edition. McGraw-Hill. 1975. (Edición anterior, disponible en biblioteca).

LOETSCH, F., ZOEHRER, F. and HALLER, K. Forest Inventory, Vol.II. BLV Verlagsgesellschaft. 1973. (Bastante completo, buen texto de consulta).

PARDE, J. Dendrométrie. Editions de l'Ecole Nationale des Eaux et Forêts – Nancy. 1961. (Texto de mensura en francés. Existe edición actualizada por Pardé y Bouchon, pero no en la biblioteca).

PRODAN, M. Holzmeßlehre. J.D. Sauërlander's Verlag, Frankfurt. 1965. (Clásico en alemán).

VAN LAAR, A. y AKQA, ? ?? (Por aparecer, pero por sus autores debería ser excelente).

SPURR, S.H. Forest Inventory. Ronald. 1952. (Un clásico, todavía interesante en su tratamiento de tablas de volumen, principios de crecimiento, etc.).

CAILLIEZ, F. Estimación del Volumen Forestal y Predicción del Rendimiento. Vol.1 - Estimación del Volumen. Estudio FAO: Montes 22/1. 1980. (Volúmenes, traducción castellana).

ALDER, D. Estimación del Volumen Forestal y Predicción del Rendimiento. Vol.2 - Predicción del Rendimiento. Estudio FAO: Montes 22/2. 1980. (Uno de los mejores en modelos de crecimiento. Traducción castellana).

BRUCE, D. and SCHUMACHER, F.X. Medición Forestal. Editorial Herrero. 1965. (Traducción de un texto bastante antiguo, aunque todavía útil en la parte de instrumentos y mediciones).

CLUTTER, J.L., FORTSON, J.C., PIENAAR, L.V., BRISTER, G.H. and BAILEY, R.L. Timber Management: A Quantitative Approach. Wiley. 1983. (Texto de manejo, también cubre bien los métodos más comunes en modelos de crecimiento).

ASSMAN, E. The Principles of Forest Yield Study. Pergamon Press. 1970. (Buena referencia, bastante sobre estudios de crecimiento en Europa. Sufre de falta de estructura y análisis).

CARRON,L.T. An Outline of Forest Mensuration with Special Reference to Australia. Australian National University Press. 1968. (Breve y sin mucho detalle, pero de interés para mensura de plantaciones).

PETERS N.,R., JOBET J.,J. y AGUIRRE A.,S. Compendio de Tablas Auxiliares Para el Manejo de Plantaciones de Pino Insigne. Instituto Forestal, Manual No.14. 1985. (Prácticas, ecuaciones, etc., usadas en Chile).

HAMILTON, G.J. Forest Mensuration Handbook. Forestry Commission Booklet No.39. Her Majesty's Stationery Office, London. 1975. (Manual interesante en la descripción detallada de procedimientos y normas. Discute grados de refinamiento en mediciones de acuerdo a sus costos y beneficios).

Capítulo 2

Trozas y productos forestales

2.1 Longitud

Longitudes se miden generalmente con huinchas o varas graduadas, y no presentan mayores dificultades. Se debe tener presente sin embargo que la definición exacta de una longitud está sujeta a convenciones que pueden variar de una situación a otra. Por ejemplo, en muchos casos se acostumbra a redondear el largo de una troza al pie o decímetro inferior. Dependiendo de la aplicación o convenciones establecidas, el largo de una troza curvada podría medirse en línea recta o siguiendo la curvatura.

♡ Aunque el medir longitudes puede parecer simple, hay algunos aspectos no del todo obvios. ¿Cuánto mide la costa de Chile? ¿Cómo medirla? Se podría medir en un mapa desplazando cuidadosamente una regla, caminando un compás sobre ella, con un curvómetro, o trazándola con un hilo. En cualquier caso, la medición sobre un mapa escolar a pequeña escala seguramente dará un valor menor que si se mide en las cartas de escala 1:50000. Se puede imaginar que si el proceso se repite sobre fotografías aéreas cada vez a mayor escala, se irán obteniendo longitudes mayores. ¿Cuál es la longitud “verdadera”? ¿Existe tal cosa?

PREGUNTAS, EJERCICIOS

1. Mida la longitud de la costa en un mapa, o del borde irregular de alguna hoja, con un compás abierto a 3 cm. Repita con aberturas cada vez menores.
2. Grafique el logaritmo de las longitudes obtenidas sobre el logaritmo de la resolución (abertura del compás). ¿Ve alguna tendencia?
3. Repita lo anterior con un semicírculo. ¿Semejanzas y diferencias?

4. El límite de la pendiente en la relación entre los logaritmos de longitud vs. resolución se llama la *dimensión fractal* de una curva. Si esta dimensión no es un entero la curva es una *fractal*. Un tema muy de moda es el uso de fractales (curvas y sus generalizaciones a superficies y volúmenes) como modelos para objetos y fenómenos naturales. Numerosos artículos sobre esto han estado apareciendo en las revistas de investigación forestal. Comente sobre las siguientes aseveraciones:
 - (a) Las fractales se distinguen en que al medir su longitud ésta aumenta con la resolución usada.
 - (b) Al contrario que en la geometría clásica, en la naturaleza todas las curvas y superficies son fractales.
 - (c) En consecuencia, hablar de la longitud o superficie de objetos reales no tiene sentido.
5. Un ejemplo que se da frecuentemente es el largo del río Misisipí. Al medirlo en segmentos de 400 Km (con un compás, por ejemplo) su longitud es de 1600 Km. Si se usan segmentos de 100 Km se obtiene un largo de 1800 Km. Así se obtienen largos cada vez mayores a medida que se aumenta la resolución. Comente.
6. A la salida de Valdivia hay un letrero que indica 187 Km a Temuco.
 - (a) ¿Cómo cree que se obtuvo ese valor? Sugiera métodos alternativos.
 - (b) ¿Encuentra alguna utilidad en esa información?

♡ Algunas referencias sobre fractales:

Scientific American, Vol.238, No.4, pág.14. Abril 1978.

STRAND, L. Crown density and fractal dimension. Medd.Nor.inst.skogforsk. 43(6):1-11. 1990.

ZEIDE, B. and PFEIFER, P. A method for estimation of fractal dimension of tree crowns. For.Sci. 37:1253-1265. 1991.

MANDELBROT, B.B. The Fractal Geometry of Nature. W.N.Freeman. 1983.

de GUZMAN, M., MARTIN, M.A., MORAN, M. y REYES, M. Estructuras Fractales y Aplicaciones. Una Introducción. Editorial Labor. 1991.

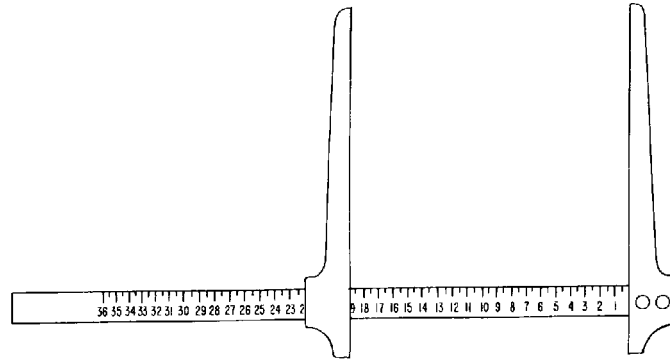
PRUSINKIEWICZ, P. and HANAN, J. Lindemeyer Systems, Fractals and Plants. Lecture Notes in Biomathematics, Vol.79. 1989.

En Física se han soslayado los problemas filosóficos sobre la naturaleza o existencia de valores “reales” tomando el punto de vista *operacional*. En éste se define una cantidad a través del procedimiento (operaciones) usado para medirla.

2.2 Diámetros

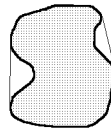
Los instrumentos comúnmente usados son la huincha (midiendo la circunferencia) y la forcípula. También se pueden medir diámetros directamente en los extremos de las trozas. No hablaremos más aquí de esta situación, sino que de

las mediciones “desde afuera”, como en puntos alejados de los extremos o en árboles en pie.



FORCÍPULA

Lo que generalmente interesa no es el diámetro, sino que el área de la sección transversal con el fin de estimar volúmenes. En primer lugar, está claro que tanto la huincha como la forcípula ignoran posibles concavidades en la sección, tratando con el cierre convexo (*convex hull*):



La diferencia (positiva) entre el área del cierre convexo y el área de la sección transversal del tronco se llama el **déficit convexo** o déficit de convexidad.

La huincha en realidad mide el perímetro P del cierre convexo. Lo que se llama “diámetro” es el diámetro que tendría un círculo con ese perímetro, es decir, $D = P/\pi$. De hecho, la *huincha de diámetro* usada en mensura está graduada en unidades de π , de modo que da ese diámetro directamente. Pero el área del cierre convexo será siempre menor o igual al área $\pi D^2/4$ de ese círculo, ya que el círculo es la figura con mayor área para un perímetro dado. La diferencia se llama el **déficit isoperimétrico**. Hemos identificado entonces dos fuentes de sobreestimación al calcular el área de la sección como $S = \pi D^2/4 = P^2/(4\pi)$.

En general, el diámetro medido con la forcípula varía con la dirección desde la cuál se mide. Por un teorema de Cauchy (1841), se puede probar que el valor esperado de un diámetro con forcípula medido en una dirección elegida al azar (o lo que es lo mismo, el diámetro promediado sobre todas las direcciones posibles) es igual al valor obtenido con la huincha (estamos ignorando posibles errores de medición). Es decir, $E[D_f] = D_h$.

♡ Más precisamente, lo que dijo Cauchy es que para cualquier figura convexa de perímetro P , el valor esperado de su proyección en una dirección aleatoria es P/π .

Claramente la medición D_f dada por la forcípula corresponde a una proyección, y el “diámetro” dado por la huincha es $D_h = P/\pi$.

Sin embargo, el valor esperado del área de la sección transversal obtenida de una medición de forcípula no coincide con la obtenida con huincha. Los diámetros por forcípula tienen una cierta varianza

$$\sigma^2 = V[D_f] = E[D_f^2] - E[D_f]^2 = E[D_f^2] - D_h^2 .$$

Las áreas calculadas para la sección son $S_f = \pi D_f^2/4$, con valor esperado

$$E[S_f] = \frac{\pi}{4} E[D_f^2] = \frac{\pi}{4} (\sigma^2 + D_h^2) ,$$

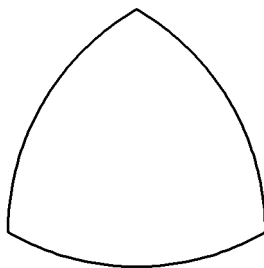
de donde

$$E[S_f] = S_h + \frac{\pi}{4} \sigma^2 .$$

Este valor es mayor o igual que S_h , ya que σ^2 es no-negativo.

Loetsch *et al.* en su texto, y B. Matérn (“On the geometry of the cross-section of a stem”, *Meddelanden från, Statens Skogsforskningsinstitut, Band46,Nr.11, 1956*) analizan los resultados de promediar pares de mediciones de forcípula tomados de varias maneras.

♡ ¿Podríamos determinar la forma de la sección transversal a través de varias mediciones con forcípula, y así idear algún tipo de corrección para el déficit isoperimétrico? La siguiente figura tiene la misma proyección en todas las direcciones, al igual que un círculo:



Por lo tanto, no se puede determinar la forma de una sección por mediciones “desde afuera”.

♡♡ La figura anterior se llama el triángulo de Reuleaux. Matérn da otros ejemplos de figuras con esta propiedad, llamadas *orbiformes*. Otro artículo interesante es: Gardner, M. “Curves of constant width, one of which makes it possible to drill square holes”. *Scientific American*, Febrero de 1963.

PREGUNTAS, EJERCICIOS

1. Ud. necesita cortar una vara de 1,3 m. Sólo tiene una huincha de diámetro graduada en centímetros y milímetros (de diámetro). ¿Qué lectura en la huincha corresponde al largo de la vara?
2. Uno de los últimos logros del programa de mejoramiento genético ha sido la producción de árboles de sección cuadrada, con los consiguientes ventajas en conversión de aserrado y costos de transporte. Para una sección de 30 cm por lado:
 - (a) Calcule el diámetro indicado por la huincha.
 - (b) Calcule el déficit isoperimétrico.
 - (c) Indique el valor esperado del diámetro con forcípula para una dirección tomada al azar.
3. Note que el triángulo de Reuleaux mostrado arriba está formado por tres segmentos de círculo, centrados en los vértices. Calcule como porcentajes del área real las áreas obtenidas de mediciones con huincha y con forcípula.
4. En una rodela, calcule lo más exactamente posible el área de la sección usando, por ejemplo, un planímetro o una red de puntos. Calcule, como porcentajes del área, el déficit de convexidad y el déficit isoperimétrico.
5. Las siguientes son algunas propiedades de la elipse con diámetro mayor $2a$ y menor $2b$:

$$\text{ecuación: } \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

$$\text{área: } \pi ab$$

$$\text{excentricidad: } e = \sqrt{1 - b^2/a^2}$$

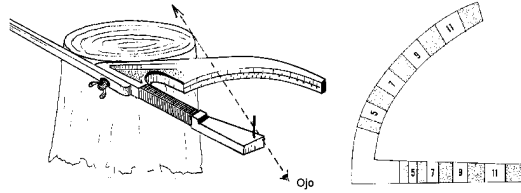
$$\text{perímetro: } \pi \left[\frac{3}{2}(a + b) - \sqrt{ab} \right] \text{ (aprox.)}$$

(esta es una aproximación al perímetro válida para excentricidades moderadas, suficiente para este ejemplo; el valor exacto es $a \int_0^{2\pi} \sqrt{1 - e^2 \sin^2 \theta} d\theta$, un integral elíptico de segundo tipo cuyos valores se pueden encontrar en tablas).

Tome algunos valores razonables de excentricidad para árboles con sección transversal elíptica (por ejemplo pensando en un diámetro mayor de 30 cm y varios diámetros menores razonables). Para estas excentricidades:

- (a) Calcule el déficit isoperimétrico, como porcentaje (demuestre que depende solo de la excentricidad).
- (b) Calcule el error porcentual máximo al calcular el área de la sección con una medición de forcípula.
- (c) Calcule el error porcentual en el área al usar el promedio de los diámetros menor y mayor.

6. Explique el principio de los instrumentos para medición de diámetros que aquí se muestran (el *Visiermeszwinkel* o sector de diámetro de Bitterlich y un tipo de forcípula parabólica):



2.3 Cubicación de trozas

Al calcular volúmenes de trozas y árboles normalmente se supone que las secciones son circulares, o al menos que se tiene diámetros tales que el área de la sección es $\pi D^2/4$. Este es un ejemplo simple del uso de *modelos*, idealizaciones que se toman como ciertas en cálculos y decisiones posteriores.

♡ “En general, un modelo es una representación simplificada de algún aspecto de la realidad (no confundir con la acepción normativa de la palabra, algo digno de ser imitado). Continuamente todos usamos modelos en alguna forma. Hay modelos mentales, que son relaciones imaginadas de causa y efecto entre componentes de algún sistema a través de las cuáles tratamos de explicar y anticipar su comportamiento. Se puede plantear modelos en forma verbal, por ejemplo la descripción en palabras del funcionamiento de alguna máquina. Los modelos físicos o materiales, como los modelos a escala de edificios y aviones, son bien conocidos.

Un modelo matemático es como un modelo verbal, pero expresado en lenguaje matemático. El lenguaje matemático difiere del lenguaje natural en que es más conciso y menos ambiguo. Esto, junto con la disponibilidad de reglas que se pueden usar mecánicamente, nos permite razonar en situaciones más complejas, con menos esfuerzo, y con menos riesgo de confundirse.

Con el progreso en computación se hace más fácil manejar modelos cada vez más complejos, y de hecho, el uso de modelos a escala en ingeniería a ido disminuyendo, reemplazados por modelos matemáticos que son más baratos y más flexibles. Los computadores han llegado a ser indispensables como herramientas para el desarrollo y uso de muchos modelos. Pero nótese la analogía entre *modelado por computador* y *poesía por máquina de escribir*. También, el realismo no es necesariamente una virtud en los modelos, es preferible abstraer sólo aquéllos aspectos más relevantes en cada caso. En un modelo de aviación para pruebas en el túnel de viento su color o el nombre del piloto pueden no ser importantes. El manual de un grabador de video no dice mucho acerca de sus mecanismos internos, pero se puede usar para predecir los efectos de apretar distintos botones.

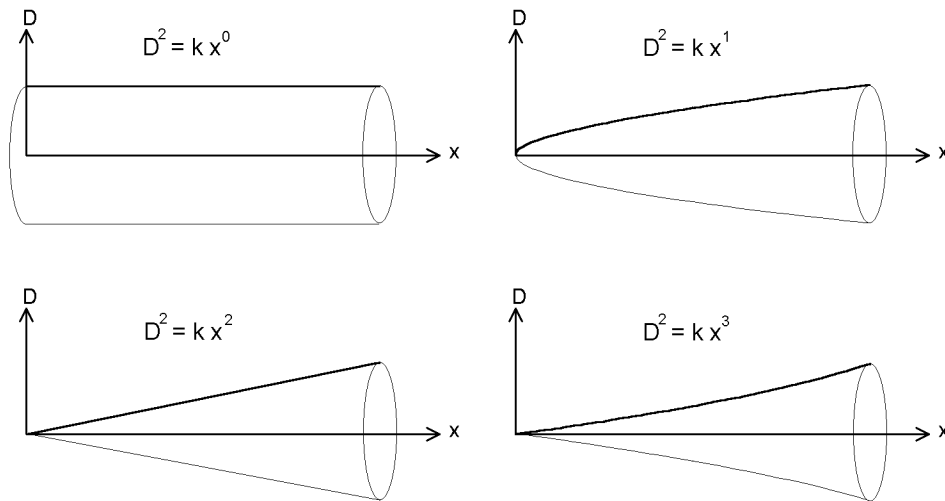
Es útil distinguir entre *modelos para predicción* y *modelos para comprensión* (Bunnell 1989). Los modelos para comprensión (por ejemplo modelos fisiológicos, de procesos) son útiles principalmente en investigación, como ayuda al entendimiento, para sintetizar y relacionar conocimientos anteriormente aislados, y para identificar vacíos donde se necesitan más estudios. Los beneficios surgen del desarrollo del modelo, y

no tanto de un uso posterior. Aquí me concentraré en modelos para predicción, destinados a la planificación del manejo. Aplicaciones típicas son las proyecciones con fines de planificación forestal, la comparación y evaluación de regímenes silviculturales (podas, raleos), y la actualización de bases de datos de rodales.

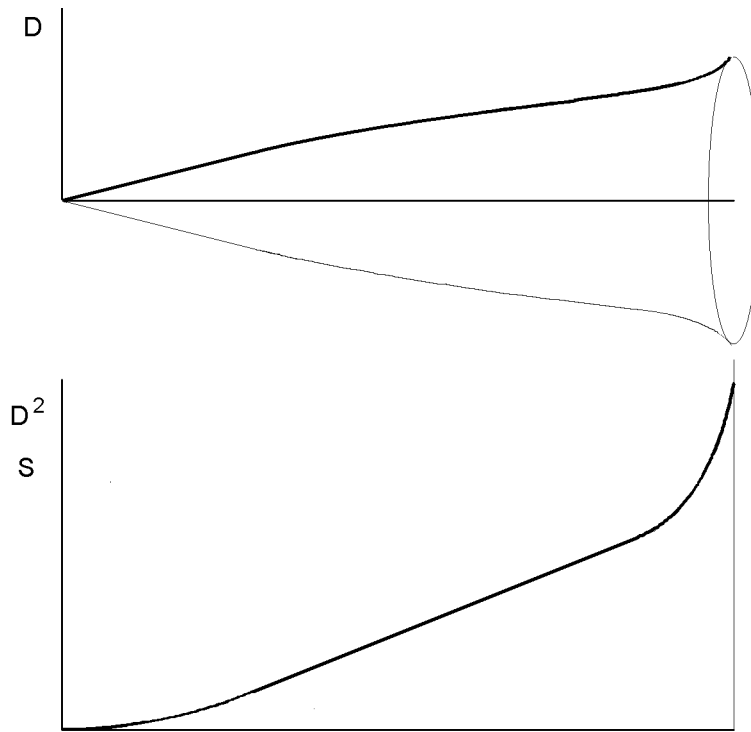
En la mayoría de los casos los modelos se presentarán como determinísticos. En general, al tomar decisiones los resultados de un modelo se toman como representativos de los sucesos más probables, y se ha encontrado que la introducción de elementos “aleatorios” en las predicciones no es muy útil en la práctica. Un componente estocástico (aleatorio, probabilístico) es sin embargo necesario para desarrollar procedimientos de estimación racionales, y estructuras estocásticas adecuadas se describen en este contexto más adelante.”

(García, O. “The state-space approach in growth modelling”. *Canadian Journal of Forest Research* 24, 1894–1903, 1994).

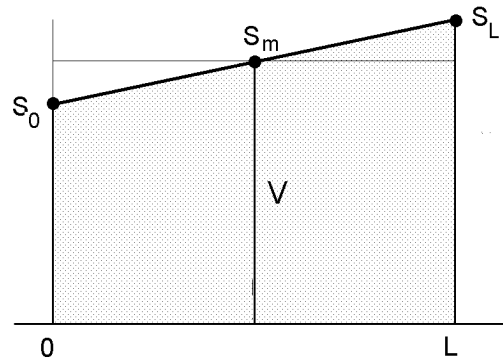
Es tradicional en mensura forestal el asemejar la forma de trozas y árboles a ciertos sólidos de revolución, el cilindro, paraboloide, cono, o neiloide. La fórmula general para la variación del diámetro D con el largo x es $D^2 = ax^n$ con $n = 0, 1, 2, 3$. En términos del área de la sección, $S = bx^n$.



Más en general, distintas partes del árbol se asemejan a porciones de estos sólidos. La parte de la copa, en coníferas, tiende a la forma de cono. La parte central del fuste se acerca al paraboloide. La base del árbol se expande en forma parecida al neiloide, aunque generalmente valores de n mayores que 3 se aproximan más.



El volumen de una sección de fuste, por ejemplo de una troza, corresponde claramente al área debajo de la curva de S sobre el largo. Para una troza en la parte central, paraboloidal, del árbol, la sección cambia de forma lineal, y el volumen es el área del siguiente trapecio.



El volumen (área del trapecio) puede en este caso calcularse dadas las secciones en los extremos y el largo:

$$V = \frac{S_0 + S_L}{2} L = \frac{\pi}{8} (D_0^2 + D_L^2) L .$$

En Mensura esta se conoce como la **fórmula de Smalian**. También se puede calcular en función del diámetro en el punto medio:

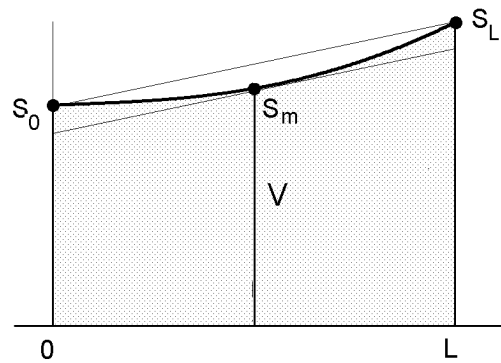
$$V = S_m L = \frac{\pi}{4} D_m^2 L ,$$

la **fórmula de Huber**. Se supone aquí que estamos usando unidades consistentes (por ej. D y L en metros, V en metros cúbicos).

PREGUNTAS, EJERCICIOS

1. De la fórmula de Smalian para D en centímetros, L en metros y V en metros cúbicos.
2. Lo mismo para D en pulgadas, L en pies y V en pies cúbicos.

Supóngase ahora que la curva de S es convexa, como es probable que ocurra en la primera troza y en el extremo superior del árbol. (Nota: la convexidad o concavidad de una curva o superficie se toma como vista desde el origen).



Se puede ver que Smalian da el área del trapecio superior, sobreestimando el volumen real. Huber da el área del trapecio inferior, produciendo una subestimación. Comparando las áreas entre las líneas de puntos a cada lado de la curva, se ve que Huber se acerca más al valor real.

La fórmula de Huber es generalmente más exacta, y requiere medir un diámetro en lugar de dos. En muchos casos, sin embargo, el centro de la troza no es fácilmente accesible, como cuando las trozas se encuentran apiladas. Además, si se necesita el volumen sin corteza es más fácil medir los diámetros bajo la corteza en las extremidades de la troza. Por esto la fórmula de Smalian, aunque produce errores mayores, tiende a usarse con mayor frecuencia.

Si se tuvieran los tres diámetros, en los extremos y en el centro, una media ponderada de Huber y Smalian reduciría los errores. Se demuestra que la siguiente fórmula, que se puede ver como una tal media ponderada, da resultados exactos para polinomios de hasta tercer grado. Es decir, es exacta para todos los sólidos de revolución aquí considerados.

$$V = \frac{S_0 + 4S_m + S_L}{6} L .$$

En Mensura ésta se conoce tradicionalmente como la **fórmula de Newton**, aunque en matemáticas esta es la base de la *regla de Simpson*. Es poco usada en la práctica, excepto en la cubicación de árboles completos como veremos más adelante.

♡ D. Bruce (“Butt log volume estimators”, *Forest Science* 28, 489–503, 1982) propone estimar el volumen de la primera troza con $\frac{\pi}{16}(3D_0^2 + D_L^2)L$ (explique por qué esto podría ser mejor que Smalian). También recomienda corregir Huber multiplicando por un factor de 1,04 o 1,08 según la especie, y analiza los errores de varios métodos en casos reales.

PREGUNTAS, EJERCICIOS

1. Una troza de 6 m de largo tiene diámetros menor, mayor y central de 22, 27 y 25 cm, respectivamente. Calcule el volumen por (a) Smalian, (b) Huber, y (c) Newton.
2. Una troza de 3,2 m de largo tiene un diámetro menor de 30 cm, mayor de 34 cm, y en su centro un diámetro de 32 cm.
 - (a) Use los diámetros menor y mayor para calcular el diámetro central:
 - i. Suponiendo que la troza es un tronco de cono.
 - ii. Suponiendo que la troza es un tronco de paraboloides.
 - (b) Calcule el volumen en metros cúbicos por las fórmulas de
 - i. Smalian.
 - ii. Huber.
3. Por geometría e integración se encuentra que el área bajo $S = bx^n$ entre dos puntos con $S = S_0$ y $S = S_L$, separados por un largo L , es

$$V = \frac{L}{n+1} \frac{S_L^{1+1/n} - S_0^{1+1/n}}{S_L^{1/n} - S_0^{1/n}}.$$

Obtenga las formas alternativas

$$V = \frac{LS_0}{n+1} \frac{r^{1+1/n} - 1}{r^{1/n} - 1} = \frac{LS_0}{n+1} \sum_{k=0}^n r^{k/n},$$

donde $r = S_L/S_0$.

4. Tome un rango de valores razonables para r (piense en árboles aproximadamente paraboloidales con varios diámetros y alturas). Calcule, para $n = 1, 2, 3, 8$, el porcentaje de error en las fórmulas de Smalian y Huber. (Sugerencia: note que $S^{1/n}$ en función del largo es una recta).

♡ **Integración numérica** El calcular el área bajo una curva $f(x)$ dados algunos puntos de ella es un problema importante en Análisis Numérico, la integración o cuadratura numérica. El enfoque usual es obtener fórmulas de la forma $L \sum w_i f(x_i)$ (el largo multiplicado por un promedio ponderado de las alturas) que sean exactas para polinomios del grado más alto posible, sujeto a diversas restricciones sobre el número de puntos y los valores de los x_i y/o los w_i . En el caso que se den los x_i uniformemente espaciados se tiene las fórmulas de *Newton-Cotes*, *cerradas* si los x_i incluyen los extremos del intervalo de integración, *abiertas* si no. Las fórmulas de Smalian y Newton usadas en mensura corresponden a las de Newton-Cotes cerradas para 2 y 3 puntos, respectivamente. Se conocen como la fórmula trapezoidal y la de Simpson, y son exactas para polinomios de 1er y 3er grado (el error es una función de la segunda y tercera derivadas). La de Huber se podría ver como la fórmula abierta con un punto, exacta para polinomios de 1er grado.

Si los x_i se pueden elegir libremente se obtienen las fórmulas de *Gauss*. Por ejemplo, con tres puntos $L[5f(L(1 - \sqrt{3/5})/2) + 8f(L/2) + 5f(L(1 + \sqrt{3/5})/2)]/18$ es exacto para polinomios de hasta quinto grado. En las fórmulas de *Chebyshev* se exige que todos los w_i sean iguales, lo que reduce el efecto de errores en $f(x)$, y facilita los cálculos ya que basta con tomar la media de los $f(x_i)$. Para dos puntos Gauss y Chebyshev coinciden, y para uno se tendría otra vez Huber.

Nótese que dependiendo de la irregularidad y errores en $f(x)$ puede convenir más integrar subintervalos por separado que usar fórmulas de grado alto. Con incertidumbre en $f(x)$ también puede ser mejor usar un polinomio que se aproxime a sus valores en lugar de uno de grado superior que los interpole exactamente. Para más detalles véanse textos de Análisis Numérico.

2.4 Madera arrumada

Productos como leña y rollizos para pulpa se comercializan frecuentemente de acuerdo al volumen en pilas o rumas. El **metro estéreo** es el volumen de una pila de $1 \times 1 \times 1$ metros (un metro cúbico apilado), y se usa para leña. Para pulpa la unidad más usada en Chile es el **metro ruma**, una sección de 1×1 metros en una ruma de trozos de 2,44 m de largo. Puede ser con corteza o sin corteza. Una unidad común en Norteamérica es la cuerda (*cord*), que corresponde a una ruma de 4×8 pies con trozos de 4 pies de largo.

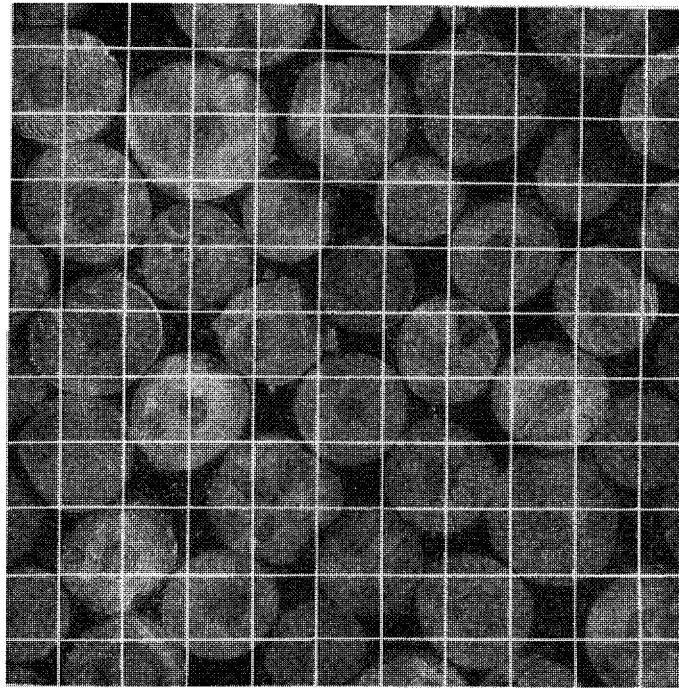
Debido a efectos de borde el contenido de madera puede variar ligeramente con las dimensiones de la ruma, y bastante con la forma de apilado, así que compradores y vendedores suelen establecer normas específicas sobre dimensiones y métodos de apilado.

Otros factores importantes en el contenido sólido son la irregularidad de los trozos, la variabilidad de los diámetros, y el espesor de corteza. El movimiento en el transporte también puede introducir cambios importantes.

Interesa disponer de factores de conversión, por ejemplo metros cúbicos sólidos por metro ruma. Estos pueden expresarse también como una proporción de volumen sólido en el volumen arrumado. Peters *et al* dan algunos factores de conversión para pino radiata en el Manual No. 14 del INFOR. Bruce y Schumacher citan una fórmula de éste último para la proporción de volumen sólido con

corteza: $0,84 - 0,04N$, donde N es el número medio de trozos por pie cuadrado en la cara de la ruma.

Es posible obtener factores de conversión más exactos para situaciones específicas a través de muestreo. Se pueden medir rumas y luego desarmarlas y determinar el volumen de los trozos usando la fórmula de Smalian, por ejemplo. Una alternativa mencionada a menudo consiste en tomar fotografías y medir en ellas la proporción de madera con una red de puntos u otro método. Se han propuesto redes similares para ser usadas directamente en el terreno. Más práctico para uso en terreno sería el método propuesto por Avery: marcar una vara o huincha a intervalos regulares y determinar la proporción de marcas que caen en madera al ponerla sobre la cara de la ruma. En cualquier caso es necesario hacer un muestreo lo suficientemente intenso para conseguir la precisión deseada.



♡ **Redes de puntos y muestreo sistemático** Es fácil obtener estimadores insesgados de precisión conocida usando muestreo aleatorio. El muestreo sistemático, del cuál la red de puntos es un caso especial, es generalmente mucho más eficiente. Su uso ha sido resistido, sin embargo, debido a la dificultad de estimar su precisión, y a la posibilidad de errores extremos en casos desfavorables. Estos errores extremos, asociados a coincidencias de periodicidad entre la población y la muestra, probablemente son raros en la práctica, y generalmente pueden prevenirse tomando precauciones adecuadas. La teoría detrás de su precisión es complicada, y los resultados exactos

dependen de características no observables de la población. No obstante, ha sido posible obtener aproximaciones adecuadas. Es así como en algunas disciplinas ha habido un resurgimiento del interés en procedimientos de muestreo sistemático, especialmente en la *Esterología*. Esta trata de la determinación de características cuantitativas principalmente en órganos y tejidos animales, y existe una teoría matemática muy desarrollada. Ver por ejemplo los artículos de Matérn, Kellerer, Mattfeldt y Cruz-Orive en el volumen 153, número 3 del *Journal of Microscopy*, 1989.

La varianza en una medición de área con red cuadrada de puntos se encuentra que es aproximadamente $0,0728Pa^3$, donde P es el perímetro del área a medir y a es el espaciamiento de la red. Esta fórmula fue obtenida por Matheron en 1965, y presentada en la literatura forestal por J. Bouchon (*Ann.Sci.For.* 32,131-134, 1975) y R.B. Chevrou (*Resource Inventory Notes* 20,3-6, 1979). El valor más exacto 0,0728 en lugar de 0,0724 lo da Matérn en la referencia ya citada, y se usa en la ecuación de más abajo. Gundersen y Jensen (*J.Microscopy* 147,229-263, 1987) usan esto para dar el coeficiente de variación del área A en función del número de puntos contados N y del indicador de forma P/\sqrt{A} :

$$\sigma/A = 0,27(P/\sqrt{A})^{1/2}/N^{3/4} .$$

El cociente P/\sqrt{A} tiene un mínimo de $2\sqrt{\pi}$ para círculos (nótese la relación con el déficit isoperimétrico). Gundersen y Jensen dan ejemplos con cocientes de hasta 33 para figuras muy complicadas, lo que da para el coeficiente de variación un rango aproximado de $0,5N^{3/4}$ a $1,5N^{3/4}$.

PREGUNTAS, EJERCICIOS

1. ¿Cuántos metros ruma son una *cuerda* (cord)?
2. Se tiene una pila de trozas pulpables, apiladas a lo largo de una pendiente fuerte. Su longitud, a lo largo del suelo, es de 12 m. La altura media medida verticalmente es 1,5 m, y medida perpendicularmente al suelo es 1,2 m. Las trozas son de 2,44 m de largo. ¿Cuál es el contenido en metros ruma?
3. Calcule la proporción sólida en una ruma de cilindros de diámetro uniforme apilados en forma rectangular. Lo mismo para la disposición más compacta (triangular).



4. En la fotografía anterior estime la precisión obtenida con la red que se muestra. Haga una tabla de precisión vs. número de puntos.

5. Loetsch, Zöhner y Haller, página 34, muestran coeficientes de variación obtenidos con redes de puntos. Compare con las predicciones de la fórmula de Gundersen y Jensen. ¿Qué opina del coeficiente de la regresión logarítmica en la página 35?
6. Estime la precisión del método sugerido por Avery.
7. Una alternativa a contar puntos a lo largo de una línea es medir el largo acumulado de las intersecciones. Idee un método práctico basado en este principio.

2.5 Medición por peso

Un método usual de medición de madera para pulpa, y en algunos casos para madera aserrable, es a través del peso. Se pesan camiones o carros de ferrocarril cargados y se resta el peso conocido del vehículo vacío. Para operaciones en menor escala y medición en terreno, en Nueva Zelandia se ha desarrollado un dispositivo que calcula el peso a partir de la presión en el sistema hidráulico de un cargador frontal.

Interesa conocer los equivalentes en volumen y/o en peso seco. Factores que afectan estas conversiones son las variaciones en contenido de humedad y en densidad de la madera. A su vez estos varían con el tiempo transcurrido desde la corta, tamaño de los árboles, estación, clima y localidad de origen. Variables como la cantidad de barro adherido a los trozos y la cantidad de combustible en los camiones también pueden ser importantes.

La fuente de variación más importante es generalmente el contenido de humedad. Debe mencionarse que en muchos casos el pagar por peso es ventajoso para el comprador al incentivar la entrega de madera fresca.

Tal vez lo más usual sea el uso de un factor de conversión promedio. En ocasiones se utilizan correcciones basadas en el contenido de humedad (estimado con instrumentos eléctricos). Si P es el peso verde, p es el peso seco, y V es el volumen, el factor de conversión a volumen deseado V/P está determinado por la *densidad básica* p/V , característica de la especie, y el *contenido de humedad* (%) $100(P - p)/p$. El volumen verde V no cambia apreciablemente mientras el contenido de humedad no baje del punto de saturación de las fibras, alrededor del 30%.

Se usan también regresiones que incluyen el tamaño medio de los trozos o el número de trozos por carga. Otro predictor que se ha usado es la fecha (mes), ya que las variaciones a través del año suelen ser importantes. Un análisis de los factores que afectan los factores de conversión, especialmente en pino radiata, se encuentra en: Ellis, J.C. "Weight/volume conversion factors for logs", New Zealand Logging Industry Research Association, *Technical Release 6(3)*, 1984.

PREGUNTAS, EJERCICIOS

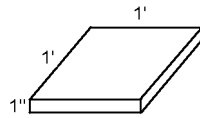
1. Dados el contenido de humedad h y la densidad básica d , obtenga fórmulas para los factores de conversión a volumen y a peso seco.

2. El precio de un metro ruma de *E. anonimus* descortezado es \$20000. Calcule un precio equivalente a pagar por la tonelada. Se sabe que: la proporción de madera sólida es 0,72; la densidad básica es 0,74 g/cm³; el contenido de humedad es 80%.

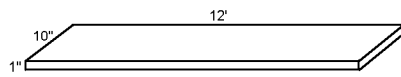
2.6 Madera aserrada

En Chile se acostumbra a expresar el volumen de madera aserrada en pulgadas madereras, una unidad relacionada al pie maderero (*board foot*) norteamericano. Aunque el sistema métrico gana terreno, en la práctica de aserraderos las medidas en pies y pulgadas todavía son de uso corriente. Una pulgada (abreviación 1") son 2,54 cm. Un pie (abreviación 1') son 12 pulgadas, o 30,48 cm.

El **pie maderero** es el volumen de una pieza cuadrada de 1 pie por lado y una pulgada de espesor. La **pulgada maderera** es el volumen de una tabla de 10" de ancho por 12' de largo y 1" de espesor.



Claramente, una pulgada maderera equivale a 10 pies madereros. Esta se usa comúnmente con maderas nativas. En plantaciones se usa más la **pulgada corta** o **pulgada pinera**, que se define con un largo de 10,5' en lugar de 12'. Estas son medidas *nominales*, es decir, pueden incluir tolerancias y/o pérdidas por contracción y cepillado.



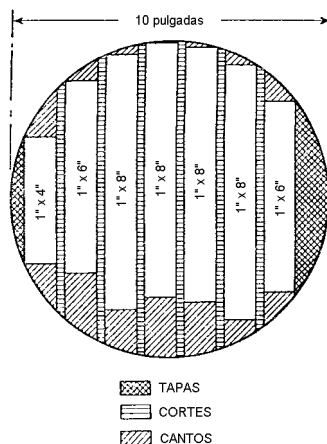
PREGUNTAS, EJERCICIOS

- Calcule el volumen en pulgadas madereras de piezas de las siguientes dimensiones:
 - $2'' \times 4'' \times 16'$
 - $3/4'' \times 3 \ 3/4'' \times 8'$
 - $4'' \times 4'' \times 10'$
- Calcule el factor de conversión entre pulgadas madereras y pulgadas cortas.
- ¿Cuántos metros cúbicos (nominales) son una pulgada maderera?

Para estimar la cantidad de madera aserrada a obtener de una troza se usan reglas madereras generales establecidas, o relaciones empíricas desarrolladas para un aserradero determinado.

Las **reglas madereras** son tablas o fórmulas que dan el volumen aserrado en función del diámetro menor y el largo de la troza. Se han obtenido a través de diagramas, o como fórmulas derivadas con razonamientos geométricos.

En el método de diagramas se dibujaron a escala círculos de varios tamaños representando el extremo menor de las trozas. En estos se marcaron las tablas que se obtendrían en el aserrado y se calculó su volumen.



La regla maderera de este tipo más conocida es la de **Scribner** (1846). Una ligera variante, la Scribner decimal C, es una de las más usadas actualmente en los E.E.U.U. Existe una fórmula de aproximación:

$$V = 0,79D^2 - 2D - 4$$

(trozas de 16' de largo, V en pies madereros, D en pulgadas truncadas al entero inferior).

Entre las reglas de fórmula, una de las más exactas es la **Internacional** (Clark, 1906). El razonamiento es como sigue (ver la figura más arriba). Primero, se considera que a cada tabla de 1" de espesor corresponde un corte donde se pierde $1/8''$ en aserrín y $1/16''$ de tolerancia por contracción. Es decir, se aprovecha 1" en $19/16''$, de modo que el área $\pi D^2/4$ del extremo menor se reduce multiplicándola por $16/19$. Luego se resta el área que se perdería en cantos y tapas. Se ve que ésta tiene una forma irregular pero se aproxima a un anillo exterior, con una profundidad independiente del diámetro. Alternativamente, se puede suponer que de cada tabla cortada se pierde en los bordes aproximadamente la misma cantidad. En todo caso, este término que se resta es proporcional a D , con una constante estimada experimentalmente. Para un largo de 4' se multiplican estos dos términos por el largo para obtener el volumen recuperable, que en en pies madereros y para diámetro en pulgadas es

$$V = 0,22D^2 - 0,71D .$$

Esta fórmula se puede modificar para anchos de corte distintos de $1/8''$, y la más usada es la de $1/4''$. Con este corte el área del círculo debería haberse reducido

por un factor de 16/21 en lugar de 16/19, así que simplemente se ajusta la fórmula obtenida multiplicándola por 19/21, obteniéndose

$$V = 0,199D^2 - 0,642D .$$

A diferencia de la regla de Scribner, la Internacional toma en cuenta la conicidad, suponiendo que tablas de 4' de largo son aprovechables y que la conicidad media es de 1/2" en 4'. La fórmula anterior se aplica entonces a partir del extremo menor a cada tramo de 4', incrementando el diámetro en 1/2" para el tramo siguiente. Así, para trozas de 16' la regla de 1/4" daría la fórmula siguiente:

$$\begin{aligned} V &= 0,199D^2 - 0,642D \\ &+ 0,199(D + 0,5)^2 - 0,642(D + 0,5) \\ &+ 0,199(D + 1,0)^2 - 0,642(D + 1,0) \\ &+ 0,199(D + 1,5)^2 - 0,642(D + 1,5) \\ &= 0,769D^2 - 1,374D - 1,23 . \end{aligned}$$

En su aplicación comercial estas reglas madereras van acompañadas por una serie de normas detalladas para la toma de mediciones, redondeo de cifras y deducciones por defectos. Especialmente en bosques nativos sobremaduros los defectos (pudrición, rajaduras, trozas torcidas) suelen ser importantes. La idea general es encerrar el defecto en un paralelepípedo pensando en la forma como se harán los cortes, y descontar su volumen antes de aplicar la regla. Las prácticas empleadas en E.E.U.U. son descritas por Bell, J.F. y Dilworth, J.R. "Log Scaling and Timber Cruising", O.S.U. Book Stores Inc., Corvallis, Oregon, 1988, 1993.

Las reglas madereras tradicionales solo pueden dar una estimación generalizada, ya que la conversión varía mucho con la tecnología del aserradero, especie, dimensiones de los productos, etc. Para una situación determinada es posible desarrollar una regla maderera empírica o función de conversión a través de estudios de aserrado. Se mide un cierto número de trozas cubriendo el rango de diámetros deseado, se aserrear, y se mide la madera obtenida. Con estos datos se puede entonces ajustar una ecuación de regresión dando el volumen aserrado en función del diámetro menor, y posiblemente del largo si éste varía. De las fórmulas para las reglas madereras ya vistas se deduce que una función razonable podría tomar la forma (para un largo dado)

$$V = b_0 + b_1D + b_2D^2 .$$

♡ F.X. Schumacher y W.C. Jones (*Journal of Forestry* 38, 889–896, 1940) propusieron un método interesante para obtener reglas madereras empíricas sin contar con datos detallados de trozas individuales. La idea básica es que la ecuación anterior puede sumarse sobre todas las trozas procesadas en un día:

$$\sum V = b_0N + b_1 \sum D + b_2 \sum D^2 .$$

Se puede entonces estimar los coeficientes a partir de cifras diarias de producción total $\sum V$, números de trozas N , y sumas de diámetros menores y de sus cuadrados. Fórmulas para trozas de largo variable se manejan de una manera análoga.

El método puede ser útil también para factores de conversión peso/volumen y en otras aplicaciones. Claramente, para obtener resultados confiables se necesitan series largas de producción diaria con variaciones importantes en las características y cantidad de trozas de un día a otro.

En muchos casos puede ser más conveniente expresar el rendimiento aserrable en términos relativos, como un factor de conversión de pulgadas o metros cúbicos aserrados por metro cúbico en rollizos. El volumen cúbico se puede estimar a partir del diámetro menor asumiendo algún valor para la conicidad, por ejemplo, la 1/2" por 4' de largo (1:96) de la regla Internacional.

PREGUNTAS, EJERCICIOS

1. Una troza de 3,2 m de largo tiene diámetros menor y mayor (sin corteza) de 30 y 34 cm, respectivamente.
 - (a) Obtenga el volumen aserrado en pulgadas madereras usando la regla Internacional de 1/8".
 - (b) Basado en Smalian, de el factor de conversión (% de recuperación).
2. L.R. Grosenbaugh (U.S.Forest Service, Southern Exp.Sta. Ocasional Paper 126, 1952) dió la siguiente aproximación simple a la regla Internacional de 1/4" para trozas de 16': $V = 0,8(D - 1)^2$. Evalúe su exactitud.
3. Expresé la regla Internacional de 1/4" para volumen aserrado en metros cúbicos, diámetro en centímetros y largo en metros.
4. Bajo el supuesto de conicidad de la regla internacional (1/2" en 4', o 1:96), obtenga fórmulas para el volumen cúbico en función del diámetro menor usando: a) Smalian, b) Huber, c) volumen de un tronco de cono.
5. Obtenga una fórmula para el factor de conversión a madera aserrada, como % en función del diámetro menor, usando la regla Internacional de 1/4" y su supuesto de conicidad. Sugerencia: considere una sección de 4'.
6. Suponga un largo de troza estándar de 12'. Se tienen valores diarios de número de trozas procesadas, producción en pies madereros de acuerdo a la regla Internacional, y volumen total en rollizos (calculado con el diámetro menor, Smalian y conicidad de 1%). ¿Puede obtener una regla maderera empírica por el método de Schumacher y Jones? ¿Cómo calcular $\sum D$ y $\sum D^2$?
7. Avery (citado por Loestch et al) escribió: "Hay poco para defender las reglas madereras aparte de su uso en el pasado y la resistencia a los cambios...". Al convertirse al sistema métrico Nueva Zelanda abandonó el uso de todo tipo de reglas madereras, presumiblemente por ser consideradas incompatibles con el nuevo sistema. Comente.

2.7 Cambio de unidades

Se tiene una fórmula, y se quiere expresarla en unidades distintas. Por ejemplo, la fórmula de Smalian con largo y diámetros en metros, $\frac{\pi}{8}(d^2 + D^2)L$, queremos convertirla a una para largos en metros y diámetros en centímetros. Para esto, simplemente se reemplazan las variables por el valor correspondiente en las nuevas unidades. Por ejemplo, D metros = D centímetros/100. Tenemos

$$\frac{\pi}{8}([d/100]^2 + [D/100]^2)L .$$

Simplificando se obtiene

$$\frac{\pi}{80000}(d^2 + D^2)L .$$

Como otro ejemplo, la fórmula Scribner

$$V = (0,79D^2 - 2D - 4)(L/16)$$

con V en pies madereros, D en pulgadas y L en pies se puede convertir a pulgadas madereras, centímetros y metros de la siguiente manera:

$$[10V] = (0,79[D/2,54]^2 - 2[D/2,54] - 4)([L/0,3048]/16)$$

$$V = (0,0025D^2 - 0,016D - 0,083)L .$$

Capítulo 3

Arboles

3.1 Diámetros

El diámetro más usado es el *diámetro a la altura del pecho* (DAP). Se define como el diámetro, con corteza a menos que se especifique lo contrario, a una altura sobre el suelo que en la mayoría de los países que usan el sistema métrico es de 1,3 metros. Algunas excepciones son Nueva Zelanda (1,4 m) y Japón (1,25 m). En Norteamérica se usa 4,5'. Estas alturas son cómodas para la medición con forcípula y están algo alejadas de la influencia del ensanchamiento que se produce en la base del árbol (aunque probablemente una altura mayor sería preferible).

Al medir el DAP es deseable atenerse a normas más precisas, las que lamentablemente no están estandarizadas. Por ejemplo, en terreno con pendiente la altura del DAP se acostumbra a medir ya sea sobre el nivel medio del suelo en la base del árbol, o sobre el nivel en el lado de arriba de la pendiente. En caso de haber deformaciones del fuste a la altura del pecho, la medición puede desplazarse hacia arriba, hacia abajo, o se puede tomar el promedio de dos mediciones.

PREGUNTAS, EJERCICIOS

1. Discuta las ventajas y desventajas de los varios criterios de medición del DAP recién mencionados.
2. ¿Cuán importante puede ser la diferencia entre medir el DAP a 1,3 o a 1,4 m? Una conicidad típica del fuste en la primera troza es del orden de 1:100.

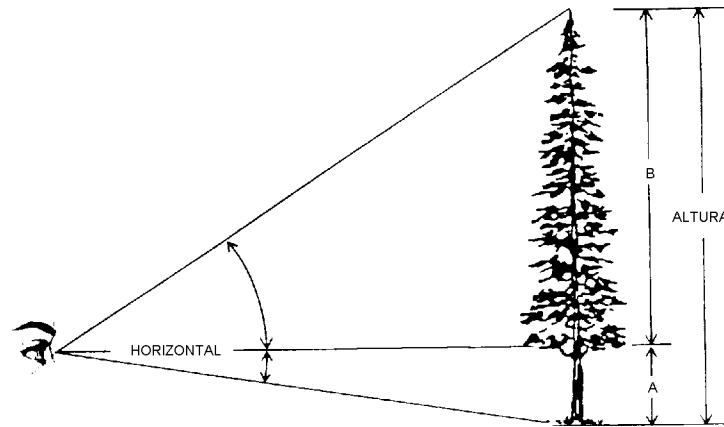
Los instrumentos más usados para medir el DAP son la forcípula y la huincha de diámetro (graduada en unidades de π). Diámetros a alturas superiores se miden trepando el árbol, con instrumentos montados en varas, o con diversos

tipos de dendrómetros ópticos. En todos los casos son aplicables las consideraciones sobre fuentes de variación y errores y las relaciones con el área de la sección transversal ya discutidas en la sección sobre medición de trozas.

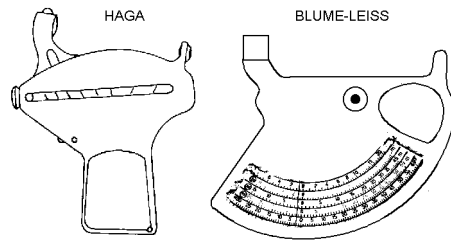
Para convertir diámetros con corteza a sin corteza, el espesor de corteza puede medirse o puede estimarse con relaciones pre-establecidas. El calibrador de corteza tipo sueco es el más empleado. Su uso adecuado requiere práctica y controles periódicos, y los sesgos pueden ser importantes. Generalmente se hacen dos lecturas, en lados opuestos del fuste, las que se suman para obtener el “doble espesor de corteza” que se toma como la diferencia entre los diámetros.

3.2 Alturas

Alturas de hasta 10–15 m se miden de preferencia con varas telescópicas. Para alturas mayores se usan clinómetros (instrumentos que miden ángulos verticales) o hipsómetros (instrumentos especializados que indican altura). Algunos hipsómetros (Christen, Merrit) emplean semejanza de triángulos, pero los más usados en la actualidad se basan en principios trigonométricos.



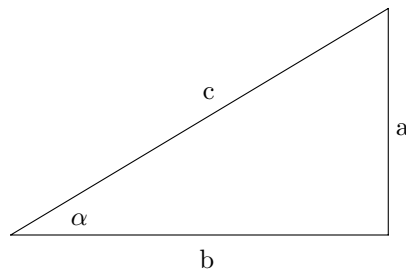
Alturas pueden ser totales (hasta el ápice) o comerciales (hasta un diámetro límite o hasta el punto donde el fuste se ramifica). Se mide la distancia al árbol, y ángulos entre la horizontal, la base, y el extremo superior. La distancia se mide con huincha o con un telémetro de distancia fija incorporado al instrumento. Generalmente la distancia es inclinada, lo que debe tomarse en cuenta en los cálculos. Los hipsómetros más recomendables son el Suunto, el Haga y el Blume-Leiss, todos basados en el uso de un péndulo o contrapeso para establecer una vertical de referencia. Clinómetros como el nivel Abney, que usa una burbuja como referencia, también pueden dar buenos resultados aunque su uso es más engorroso. Para precisiones altas es necesario recurrir a teodolitos o taquímetros.



Los hipsómetros están sujetos a errores altos sino se usan con cuidado. La vara es preferible cuando sea práctico usarla. Siempre deben de tomarse varias lecturas; la mediana de tres sería recomendable. Se debe también verificar periódicamente la calibración del instrumento. En especial la exactitud del telémetro debe comprobarse, ya que se han encontrado variaciones de fábrica importantes.

PREGUNTAS, EJERCICIOS

- Basándose en la figura más arriba, deduzca los principios y fórmulas para el cálculo de alturas. El enfoque general en problemas de este tipo consiste en establecer o identificar triángulos rectángulos y aplicar algunas de las fórmulas siguientes.



$$\frac{a}{c} = \sin \alpha$$

$$\frac{b}{c} = \cos \alpha$$

$$\frac{a}{b} = \tan \alpha$$

$$a^2 + b^2 = c^2$$

- ¿Qué sucede si el nivel de la visual está por debajo de la base del árbol?
- Ud. se encuentra a una distancia horizontal de 20 m de un árbol. Las siguientes lecturas han sido tomadas con clinómetro. Lectura al ápice: 48° . Lectura a la base: -8° .
 - ¿Cuál es la altura del árbol?
 - Los ángulos tienen un error de $\pm 1^\circ$. Por substitución, calcule límites de error para la altura.

3.3 Cubicación de árboles

Se distinguen varios tipos de volumen en un árbol. El volumen *cúbico* es el volumen de madera contenido en una porción del fuste. Puede ser volumen *total*,

para todo el árbol desde su base hasta el ápice, o *comercial*, desde la altura del tocón hasta la altura donde se alcanza un *diámetro límite* determinado. A veces se habla de volumen total refiriéndose a todo el volumen comercial utilizable, por ejemplo en las tablas de volumen para pino del INFOR que dan volúmenes desde el tocón hasta un diámetro límite de 10 cm (actualmente se aceptan diámetros menores, de 8 cm o menos). Se pueden considerar volúmenes cúbicos entre varios diámetros límites, por ejemplo volumen *aserrable* entre el tocón y un diámetro límite de 25 cm, y volumen *pulpable* entre diámetros límites de 25 y 10 cm. A menudo por volumen aserrable se entiende no el volumen cúbico, sino el volumen de madera aserrada estimado en pulgadas o pies madereros (volumen *aserrado*). Igualmente, el volumen pulpable podría especificarse en metros ruma. Por último, los volúmenes pueden ser *con corteza* o *sin corteza*.

Como ya se vio en el caso de trozas, el volumen cúbico de todo el árbol o de una porción de éste está dado por el área bajo la curva de sección transversal en función de la longitud o altura. Las secciones se calculan normalmente como áreas circulares para una serie de mediciones de diámetro a distintas alturas. Los diámetros se miden en el árbol volteado, trepando, o a la distancia con instrumentos ópticos (dendrómetros).

El cálculo del volumen cúbico es entonces esencialmente un problema de integración. Hay que calcular el área bajo la relación sección–altura, conociendo las secciones a varias alturas. Algunas posibilidades son las siguientes:

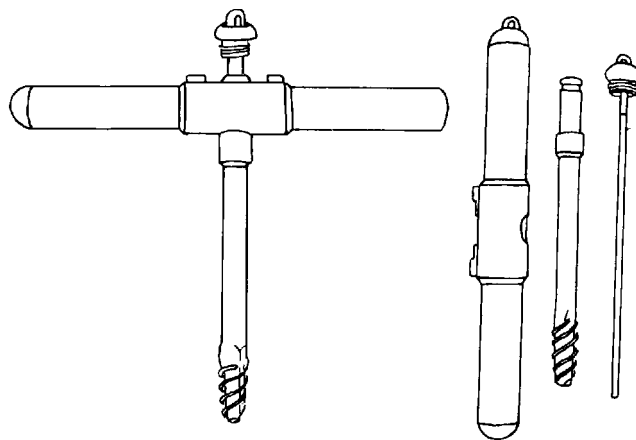
- Método gráfico. Graficar en papel milimetrado para cada altura de medición la sección transversal o el diámetro al cuadrado. Dibujar una curva que pase por estos datos. Calcular el área bajo la curva con un planímetro, contando puntos, o de otro modo. Aplicar el factor de proporcionalidad apropiado para obtener el volumen. Actualmente este método es poco usado.
- Por secciones. Para pares de mediciones consecutivas se calcula el volumen entre esas alturas con la fórmula de Smalian o con la fórmula para el cono. Se suman estos volúmenes para la porción deseada del fuste. Para intervalos de altura uniformes se puede usar también Huber o Newton. Si se incluyen, generalmente se toma el tocón como un cilindro y el extremo apical como un cono. Tal vez el procedimiento más usado. Si las secciones son suficientemente cortas, los errores y las diferencias entre las varias fórmulas son despreciables.
- Métodos más elaborados de integración numérica. Para intervalos uniformes se puede usar el método de Simpson u otro similar. Hay métodos para intervalos irregulares que podrían usarse, tal como el de Gill y Miller (*The Computer Journal* 15,80–83, 1972). J. Bouchon (“Les Tarifs de Cubage”, Ecole Nationale du Génie Rural des Eaux et Forêts, 1974), y más tarde otros, han usado integración con *splines*. La integración de curvas de ahusamiento (ver más adelante) podría considerarse como otro ejemplo de esto.

- Grosenbaugh propuso un método basado en mediciones espaciadas a intervalos constantes de diámetro. Está dirigido al uso de dendrómetros ópticos. El principio es similar a intercambiar variables e integrar bajo la curva de altura versus sección, en lugar de bajo la de sección versus altura.
- Muestreo. Ha habido interés recientemente en métodos basados en una selección cuidadosa de uno o más puntos de medición. El objetivo es el estimar en inventarios volúmenes totales por hectárea sin el uso de tablas de volumen. Se emplean estrategias de muestreo sistemático, restringido, y/o con probabilidad variable. Ver, por ejemplo, H.T. Valentine *et al.*, *Forest Science* 38,160–172, 1992, y H.V. Wiant *et al.*, *Forest Science* 38,187–191, 1992.

En trabajos de investigación, se han medido volúmenes por desplazamiento de agua en dispositivos llamados *xilómetros*. Se trata básicamente de tanques con agua en los que se sumergen las trozas. El volumen se determina por el rebalse o cambio de nivel del agua, o a través del cambio en el peso de la troza al sumergirla.

3.4 Análisis fustal

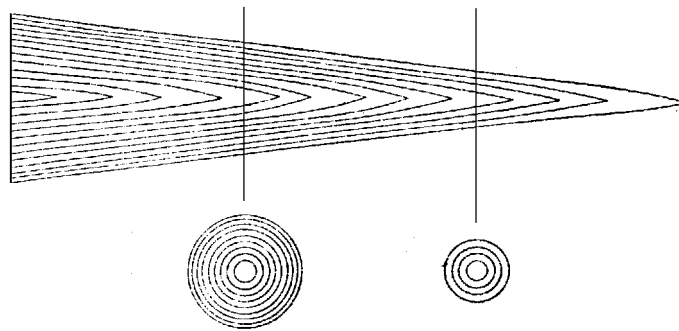
En muchas especies de árboles la madera producida al comienzo de la temporada de crecimiento difiere en su anatomía de la producida al final. En una sección transversal del fuste se pueden distinguir entonces *anillos de crecimiento* o anillos anuales, con una discontinuidad visible marcando el límite entre el crecimiento de años sucesivos. Los límites entre los anillos se pueden observar en secciones transversales, o en tarugos cilíndricos extraídos perpendicularmente a la superficie del fuste con un *taladro de incremento*:



La visibilidad de los anillos puede mejorarse con colorantes y otros tratamientos. Se usan también densitómetros de rayos X o de rayos gama para distinguir los cambios de densidad asociados con los anillos. Hay una serie de posibles

fuentes de error que hay que tener presente. La presencia de *anillos falsos*, producidos por variaciones climáticas bruscas u otros factores, causa dificultades. Excentricidad de los anillos y la inclinación del taladro pueden producir errores serios. Puede haber compresión de los anillos exteriores del tarugo, especialmente si el taladro no está bien afilado. La pérdida de humedad por debajo del punto de saturación de las fibras (aproximadamente 30%) produce contracción de la madera. Un buen manual sobre tarugos de incremento es: Jozsa, L. "Increment core sampling techniques for high quality cores", Forintek, Spec. Pub. No. SP-30, 1988.

El número de anillos anuales entre la médula y el cambio indica el año en que el árbol alcanzó la altura correspondiente. Esto puede usarse para estimar la edad del árbol. También se puede estimar el tiempo transcurrido entre dos alturas, y de ahí la tasa de crecimiento en altura. Medición de los anillos proporciona estimaciones del crecimiento en diámetro y en área basal. Para valores con corteza es necesario estimar indirectamente el espesor de corteza, generalmente con relaciones corteza-diámetro.



Tri-dimensionalmente lo que se tiene son capas de madera que se forman anualmente unas sobre otras. Observando las intersecciones de estas capas con cortes transversales del fuste a varias alturas (los anillos), se pueden reconstruir las dimensiones del fuste en el pasado. Así se pueden obtener datos para tablas de volumen, curvas de ahusamiento, e índices de sitio (crecimiento en altura). Los principios de esta reconstrucción son más o menos obvios, excepto posiblemente por las estimaciones de altura.

El problema con la altura es que el término de cada incremento en altura puede ocurrir a cualquier altura entre los niveles de dos cortes sucesivos, desconociéndose la altura exacta. A veces es posible guiarse por indicadores externos (verticilos, cicatrices dejadas por las brácteas de la yema apical, disposición de los primordios foliares), y hacer cortes que coincidan con el término de cada incremento anual. En caso contrario es necesario hacer algún tipo de interpolación. Dyer y Bailey (*Forestry Science* 33,3-13, 1987) encontraron que el simple método propuesto por Carmean da buenos resultados.

El método de Carmean se basa en suponer un incremento constante en altura para cada año comprendido entre dos niveles sucesivos, con los cortes ocurriendo

en el centro de un incremento. Se divide la distancia entre dos cortes sucesivos por la diferencia en número de anillos, obteniendo un incremento medio k . Las alturas estimadas sobre el nivel del corte inferior son entonces $k/2, k/2+k, k/2+2k, \dots$ (ver la figura).

PREGUNTAS, EJERCICIOS

1. En un análisis fustal un árbol con 67 anillos en el tocón se corta en trozas de 5 m de largo. Los números de anillos en los extremos superiores de las trozas son 53, 37 y 19. Considere los períodos de tiempo en los que el árbol pasó del nivel de un corte al siguiente. Para cada uno de estos períodos estime la tasa media de crecimiento en altura.
2. En un análisis fustal se acaban de obtener diámetros para los anillos números 5 y 10, contados desde el exterior de cada sección. Con estos diámetros se calcularon las siguientes áreas (m^2):

Altura (m)	Anillo 5	Anillo 10
0,3	0,1080	0,0740
3,3	0,0829	0,0581
6,3	0,0558	0,0412
9,3	0,0384	0,0283
12,3	0,0137	0,0093
14,1	–	0
15,0	0	–

Calcule la tasa de crecimiento en $\text{m}^3/\text{año}$ entre 1984 y 1989. Use Smalian tomando el tocón (altura 30 cm) como un cilindro.

3.5 Funciones (tablas) de volumen

Las mediciones necesarias para cubicar un árbol y calcular su volumen son costosas y lentas. Es de interés entonces el poder estimar el volumen indirectamente a través de variables más fáciles de medir, como el DAP y la altura. Las relaciones que permiten lograr esto son las funciones de volumen por árbol, llamadas también tablas por razones históricas.

Las funciones de volumen se obtienen por regresión, usando una muestra de árboles en los que se mide el volumen y las variables predictoras. El volumen puede ser total, comercial, aserrado, etc. Una vez teniendo la función, el volumen de otros árboles se puede estimar conociendo solo el valor de los predictores.

Las funciones de volumen cuyo único predictor es el DAP se llaman tablas *locales*. Solo son aplicables al rodal para el que se construyeron, o a lo más para rodales muy similares, ya que el volumen depende también en forma importante de la altura, y la relación entre ésta y el DAP varía con la densidad del rodal,

con su edad, etc. Típicas funciones de volumen local pueden tomar las formas $V = a + bD^2$, $V = aD^b$, o $V = aD + bD^2$. Los parámetros a y b pueden estimarse por regresión lineal simple, usando en los dos últimos casos las transformaciones $\log V = \log a + b \log D$ y $V/D = a + bD$.

Las tablas de volumen de uso general incluyen como predictores, además del DAP D , la altura (total o comercial) H , y en algunos casos también algún diámetro superior o indicador de forma. Un indicador de forma bastante usado es el *cuociente de forma de Girard*, definido como la razón entre el diámetro sin corteza a 5,19 m de altura y el DAP con corteza. Los 5,19 m corresponden al extremo de una primera troza de 16 pies.

Algunas formas usuales para funciones de volumen son

$$V = a + bD^2 + cH + dD^2H ,$$

y las variantes que se obtienen con varias combinaciones de a, b y c iguales a cero, y

$$\log V = a + b \log D + c \log H .$$

Se trata de típicos problemas de regresión lineal, sin mayores complicaciones. Sin embargo, tres aspectos particulares pueden mencionarse.

Ocurre a menudo que la dispersión de los residuos de la regresión para V tiende a aumentar con los valores de la variable, un caso de heterocedasticidad. La transformación logarítmica, cuando se usa, puede eliminar o reducir este efecto, ya que si σ es proporcional a V entonces los desvíos de $\log V$ tienen una varianza aproximadamente constante.

Otra forma de enfrentar la heterocedasticidad es el uso de *mínimos cuadrados ponderados*, aplicados a las tablas de volumen por Cunia en 1964. El supuesto es que la varianza de ε_i es $\sigma^2 w_i$, donde los w_i son conocidos. Esto es un caso especial de los mínimos cuadrados generalizados, con W diagonal (solo sus elementos diagonales $w_{ii} \equiv w_i$ son distintos de cero). Los parámetros se estiman entonces minimizando la suma ponderada de cuadrados $\sum e_i^2/w_i$. Se puede usar también un programa de regresión lineal ordinaria, notando que el modelo $y_i = \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i$ con varianza $\sigma^2 w_i$ se reduce a uno con varianza σ^2 si dividimos ambos lados por $\sqrt{w_i}$.

Un caso típico es la ecuación de volumen $V = a + bD^2H$. Se encuentra frecuentemente que los residuos sugieren un σ proporcional a la variable independiente D^2H . Se consiguen entonces mejores estimaciones de los parámetros a y b ajustando la ecuación $\frac{V}{D^2H} = a \frac{1}{D^2H} + b$.

Otro tema que se comenta a menudo es el hecho de que al usar transformaciones logarítmicas de la variable dependiente, tales como $\log V$, se obtienen estimadores sesgados para la variable original, en este caso V . Se han propuesto correcciones a los resultados de la regresión lineal para reducir el sesgo. No está claro, sin embargo, si esto realmente se justifica, ya que generalmente el sesgo se reduce a costa de aumentar el ECM y reducir la verosimilitud.

También es posible evitar el sesgo usando regresión no lineal, por ejemplo ajustando por mínimos cuadrados la ecuación $V = kD^b H^c$ en lugar de la ecuación logarítmica mostrada más arriba. Esto se ha puesto bastante de moda con

los progresos en computación. Considerando la ya mencionada estabilización de la varianza que generalmente se consigue con los logaritmos, es probable que en la mayoría de los casos este remedio sea peor que la enfermedad.

El tercer tema es la comparación de modelos que usan distintas transformaciones de la variable dependiente. Obviamente en ese caso no tiene sentido comparar los ES, ECM o R^2 de la regresión, ya que estos se refieren a variables distintas. Probablemente lo mejor es el análisis gráfico de los residuos, ya que la calidad del ajuste puede ser distinta para diferentes valores de las variables. Otra posibilidad es calcular el ES o ECM para las variables sin transformar, o con una misma transformación, preferiblemente en forma separada para varios rangos de los predictores. Furnival (*Forest Science* 7,337–341, 1961) propuso un índice que se usa frecuentemente con este fin. Es esencialmente una transformación aproximada de la función de verosimilitud. Debe tenerse presente entonces que mide tanto la plausibilidad de la función de regresión como la de la distribución de los desvíos implícita en ella.

3.6 Cuocientes y factores de forma, etc.

Especialmente en los países centro-europeos, se utilizan diversas combinaciones de diámetros para describir la forma y estimar volúmenes de árboles. Ya se mencionó el cuociente de forma de Girard, muy usado en Norteamérica.

Los *factores de forma* son razones entre el volumen de un árbol y el volumen de cierto cilindro. Se habla del factor de forma *artificial*, basado en el volumen sin corteza y el volumen de un cilindro de diámetro igual al DAP (con corteza), y de factores *naturales* basados por ejemplo en el cilindro de diámetro igual a aquel en la mitad del árbol o en alguna otra proporción de la altura total. El de Hohenadl, usado en Alemania, se basa en el diámetro a 1/10 de la altura.

Así, el factor de forma artificial es

$$f = \frac{V}{\frac{\pi}{4}D^2H},$$

usando unidades consistentes. Se usan regresiones para estimar f en función de predictores como D y H . Esto no difiere substancialmente de las funciones de volumen ya discutidas, especialmente al usar transformaciones como la de Cunia. Por lo tanto no lo estudiaremos en detalle.

Una aplicación útil es para estimaciones simplificadas y rápidas de volumen, suponiendo un factor de forma constante. Claramente esto equivale a usar una función de volumen de la forma $V = bD^2H$.

PREGUNTAS, EJERCICIOS

1. ¿Qué diferencia porcentual en el volumen calculado se produciría al medir el DAP a 1,4 en lugar de 1,3 m sobre el suelo? Suponga un factor de forma constante, y la misma conicidad usada en la Regla Internacional (0,5" en 4', o sea, 1:96). Calcule para DAPs de 20 y 40 cm.

2. En un árbol se ha medido un DAP (con corteza) de 25 cm, altura de 20 m, y un doble espesor de corteza al DAP de 30 mm. Se supone que la forma del fuste sobre el DAP se encuentra entre un cono y un paraboloides. Bajo el DAP aproxime con un cilindro.
 - (a) Calcule un mínimo y un máximo para el cociente de forma de Girard.
 - (b) Calcule el volumen cúbico total sin corteza en ambos casos.
3. Por integración, pruebe que el volumen de los sólidos de revolución clásicos con ecuación $d^2 = kh^n$ también está dado por $V = bD^2H$, en función de la altura H y del diámetro en la base D .
4. Verifique los valores de b iguales a 0,20 para el neiloide, 0,26 para el cono, 0,39 para el paraboloides, y 0,79 para el cilindro.

Vimos que se pueden especificar volúmenes hasta y entre diversos diámetros límites. Es posible desarrollar funciones de volumen con diámetros límites variables, incluyendo el diámetro límite como otra variable independiente en la regresión. Las funciones de ahusamiento son otro camino para estimar estos volúmenes.

3.7 Funciones (curvas) de ahusamiento

Las curvas o funciones de ahusamiento o conicidad (taper curves) describen los diámetros esperados, con o sin corteza, a distintas alturas en el fuste. Las funciones contienen también el DAP y altura total, y ocasionalmente otras variables. Entre otras cosas, con esta información se pueden calcular volúmenes cúbicos, aserrados, etc., para cualquier porción del fuste, por ejemplo entre dos diámetros límites dados. Incluso, dadas las especificaciones (diámetros y largos mínimos y máximos, posición en el árbol) para productos tales como trozos pulpables, aserrables o para chapas de varias calidades, es posible simular el trozado del árbol, o determinar la mejor forma de hacerlo.

Una de las funciones de ahusamiento más simples, recomendada por Kozak, Munro y Smith en Canadá, es una regresión de la forma

$$\frac{d^2}{D^2} = b_1 + b_2 \frac{h}{H} + b_3 \frac{h^2}{H^2},$$

donde d es el diámetro a la altura h , y D y H son el DAP y la altura total, respectivamente.

Nótese en primer lugar que podría ser deseable forzar la función para que cuando $h = H$ (en el ápice) el diámetro sea cero. Claramente, se debe tener $b_1 + b_2 + b_3 = 0$. Substituyendo b_1 en función de b_2 y b_3 , se ve que esto se puede lograr con una regresión

$$\frac{d^2}{D^2} = b_2 \left(\frac{h}{H} - 1 \right) + b_3 \left(\frac{h^2}{H^2} - 1 \right),$$

o

$$\frac{d^2/D^2}{h/H - 1} = b_2 + b_3\left(\frac{h^2}{H^2} + 1\right),$$

por ejemplo. Si d es un diámetro con corteza, se podría también hacer que $d = D$ para $h = 1,3$ m, dejando un solo parámetro libre.

Una segunda observación es que un polinomio de segundo grado en h no representará bien la forma del fuste cerca de la base. Esto podría mejorarse agregando un término en h^3/H^3 (recuérdese el neiloide $d^2 = kh^3$). En la práctica se ha observado que para representar bien el ensanchamiento basal se necesita generalmente un término con una potencia bastante alta de h , tal como h^8 o h^{40} .

Otra característica de este tipo de ecuación es que implica una “forma” que no cambia con el tamaño de los árboles. Si se grafica d versus h para árboles con distintos D y H , es claro que las curvas se pueden hacer coincidir en toda su longitud eligiendo factores de escala adecuados para los ejes d y h (las curvas de d/D versus h/H son iguales). Frecuentemente se obtienen mejores resultados si se permite que la forma varíe con el tamaño. Para esto, en ecuaciones como la anterior algunos de los b_i se reemplazan por funciones de D y/o H . Nótese que si estas funciones son lineales en sus parámetros, la regresión sigue siendo lineal después de la substitución. Una manera de encontrar formas adecuadas para estas funciones es el ajustar la regresión inicial para cada árbol por separado, y luego graficar los b_i en función de D y H . Este es un ejemplo de problemas llamados a veces, por razones históricas, “armonización de curvas”, y que aparecen frecuentemente en mensura forestal.

Se ha usado una gran variedad de modelos y métodos de estimación para obtener funciones de ahusamiento. Sólo examinaremos dos ejemplos más.

P. L. Real y J. A. Moore (páginas 1037–1044 en *Forest Growth Modelling and Prediction*, USDA Forest Service, General Technical Report NC-120, 1988), usaron el siguiente modelo inicial para pino Oregón, ajustándolo independientemente a los datos de cada árbol:

$$y = b_1(x^3 - x^2) + b_2(x^8 - x^2) + b_3(x^{40} - x^2),$$

donde y es $d^2/D^2 - x^2$ y x es $(H - h)/(H - 4,5)$. Nótese el uso de potencias altas de h , y el condicionamiento para asegurar un diámetro cero al ápice y $d = D$ a la altura del DAP (los d son con corteza, las alturas están en pies y 4,5 es la altura del DAP). Luego los b_1 , b_2 y b_3 de cada árbol se ajustaron a tres regresiones conteniendo (no necesariamente en forma lineal) H , D y la longitud de la copa.

Este modelo es atípico en incluir otra variable predictora además de D y H , la longitud de copa. Esta parece una buena idea, recuérdese las diferencias en forma entre el fuste dentro de la copa (aprox. cónico) y por debajo de ella (aprox. parabólico). Además el largo de copas refleja características del rodal, estando asociado con su densidad, y se sabe que dos árboles con un mismo DAP y altura, uno dominante en un rodal denso y el otro suprimido en un rodal más abierto tendrían formas distintas. Por otro lado, probablemente habría sido

recomendable re-estimar los parámetros directamente con todos los datos luego de substituir las expresiones para los b_i .

El otro ejemplo es de A. Gordon (*New Zealand Journal of Forestry Science* 13,146–155, 1983), para pino radiata. La función es

$$d^2 = \frac{4V}{\pi H} (b_1 z + b_2 z^2 + b_5 z^5 + b_{16} z^{16}),$$

con la restricción $\sum b_i/(i+1) = 1$ impuesta a través de regresión con una función transformada. La variable z es $1 - h/H$, y V es el volumen calculado con una función de volumen cúbico total obtenida con los mismos datos. Esta es lo que se llama una función de ahusamiento *compatible*, concepto desarrollado por Demaerschalk en Canadá. Tiene la propiedad se que al integrar $\pi d^2/4$ con respecto a h entre 0 y H se obtiene exactamente el mismo V estimado por la función de volumen. Aunque la compatibilidad entre funciones de volumen y de ahusamiento es atractiva, no está muy claro si los posibles sacrificios en flexibilidad y precisión valen la pena.

Las funciones basadas en polinomios parecen ser las más comunes, pero también se han usado otras formas, tales como funciones racionales y trigonométricas. El uso de *splines* también se ha popularizado. Estas son combinaciones de varias funciones, comúnmente polinomios cúbicos, cada una válida para cierto rango de la variable independiente y con restricciones de continuidad y suavidad en los puntos donde se juntan.

Desde el punto estadístico, se puede acotar que en las funciones de ahusamiento los supuestos necesarios para la optimalidad de los mínimos cuadrados están lejos de cumplirse. El supuesto de homocedasticidad es poco realista, considerando que hay puntos perfectamente conocidos (el ápice y posiblemente el DAP) cerca de los cuáles los errores deberían ser más pequeños. Igualmente, el supuesto de independencia es insostenible, ya que, en un árbol determinado, diámetros cercanos tienden a desviarse del promedio en la misma dirección.

PREGUNTAS, EJERCICIOS

1. Se tiene una función de ahusamiento (sin corteza):

$$(d/D)^2 = 1,0 - 1,5h/H + 0,5(h/H)^2.$$

Para un árbol de 25 m de altura y 30 cm de DAP calcule:

- (a) El cociente de forma de Girard.
- (b) El volumen cúbico total.
- (c) El número de trozas aserrables de 4 m y diámetro menor de por lo menos 20 cm (suponga un tocón de 30 cm).

3.8 Funciones de corteza

En muchos casos es necesario estimar el espesor de corteza, su volumen, o el diámetro sin corteza a partir del diámetro con corteza o vice-versa. Como una

primera aproximación se pueden obtener regresiones de espesor de corteza, o de la razón entre los diámetros con y sin corteza, en función del diámetro en ese punto. Frecuentemente esta relación entre corteza y diámetro varía a lo largo del fuste, y se consiguen estimaciones más precisas incluyendo la altura de medición como otro predictor.

Debe tenerse presente que las mediciones con el calibrador de corteza son poco precisas y pueden tener sesgos considerables, de modo que en muchos casos es preferible utilizar una estimación en lugar de la medición directa.

Capítulo 4

Rodales

Nos referimos aquí a características agregadas, aplicables al conjunto de los árboles en un pedazo de terreno. Este puede ser un rodal, una hectárea o una parcela de muestreo.

4.1 Diámetro, área basal

Ya hemos considerado antes la medición de diámetros en trozas y árboles individuales, en especial el DAP y su definición. Vimos que generalmente había más interés en el área de la sección transversal, derivada de la medición de diámetro, que en el diámetro en sí. Lo mismo ocurre en un rodal. A este nivel son importantes el área basal y el DAP medio asociado a ésta.

El *área basal* es la suma de las secciones transversales a la altura del pecho, por unidad de área del terreno. Normalmente la sección para cada árbol se calcula a partir del DAP, suponiendo una sección circular, y el área basal se expresa en metros cuadrados por hectárea. En EEUU y algunos otros países se usa pies cuadrados por acre. A veces se obtienen estimaciones de área basal sin medir el DAP, por ejemplo con el método de muestreo puntual de Bitterlich y con regresiones basadas en mediciones sobre fotos aéreas, como se verá en el curso de Inventarios Forestales. El área de un círculo con diámetro igual al DAP de un árbol determinado también se llama su área basal (área basal por árbol).

En Mensura Forestal se entiende generalmente por *DAP medio* o diámetro medio el diámetro del árbol de área basal media. Esto es, si no se especifica otra cosa, el DAP medio es el diámetro medio cuadrático $\sqrt{\sum d_i^2/n}$. Dados el área basal B y el número de árboles por hectárea N , el DAP medio (en metros) es $D = \sqrt{\frac{4}{\pi}B/N}$.

El diámetro medio aritmético es menos usado. Ocasionalmente se definen diámetros dominantes o diámetros tope, basados en los árboles más grandes. Estos se relacionan con las alturas dominantes descritas más abajo.

Nótese que la varianza de una muestra de diámetros es

$$s^2 = \overline{d^2} - \bar{d}^2$$

(la barra sobre una expresión denota una media). Se ve entonces que el diámetro medio (cuadrático) es siempre mayor que el diámetro medio aritmético (a menos que la varianza sea cero).

♡ En 1959 un comité de la Unión Internacional de Instituciones de Investigación Forestal (IUFRO) preparó recomendaciones para estandarizar la notación en Mensura. Se usarían letras minúsculas para variables a nivel del árbol y mayúsculas para rodales. Las letras se eligieron considerando la inicial mayoritaria de las palabras en inglés, francés y alemán: H para altura, D para diámetro, G para área basal. Diversas variaciones se indicarían con subíndices (IUFRO, 1959, “The standardization of symbols in Forest Mensuration”, Reimpreso en: *Technical Bulletin 15*, U. of Maine, Orono, 1965).

El intento de estandarización no ha sido del todo exitoso. Tanto G como B son comunes para denotar área basal. Aunque la recomendación de emplear minúsculas para árboles y mayúsculas para rodales se aplica con cierta frecuencia, muchas veces las minúsculas se confundirían con parámetros.

4.2 Alturas

4.2.1 Curvas altura–diámetro

Por razones de costo y de tiempo, frecuentemente no se miden las alturas de todos los árboles en un rodal o parcela de muestreo. Las alturas medidas en una sub-muestra se utilizan en una regresión para estimar alturas en función del DAP. Generalmente se usan regresiones lineales simples con transformaciones de la variables.

Comparaciones de curvas altura–DAP han sido hechas, por ejemplo, por Curtis (*Forest Science* 13,365–375, 1967), A.R. Ek (p. 67–80 *Statistics in Forest Research*, Proc. of meeting of IUFRO subject group S6.02, Vancouver, 1973), García (INFOR, Nota Técnica N° 19, 1973), y Arabatzis y Burkhart (*Forest Science* 38,192–198, 1992). Algunas ecuaciones que han dado buenos resultados son $H = b_1 + b_2 \log D$, $H = b_1 + b_2/(D + 10)$ y $\log H = b_1 + b_2/D$.

Se trata de un problema típico de regresión, sin mayores complicaciones. Conviene, sin embargo, asegurarse que la curva resultante no sea decreciente, lo que sucede con frecuencia con algunas funciones, muestras de tamaño reducido, y/o grandes errores de medición.

PREGUNTAS, EJERCICIOS

1. Considere funciones de volumen general y local de la forma $V = a + bD^2H$ y $V = c + dD^2$ (a, b, c y d son parámetros). Si estas relaciones fueran exactas, ¿qué forma debería tener la ecuación altura–diámetro?

4.2.2 Altura dominante

Además de la altura media (aritmética), se utilizan varias definiciones de *altura dominante*, también llamada altura tope (top height) o altura superior. El objetivo es tener una variable que no sea apreciablemente afectada por la densidad del rodal (árboles por hectárea), ni por los tratamientos silviculturales, especialmente los raleos. Entre otras cosas, una variable de ese tipo es útil para evaluar la productividad del terreno, como veremos más adelante. La altura no varía tanto con la densidad como el diámetro o área basal. Sin embargo, al hacer un raleo selectivo en el que se extraen árboles de menor tamaño que el promedio, la altura media del rodal residual aumenta. Una altura dominante, basada en los árboles más grandes, varía mucho menos.

Podemos distinguir tres enfoques generales al problema de definir una medida de altura dominante. El primero se basa en la clasificación visual de los árboles en clases de copa, típicamente en dominantes, codominantes, intermedios y suprimidos. Se toma entonces la altura promedio de los dominantes, o la de los dominantes y codominantes. Esta altura es poco sensible a la densidad y raleos (excluyendo los raleos por lo alto, que son poco usuales), y se ha usado bastante. Una objeción se relaciona con su subjetividad; depende de la clasificación de copas, y distintos observadores probablemente obtengan valores diferentes.

Para evitar este problema, se propusieron diversas medidas basadas en alguna proporción de los árboles mayores determinada objetivamente. Este es el segundo tipo de medidas de altura dominante, desarrolladas principalmente en Alemania. Algunas corresponden a la altura media de una proporción fija de los árboles más grandes, tal como el 20% de los árboles de mayor diámetro (propuesta por Weise en 1880). Otras propuestas toman los árboles que exceden el promedio en más de una o dos desviaciones estándar. Es claro que este enfoque es algo deficiente en su sensibilidad a los raleos por lo bajo.

El tercer tipo de alturas dominantes usa un número fijo de árboles por hectárea en lugar de una proporción fija. Por ejemplo, los 100 o 200 árboles mayores por hectárea. Esto se hace en parcelas de muestreo, donde se toma la proporción correspondiente al área de la parcela: para los 100 mayores por hectárea se toman los 10 mayores en una parcela de 1000 m² o los 5 mayores en una parcela de 500 m². Hay una serie de variantes en este enfoque. Se pueden elegir los m árboles más altos, o los m de mayor diámetro. Esto último es frecuentemente bastante más fácil. Cuando las alturas se estiman con una curva altura-DAP los dos métodos son obviamente equivalentes. Una vez seleccionados los m árboles, se puede calcular la media aritmética de sus alturas, ya sean medidas o estimadas. Otra alternativa es calcular la media cuadrática de los DAP para los m árboles (llamada a veces el DAP dominante o DAP tope), y tomar la altura dada por la curva altura-DAP para esa media. Se usa también la altura del árbol más alto o más grueso en parcelas o sub-parcelas de 1/100 ha. En todos los casos se excluyen árboles malformados (quebrados, doble flecha, etc.).

♡ Este tipo de alturas define un procedimiento objetivo de cálculo, y al mismo tiempo no se ve afectado apreciablemente por los raleos por lo bajo o la mortalidad natural. Hay sin embargo un problema de definición que es generalmente ignorado, pero que puede ser importante en algunas situaciones. La altura media esperada de los 100 árboles mayores en una hectárea no es la misma que la de los 10 mayores en 1/10 ha, o la del árbol mayor en 1/100 ha, etc. Este efecto del tamaño de la parcela es fácil de entender pensando en un área de 1/10 ha subdividida en 10 partes de 1/100. Si se elige el árbol más grande en cada una de las 10 partes, es claro que la altura media será en general menor que si se eligen los 10 mayores independientemente de donde se encuentren.

El problema fue identificado por J. Fries (Sveriges Skogsvårdsförbunds Tidskrift:72, 559–563, 1974) y B. Matérn (Sv.Skogs.Tids.:74, 51–53. 1976). En base a observaciones y a aproximaciones teóricas encontraron, por ejemplo, que el valor esperado de la altura media de los 10 árboles mayores en 0,1 ha corresponde al de la altura del árbol mayor en 0,015 a 0,018 ha (en lugar de 0,01 ha). Por lo tanto, en Suecia se ha definido más precisamente la altura dominante como la media de los 10 mayores en 0,1 ha. K. Rennolls (*Comm.For.Rev.* 57:215–219, 1978) también trató el tema y propuso como norma el más grande en 1/100 ha.

♡♡ Se menciona a veces la altura de Lorey. Esta es la altura media ponderada por las áreas basales de los árboles, $\sum b_i h_i / B$. La idea es que multiplicada por el área basal B da una cantidad más estrechamente relacionada con el volumen por hectárea que si se usa la media aritmética. Esto porque el volumen del árbol i es aproximadamente una función lineal de $b_i h_i$.

PREGUNTAS, EJERCICIOS

1. El curso de 1994 obtuvo los siguientes datos de una parcela de 500 m² en *Eucalyptus nitens*:

DAPs: 30 7 33 22 23 24 26 13 30 25 6 23 19 14 19 29 28 25 26 42 20 21 35
42 14 45 24 14 41 23 13 28 42 26 20 22 25 40 27 21 26 27 38 23 46 21 20
16 22 25 31 29 21 21 22 25 25 22 21 17

Arboles muestra:

DAP: 30 22 26 30 28 42 20 21 35 42 14 45 27 26 27 31 29 25

ALT: 22 24 26 30 23 28 20 20 25 27 14 32 27 25 27 27 28 24

Calcule DAP y altura dominantes de acuerdo a las varias alternativas descritas.

2. Separe los datos de DAP en cuatro grupos de aproximadamente igual tamaño. Encuentre los dos árboles mayores en cada grupo. Compare con los ocho árboles más grandes en la parcela.

4.3 Cubicación de rodales/parcelas

En inventarios se estima el volumen del rodal a partir de los volúmenes para cierto número de parcelas de muestreo. Para calcular el volumen de una parcela

se podría calcular el volumen de cada árbol, ya sea por cubicación directa o con una función de volumen general, y sumar. Sin embargo, en la mayoría de los casos el cubicar todos los árboles, o aún medir todas las alturas para aplicar una función de volumen, es demasiado costoso. Se recurre entonces a una estimación indirecta utilizando los DAP de todos los árboles y las alturas o volúmenes en una parte de ellos. Los árboles de la parcela que se miden en detalle se llaman a veces los “árboles muestra”.

Hay dos métodos comúnmente usados. El primero se emplea con funciones generales de volumen (volumen del árbol en función del DAP y altura). Se supone que se cuenta con una función de volumen adecuada. Se empieza por usar los árboles muestra, en los que se ha medido la altura y el DAP, para ajustar una curva de altura-diámetro. Con ésta se estima la altura de los árboles en los que se ha medido sólo el DAP. Teniendo ahora DAP y altura para todos los árboles, se calculan los volúmenes con la función de volumen dada, y se suma. El volumen por hectárea se puede calcular dividiendo por el área correspondiente.

Este método puede verse también como una sustitución en la función de volumen general $V = f(D, H)$ de la altura dada por la función altura-DAP $H = g(D)$, para obtener una función de volumen local $V = f[D, g(D)] \equiv h(D)$.

En el segundo método se determina primero el volumen de cada árbol muestra. Esto se puede hacer por cubicación directa (con diámetros medidos a varias alturas), o con una función de volumen adecuada. Con los volúmenes y DAP de los árboles muestra se ajusta entonces una regresión de volumen en función del DAP (una función local de volumen). Con ésta se calculan los volúmenes de todos los árboles y se suman. A menudo los datos se ajustan bien con una regresión lineal simple de volumen sobre DAP al cuadrado (o sobre área basal por árbol), la llamada *línea de volumen* o recta volumen-área basal¹.

El primer método es muy usado, y puede ser algo menos trabajoso cuando es necesario obtener la curva altura-diámetro de todos modos, por ejemplo para calcular la altura dominante. La función de volumen local puede ser algo sesgada por la forma indirecta en que se obtiene. El segundo método es más directo y más general (no está limitado al uso de funciones de volumen). La línea de volumen presenta normalmente una relación mucho más estrecha entre las variables que la curva altura-diámetro. Usualmente el segundo método da resultados algo mejores, aunque las diferencias no son grandes.

Se acostumbra no incluir como árboles muestra aquéllos que presenten malformaciones tales como flechas múltiples, torceduras o ápices quebrados. Esto proporciona medidas más consistentes a nivel de rodal, tanto para alturas como para volúmenes. Por otro lado, se produce cierta sobreestimación. Es importante también tener presente que lo calculado es un volumen en pie, que difiere del volumen extraído por las pérdidas y deshechos de explotación.

Como comentario de índole general, se podría decir que tal vez se tiende a exagerar la importancia de calcular volúmenes cúbicos. En la práctica al volumen cúbico se le aplican factores de conversión promedios de dudosa exactitud

¹Nada que ver con la también llamada “línea de volumen” de Gray, que es la recta de diámetro al cuadrado sobre altura en el fuste en la porción parabólica de un árbol.

para estimar volúmenes aserrables, valores monetarios, etc. Su utilidad directa es más que nada como una unidad convencional usada tradicionalmente con fines comparativos.

PREGUNTAS, EJERCICIOS

1. Con los datos de la parcela de *E. nitens* de la sección 4.2.2, estime el volumen por hectárea con ambos métodos. Grafique los datos y relaciones.
2. Compare el ECM de las funciones locales de volumen obtenidas.
3. Compare el ajuste a sus datos de esta función altura–DAP con el de las otras funciones indicadas anteriormente, gráficamente y a través del ECM.
4. ¿Qué forma de función de volumen local implicaría la mejor función altura–DAP encontrada en el punto anterior? Compare con la línea de volumen.
5. En una parcela de 500 m² se midieron los siguientes DAP (cm), ordenados de menor a mayor:

6,3 7,4 12,9 13,1 13,6 13,7 13,9 16,2 16,7 19,0 19,1 19,9 20,1 20,5 20,6 21,1
21,2 21,2 21,2 21,5 21,7 21,7 21,8 22,0 22,5 22,7 22,7 22,9 23,4 24,3 24,4
24,6 24,9 25,1 25,2 25,4 25,4 25,7 26,0 26,3 26,4 26,9 27,2 28,1 28,1 28,7
29,4 29,8 30,0 31,0 33,0

($n = 51$, $\sum D = 1136,5$, $\sum D^2 = 26935,99$). Con árboles muestra se obtuvieron la curva altura–DAP $H = -15,1 + 12,0 \ln D$ y la línea de volumen $V = -0,115 + 0,00082D^2$.

- (a) Calcule el volumen por hectárea.
- (b) Calcule una altura dominante basada en los 100 árboles mayores por hectárea.
- (c) Lo mismo que (b), pero con otro método/definición que dé un resultado distinto.

4.4 Funciones de volumen

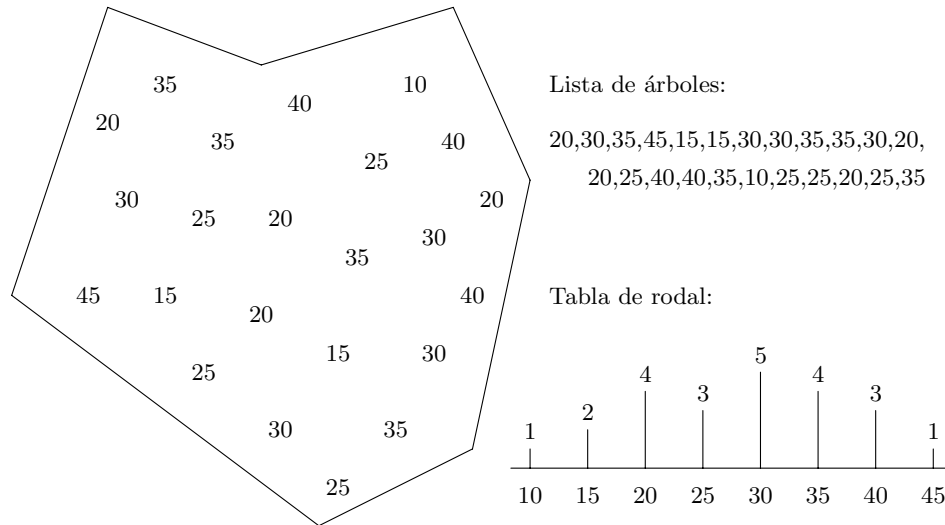
Usando datos obtenidos por cubicación de un número de parcelas, se pueden obtener regresiones de volumen en función de variables de rodal como área basal y altura dominante. Estas son funciones (o “tablas”) de volumen por hectárea.

Es más fácil usar la función de volumen que cubicar una parcela directamente, pudiendo emplearse por razones de conveniencia. Más frecuentemente las funciones de volumen por hectárea son útiles cuando no se tiene la información de los árboles individuales. Un ejemplo podría ser al determinar el área basal directamente con el método de Bitterlich. Otro es cuando el área basal y altura corresponden a la proyección de un modelo de crecimiento.

Se trata de un típico problema de regresión, sin mayores complicaciones. Se usan funciones similares a las de volumen por árbol, con área basal en lugar del DAP al cuadrado. Por ejemplo, $V = b_1 + b_2BH$. Una función que ha dado buenos resultados en plantaciones de coníferas es $V/B = b_1 + b_2H$, donde la división por B ayuda con la heterocedasticidad.

4.5 Tablas de rodal, distribuciones

Los DAP individuales medidos en una parcela o rodal se pueden usar como una descripción más detallada que el resumen proporcionado por el DAP medio o el área basal. Se puede usar directamente el conjunto de diámetros observados (una “lista de árboles”), o estos se pueden presentar en la forma de un histograma o tabla de frecuencias (una “tabla de rodal”).



En la tabla de rodal (*stand table*) tradicionalmente los DAP se agrupan en *clases de DAP*, y el número de árboles en cada clase se presenta en términos de su equivalente por hectárea. Se acostumbra incluir también el volumen por hectárea calculado para cada clase de DAP, lo que se llama una “tabla de existencias” (*stock table*), y a veces además las alturas estimadas y otras variables. El agrupamiento en clases se hace *a posteriori*, o resulta de registrar solamente la clase de diámetro al momento de la medición.

Las tablas de rodal eran muy usadas para facilitar el cálculo de variables agregadas tales como el área basal, DAP medio y volumen por hectárea, sumando valores por clase ponderando por la frecuencia. Con el mejoramiento de los medios computacionales esta aplicación ha perdido su importancia. Hoy día no es recomendable agrupar datos de esta manera con fines de cálculo, ya que la pérdida de precisión no se justifica. Las tablas de rodal y de existencias pueden

ser aún de utilidad como un resumen simple y conveniente de las características de un rodal.

Una alternativa a la tabla de rodal, la cuál es esencialmente un histograma de la distribución diamétrica, es la aproximación de la distribución por funciones continuas. Se usan funciones de distribución de probabilidad, haciendo la analogía entre la proporción de árboles con diámetros en un cierto intervalo y la probabilidad de que una variable aleatoria tome valores dentro del intervalo.

La *función de distribución* $F(x)$ de una variable aleatoria X es la probabilidad de que ésta tome un valor menor o igual que x :

$$F(x) = \Pr\{X \leq x\} .$$

Si la variable es continua y la derivada existe, $f(x) = dF(x)/dx$ es la función de *densidad*. Dada una función de distribución o densidad, la probabilidad de que X tome un valor entre a y b , o análogamente la proporción de árboles con DAP entre a y b , se puede calcular entonces como

$$\Pr\{a < X \leq b\} = F(b) - F(a) = \int_a^b f(x) dx .$$

Se habla en general de una *distribución* de probabilidad, definida por su función de distribución (o distribución acumulativa) o por su densidad.

Para rodales multietáneos, la *ley de Liocourt* (1898) indica que las frecuencias por clases de diámetro disminuirían con el diámetro en progresión geométrica. Esto corresponde a una distribución exponencial con densidad $f(x) \propto e^{-kx}$. En la realidad esto no siempre concuerda con las observaciones.

En rodales coetáneos se usan normalmente densidades unimodales, es decir, con un único máximo bien definido. Entre las distribuciones que se han usado o propuesto están la normal, log-normal, beta, gama, Johnson S_B , y la Weibull. En general no se encuentran grandes diferencias en la representación de los datos por cada una de ellas. Esto no es sorprendente ya que para determinar con seguridad la forma de una distribución se necesitarían muestras extraordinariamente grandes, lo que no ha sido apreciado por la mayoría de los investigadores y usuarios de estas distribuciones.

En la actualidad la más usada es la distribución de Weibull, ya que da resultados aceptables y es matemáticamente conveniente al tener expresiones explícitas tanto para la densidad como para la función de distribución:

$$F(x) = 1 - \exp[-(\frac{x-a}{b})^c], \quad f(x) = \frac{c}{b}(\frac{x-a}{b})^{c-1} \exp[-(\frac{x-a}{b})^c], \quad a \leq x < \infty .$$

Esta es la Weibull con tres parámetros, a , b y c . El supuesto de que es imposible encontrar un DAP menor que a es conceptualmente algo cuestionable, y produce dificultades estadísticas en la estimación. Probablemente es preferible usar la Weibull con dos parámetros, donde $a = 0$. Nótese que en este caso si la distribución de los DAP es Weibull, la de los DAP al cuadrado (o de las áreas basales por árbol) también lo es.

Son comunes los modelos de crecimiento que usan una distribución diamétrica como descripción (estado) del rodal, y predicen como cambian los parámetros de esta distribución en el tiempo. Otros modelos proyectan listas de árboles o tablas de rodal. Las distribuciones, listas y tablas de rodal se usan también para estimar mezclas de productos con diversas especificaciones que se obtendrían de un rodal, siendo la base en muchos casos de sistemas computacionales sofisticados incluyendo simuladores y optimizadores de trozado. Ver, por ejemplo, D.M. Hyink y J.W. Moser (*Forest Science* 29, 85–95, 1983), D.J. Depta (“Integrated forest planning systems at Weyerhaeuser Company”, en Nagumo H. et al (eds) *Proc. IUFRO Symp. on Forest Management Planning and Managerial Economics*, U. of Tokyo, 1984), B.E. Borders y W.D. Patterson (*Forest Science* 36, 413–424, 1990).

Aunque indudablemente las distribuciones (en todas sus formas, incluyendo listas de árboles y tablas de rodal) son útiles y necesarias para estimar mezclas y tamaños de productos, su confiabilidad tiende a sobreestimarse. Aparte de la variabilidad en el muestreo ya mencionada, se ha llevado la analogía con las distribuciones de probabilidad demasiado lejos. El uso de parcelas de muestreo produce una sobre-representación de pares de árboles separados por distancias cortas, y los DAP no están distribuidos al azar en el terreno. La competencia induce correlaciones negativas en los DAP de árboles cercanos, mientras que las variaciones de micrositio inducen correlaciones positivas. En consecuencia, las distribuciones derivadas de datos de parcelas, las que generalmente se obtienen, pueden ser considerablemente distintas de las distribuciones para todo un rodal, las que generalmente se requieren en las aplicaciones (O. García, “What is a Diameter distribution?”, en Minowa, M. y Tsuyuki, S. (eds) *Proc. Symp. Integrated Forest Management Systems*, Japan Soc. of Forest Planning Press, 1992). Estos modelos deben usarse con cautela, apreciando sus limitaciones.

Se han usado diversos métodos para estimar los parámetros de distribuciones de DAP, siendo los principales el de máxima verosimilitud (MV) y el método de momentos. Este último consiste en hacer que los dos o tres primeros momentos de la distribución (dependiendo de si hay que estimar 2 o 3 parámetros) coincidan con los respectivos momentos de la muestra. O sea, con dos parámetros se toman los valores de estos para los cuales la media y la varianza de la muestra y de la distribución teórica son las mismas. Aunque estadísticamente no tan eficiente como MV, hay ciertas ventajas en la consistencia del DAP medios cuadrático observado con el calculado con la distribución (claramente esto ocurre ya sea que se considere la distribución de los diámetros o de las áreas basales).

♡ La media y la varianza de la Weibull (con $a = 0$) son

$$\mu = b\Gamma(1 + 1/c),$$

$$\sigma^2 = b^2[\Gamma(1 + 2/c) - \Gamma^2(1 + 1/c)].$$

(Γ es la llamada *función gama*, definida por $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t} dt$, cuyos valores están tabulados y para los cuáles existen diversas aproximaciones). Los estimadores

de momentos se obtienen resolviendo estas dos ecuaciones para b y c en función de μ y σ^2 . Este despeje no puede hacerse algebraicamente, siendo necesario recurrir a procedimientos de aproximaciones sucesivas.

Más conveniente es usar una fórmula aproximada que da c en función del coeficiente de variación $z = \sigma/\mu$, suficientemente exacta para fines prácticos (O. García, *New Zealand Journal of Forestry Science* 11, 304-306, 1981):

$$1/c = z[1 + (1 - z)^2(-0,22102 + z(0,010061 + z(0,11736 - 0,050999z)))] .$$

Luego se obtiene $b = \mu/\Gamma(1 + 1/c)$. Para tablas y aproximaciones a la función gamma véase por ejemplo Abramowitz, M. y Stegun, I. A., "Handbook of Mathematical Functions", o el artículo recién citado (en APL la función ! calcula $\Gamma(z + 1)$, que para z entero es igual al factorial de z). La aproximación dada no es válida para $c \leq 1$, donde la Weibull toma forma de J invertida en lugar de ser unimodal.

PREGUNTAS, EJERCICIOS

1. Calcule la función de distribución y el factor de proporcionalidad para la exponencial $f(x) \propto e^{-kx}$. Note que el área total bajo $f(x)$, que es lo mismo que $F(\infty)$, debe ser 1.
2. Ajuste una Weibull de dos parámetros a los DAP de *E. nitens* de la sección 4.2.2. Use el método de momentos.
3. Grafique la densidad y la distribución acumulativa obtenidas. Compare con un histograma de los datos.
4. ¿Cuál es el número de árboles por hectárea estimado entre 10 y 20 cm?
5. La distribución diamétrica en un rodal multietáneo es exponencial, con densidad $f(D) = 0,03e^{-0,03D}$. ¿Qué porcentaje de los árboles tiene más de 40 cm de DAP?
6. Dadas n observaciones, la distribución empírica es aquella que asigna la probabilidad $1/n$ a cada valor observado. En una parcela de 500 m² se midieron los siguientes DAP (cm), ordenados de menor a mayor:
6,3 7,4 12,9 13,1 13,6 13,7 13,9 16,2 16,7 19,0 19,1 19,9 20,1 20,5 20,6 21,1 21,2 21,2 21,2 21,5 21,7 21,7 21,8 22,0 22,5 22,7 22,7 22,9 23,4 24,3 24,4 24,6 24,9 25,1 25,2 25,4 25,4 25,7 26,0 26,3 26,4 26,9 27,2 28,1 28,1 28,7 29,4 29,8 30,0 31,0 33,0

Usando la distribución empírica estime el número de árboles por hectárea con DAP entre 20 y 25 cm.

7. La media de la exponencial con función de distribución $F(x) = 1 - e^{-kx}$ es $1/k$. Suponga que en un rodal con 800 árboles por hectárea y área basal de 40 m²/ha los diámetros al cuadrado (o las áreas basales por árbol) siguen una distribución exponencial. Calcule el número de árboles menores de 25 cm, estimando k por el método de momentos.

Apéndice A

Errores

Toda medición está sujeta a error o incertidumbre. Las fuentes de error son variadas, y podrían clasificarse de muchas maneras. Por ejemplo están lo que podríamos llamar “equivocaciones”, debidas a lecturas incorrectas de la escala en un instrumento, errores de transcripción, etc. Hay errores instrumentales, debidos a deficiencias o mal uso de un instrumento. Errores personales, causados por deficiencias en los sentidos del observador, y por manifestación subconsciente de sus intereses o preferencias. Muy importantes y frecuentemente ignorados son los errores debidos al modelo; por ejemplo, en la mayoría de los cálculos con diámetros y secciones transversales de árboles (área basal, etc.) se supone que la sección es circular. Errores sistemáticos son aquellos que actúan siempre en la misma dirección.

Al considerar un instrumento o método que genera una serie (real o hipotética) de mediciones es útil distinguir entre *exactitud* y *precisión*. Exactitud se refiere a la cercanía entre las mediciones y el valor verdadero. Precisión tiene que ver con consistencia, cercanía de las mediciones entre sí. Mediciones pueden ser precisas pero inexactas. Algunos autores entienden por exactitud la ausencia de errores sistemáticos (“sesgos”), cercanía de la media de las mediciones al valor verdadero.

A.1 Límites de error

En cálculos de ingeniería se acostumbra a trabajar con incertidumbres o errores estimados que se supone cubren el valor verdadero. Es decir, se da un valor de la forma $x \pm \Delta x$, donde x es el valor estimado y Δx es un error máximo acotando el valor verdadero (se toma como un número positivo, el valor absoluto del error). En otras palabras, por *error* entendemos aquí un límite de error.

En casos particulares el error en el resultado de un cálculo con cantidades sujetas a error puede determinarse substituyendo todas las combinaciones de errores negativos y positivos, y tomando los resultados extremos (si está claro cuales serían las situaciones más desfavorables las combinaciones a probar

pueden reducirse). Es una buena idea hacer esto en casos importantes. Los métodos descritos a continuación son más convenientes y producen relaciones entre errores y variables que pueden ser útiles.

Está claro que en una suma o resta los errores se suman, ya que se suponen independientes y no se sabe en que dirección actúan (para una cota hay que tomar la situación más desfavorable):

$$(x \pm \Delta x) + (y \pm \Delta y) = (x + y) \pm (\Delta x + \Delta y)$$

$$(x \pm \Delta x) - (y \pm \Delta y) = (x - y) \pm (\Delta x + \Delta y) .$$

La multiplicación o división es algo más complicada:

$$(x \pm \Delta x)(y \pm \Delta y) = xy \pm x\Delta y \pm y\Delta x \pm \Delta x\Delta y .$$

El último término es pequeño en relación a los otros, y omitiéndolo podemos escribir (suponiendo que x e y son positivos)

$$(x \pm \Delta x)(y \pm \Delta y) = xy \pm xy(\Delta x/x + \Delta y/y) .$$

$\Delta x/x$ es el *error relativo* de x (mientras que Δx es el *error absoluto*). Se ve entonces que el error relativo del producto es aproximadamente la suma de los errores relativos de los factores. Lo mismo sucede con la división.

Más en general, el error en una función de x e y puede aproximarse usando los primeros términos de la serie de Taylor:

$$g(x + \Delta x, y + \Delta y) = g(x, y) + \frac{\partial g(x, y)}{\partial x} \Delta x + \frac{\partial g(x, y)}{\partial y} \Delta y + \dots .$$

Los términos omitidos contienen productos de errores y, al igual que en el caso de la multiplicación, se pueden despreciar si los errores no son muy grandes. Considerando la incertidumbre en los signos de los errores tenemos entonces que en el peor de los casos el error en g es aproximadamente

$$\Delta g = \left| \frac{\partial g(x, y)}{\partial x} \right| \Delta x + \left| \frac{\partial g(x, y)}{\partial y} \right| \Delta y .$$

La generalización a cualquier número de variables es obvia.

Veamos dos ejemplos simples.

(i) Sea $z = g(x, y) = xy$. Entonces

$$\Delta z = |y|\Delta x + |x|\Delta y ,$$

que coincide con lo obtenido más arriba.

(ii) El error de la función de una variable $g(x) = \ln x$ es

$$\Delta \ln x = \left| \frac{1}{x} \right| \Delta x = \frac{\Delta x}{x}$$

(x debe ser positivo), de modo que *el error relativo de x es aproximadamente igual al error absoluto de $\ln x$.*

PREGUNTAS, EJERCICIOS

1. Use la relación $\ln xy = \ln x + \ln y$ y el resultado del ejemplo (ii) para establecer la relación entre los errores relativos de xy , x e y . Obtenga también el error relativo de x/y .
2. Calcule el (límite de) error en la altura de un árbol dados los errores en la medición de distancia y en las mediciones de los ángulos al ápice y a la base.
3. Suponga que el error de la altura está dominado por el error en el ángulo α entre la horizontal y el ápice, y éste es independiente de α (los otros errores son prácticamente cero). Demuestre que el error es mínimo cuando $\alpha = 45^\circ$.

A.2 Cifras significativas

Una alternativa a expresar el error en la forma $x \pm \Delta x$ es el uso de cifras significativas. Las cifras significativas son los dígitos, excluyendo ceros usados meramente para establecer la posición de la coma decimal. Por ejemplo, los números 1302, 0,8206, 0,0002135, 60,60 y $1,490 \times 10^3$ tienen todos 4 cifras significativas. Sin mayor información no se sabe si el número 1490 contiene 3 o 4 cifras significativas.

Las prácticas para la expresión de errores con cifras significativas no son totalmente uniformes. Usualmente se asume incertidumbre en la última cifra dada, con ésta dando una idea del valor más probable (las cifras “significan algo”). Algunos autores (por ejemplo Husch) usan el criterio más estricto de que el error no debe sobrepasar una unidad en la última cifra. Otros aceptan alguna incertidumbre en la penúltima cifra. En general, se considera que no tiene sentido especificar más de una o dos cifras en Δx , y que x debe especificarse hasta el dígito correspondiente a la última cifra del error. Lo correcto es $15,04 \pm 0,15$ y no $15,036 \pm 0,15$. Más cifras aparentarían una falsa exactitud, y menos producirían una pérdida de exactitud innecesaria.

En cualquier caso, el número de cifras significativas refleja el error relativo, mientras que el número de decimales refleja el error absoluto. La precisión indicada por las cifras significativas o el error relativo es independiente de la unidad de medida: 3,24 m y 324 mm tienen la misma precisión.

Estas relaciones entre cifras y errores permiten establecer ciertas reglas sobre las cifras significativas a usar en los resultados de operaciones aritméticas. El error de una suma o resta está dominado por el error absoluto más grande en sus componentes (los errores máximos se suman, como vimos arriba; otras medidas de error se combinan con menor peso de los errores menores, como veremos luego). Se adopta entonces la norma de dar el resultado con un *número de decimales* igual al número de decimales más pequeño entre los términos que se suman o restan:

$$\begin{array}{r}
123 \\
32,3 \\
+ 0,276 \\
\hline
156
\end{array}$$

En multiplicación y división ocurre lo mismo con los errores relativos, de modo que en el resultado se usa el *número de cifras significativas* menor entre los factores:

$$754,1 \times 0,052 = 39$$

En los pasos intermedios de una secuencia de cálculos conviene retener cifras adicionales y redondear el resultado final.

Es importante tener en cuenta que en algunas operaciones pueden ocurrir pérdidas importantes de cifras significativas (precisión). Este es el caso de la “cancelación catastrófica” al restar números grandes cercanos entre sí.

PREGUNTAS, EJERCICIOS

1. Indique el número de cifras significativas en: (a) 1,00025 (b) 0,002710 (c) 10,003 (d) 100000
2. En los ejemplos de suma y multiplicación dados recién, suponga errores de ± 2 unidades en la última cifra significativa. Calcule los límites de error por sustitución de valores extremos. Compare con las cifras significativas.
3. En una evaluación de regímenes de manejo se calculan ingresos de \$3.274.531 y costos de \$3.256.890. Calcule la utilidad esperada.
 - (a) Suponga ahora un error de aproximadamente 1%. Repita el cálculo de utilidad usando las cifras significativas apropiadas. ¿Qué puede decir sobre la rentabilidad?
 - (b) Con los errores de 1%, obtenga límites de error sustituyendo los valores más optimistas y los más pesimistas.
4. La varianza de una muestra se puede calcular como $\frac{1}{n} \sum (x_i - \bar{x})^2$, donde \bar{x} es la media $\frac{1}{n} \sum x_i$. Frecuentemente se recomienda simplificar los cálculos usando la fórmula $\frac{1}{n} \sum x_i^2 - \bar{x}^2$.
 - (a) Demuestre que las dos fórmulas son matemáticamente equivalentes.
 - (b) Calcule con ambas fórmulas la varianza de los tres números $x_1 = 100001$, $x_2 = 100002$ y $x_3 = 100003$. ¿Qué sucede?

♡ El enfoque estadístico

Al calcular límites de error tomamos la situación más desfavorable, con los signos de los varios errores tales que produzcan el mayor error en el resultado. Por ejemplo, al sumar x e y se supone que Δx y Δy actúan en la misma dirección, positiva o negativa, sumándose sus efectos. Esto es útil porque nos da una cota superior para el máximo error que podría ocurrir. Sin embargo, especialmente al intervenir varias variables, estos límites pueden ser tan amplios que pierden su utilidad, y puede parecer poco realista que todos los errores se confabulen para producir el peor resultado posible. En lugar de los límites de error, se puede trabajar entonces con un modelo estadístico o probabilístico de la incertidumbre en las mediciones.

La *Estadística* trata del uso de información en situaciones de incertidumbre. Utiliza la *Teoría de Probabilidades*, que trata de las propiedades matemáticas de ciertos modelos de incertidumbre.

Una cantidad incierta puede tomar cualquier valor dentro de un conjunto de valores posibles. Algunos valores son más plausibles que otros, de modo que les damos ponderaciones distintas. Estas ponderaciones pueden representar frecuencias relativas de ocurrencia en observaciones repetidas, un grado de creencia subjetiva en los diversos valores, etc. En el modelo de esta situación representamos la cantidad incierta por una *variable aleatoria*, y las ponderaciones por una probabilidad. Como siempre, la teoría y manipulaciones matemáticas del modelo no dependen de su interpretación, pero obviamente ésta es importante al considerar la aplicabilidad de los resultados.

Por ahora consideraremos cantidades que toman valores numéricos de modo que la ponderación se pueda representar por una función de densidad de probabilidad definida sobre los números reales. La probabilidad de que la variable aleatoria X se encuentre entre a y b es $\int_a^b f(x) dx$. Obviamente, $\int_{-\infty}^{\infty} f(x) dx = 1$. A veces es conveniente distinguir la variable aleatoria X de los valores observados x .

SITUACION PRACTICA		MODELO PROBABILISTICO
Incertidumbre en x	\rightsquigarrow	X es una variable aleatoria
Ponderación de posibles valores	\rightsquigarrow	densidad $f(x)$
Media ponderada de $g(x)$	\rightsquigarrow	valor esperado $E[g(X)]$
El valor esperado o esperanza de una función $g(X)$ es la media ponderada		

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x) dx .$$

Casos especiales importantes:

$$\text{Media: } E[X] = \mu$$

$$\text{Varianza: } E[(X - \mu)^2] = E[X^2] - (E[X])^2 = \sigma^2 = V[X] .$$

La media es una medida de ubicación, el centro de gravedad alrededor del cuál se distribuye la incertidumbre. La desviación estándar (o típica) $\sigma = \sqrt{\sigma^2}$ es una medida importante de dispersión.

Volviendo a los errores, representemos una medición o valor observado o calculado por una variable aleatoria X , y denotemos el valor verdadero por x_0 . El error (otra variable aleatoria) es $\varepsilon = X - x_0$, o sea, $X = x_0 + \varepsilon$. Entonces, $E[\varepsilon] = \mu_\varepsilon$ es el *sesgo*. Una medida de precisión es el *error estándar* $\sqrt{V[\varepsilon]} = \sigma_\varepsilon$ (es común llamar error estándar a la desviación estándar de un estimador). Otra medida que combina exactitud y precisión es el *error cuadrático medio*: $ECM = \sqrt{E[\varepsilon^2]}$. Nótese que

$$ECM^2 = \sigma_\varepsilon^2 + \mu_\varepsilon^2 = \text{varianza} + \text{sesgo}^2 .$$

El error límite o máximo absoluto que usamos anteriormente sería (si es que existe): $\Delta x = \max |\varepsilon|$, y el relativo, $\Delta x/x_0$ (o $\Delta x/(x_0 + \varepsilon)$) que es casi lo mismo si el error es pequeño).

Para estudiar la propagación de errores al calcular con variables sujetas a error (variables aleatorias) necesitamos algunas propiedades simples de la esperanza y de la varianza. De su definición como integral se deduce fácilmente que el valor esperado es un operador lineal:

$$E[aX + bY] = aE[X] + bE[Y] .$$

Veamos la varianza de una combinación lineal.

$$\begin{aligned} V[aX + bY] &= E[(aX + bY - E[aX + bY])^2] = E[\{a(X - E[X]) + b(Y - E[Y])\}^2] \\ &= E[a^2(X - E[X])^2 + 2ab(X - E[X])(Y - E[Y]) + b^2(Y - E[Y])^2] \\ &= a^2V[X] + b^2V[Y] + 2abE[(X - E[X])(Y - E[Y])] . \end{aligned}$$

La esperanza en el último término es la *covarianza* de X e Y , $Cov[X, Y]$. Tenemos entonces

$$V[aX + bY] = a^2V[X] + b^2V[Y] + 2abCov[X, Y] .$$

La covarianza está relacionada con el *coeficiente de correlación*

$$\rho = \frac{Cov[X, Y]}{\sqrt{V[X]V[Y]}} ,$$

el que es cero si X e Y son independientes (más precisamente, no correlacionadas), y puede llegar a 1 si X e Y tienden a variar en forma conjunta y a -1 si varían en forma opuesta. Notemos por último que si a no es aleatoria,

$$V[X + a] = V[X] .$$

♡♡ La densidad $f(x)$ que definía la probabilidad de intervalos en la recta de x se generaliza a espacios de más dimensiones. Por ejemplo la *densidad conjunta* $f(x, y)$ aplicada al plano de puntos especificados por pares de coordenadas (x, y) . (Estos pares y sus análogos en más dimensiones pueden considerarse como listas de números, o *vectores*). Las variables aleatorias X e Y se dice que son independientes si su densidad conjunta es de la forma $f(x, y) = f_1(x)f_2(y)$. Una consecuencia que se deduce de la definición de esperanza como integral múltiple es que si las variables son independientes entonces $E[XY] = E[X]E[Y]$. Se verifica fácilmente que esto implica $Cov[X, Y] = 0$. Cabe mencionar que covarianza cero (variables no correlacionadas) no implica necesariamente que las variables sean independientes, excepto en el importante caso de la distribución Normal.

Estamos listos ahora para examinar la propagación de errores. Veamos primero el caso de una suma.

$$E[\varepsilon_{x+y}] = E[(X + Y) - (x_0 + y_0)] = E[\varepsilon_x + \varepsilon_y] = E[\varepsilon_x] + E[\varepsilon_y] ,$$

de modo que los sesgos se suman.

$$V[\varepsilon_{x+y}] = V[\varepsilon_x + \varepsilon_y] = V[\varepsilon_x] + V[\varepsilon_y] + 2Cov[\varepsilon_x, \varepsilon_y] .$$

Si los errores actúan en forma independiente, se ve que el error estándar de la suma es

$$\sigma_{x+y} = \sqrt{\sigma_x^2 + \sigma_y^2} .$$

Medido en esta forma el error crece menos que el error máximo Δ .

Para el caso general usamos, como antes, la serie de Taylor:

$$\begin{aligned}\varepsilon_g &= g(X, Y) - g(x_0, y_0) = g(x_0 + \varepsilon_x, y_0 + \varepsilon_y) - g(x_0, y_0) \\ &\approx \frac{\partial g(x_0, y_0)}{\partial x_0} \varepsilon_x + \frac{\partial g(x_0, y_0)}{\partial y_0} \varepsilon_y ,\end{aligned}$$

Suponiendo errores independientes, se tiene entonces aproximadamente

$$\sigma_g^2 = \left(\frac{\partial g(x_0, y_0)}{\partial x_0} \right)^2 \sigma_x^2 + \left(\frac{\partial g(x_0, y_0)}{\partial y_0} \right)^2 \sigma_y^2 .$$

También podría haberse usado en las derivadas las medias o los valores observados en lugar de los valores reales x_0 e y_0 . Las aproximaciones seguirían siendo válidas siempre que los errores no sean muy grandes.

Usemos esto para calcular el error estándar de un logaritmo:

$$\sigma_{\ln x}^2 = (1/x_0)^2 \sigma_x^2 .$$

Usando la media en lugar de x_0 ,

$$\sigma_{\ln x} = \sigma_x / \mu_x .$$

La expresión en el lado derecho es el coeficiente de variación (CV) de x .

PREGUNTAS, EJERCICIOS

1. Obtenga una expresión para el coeficiente de variación del producto de dos variables independientes X e Y en función de los coeficientes de variación de los factores.
2. Para el problema anterior, grafique $CV(XY)/CV(X)$ sobre $CV(Y)/CV(X)$ para $CV(X) > CV(Y)$. ¿Qué efectos tienen el error menor y el error mayor en el error del resultado? ¿Implicaciones para la construcción de modelos?

Apéndice B

Regresión

Los métodos de regresión son fundamentales en Mensura. Para presentar el tema en una forma más compacta y general primeramente revisaremos algunos conceptos de matrices.

B.1 Matrices

Una *matriz* de orden $n \times m$ es simplemente una tabla de números con n filas (líneas) y m columnas:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}].$$

Los a_{ij} son los *elementos* de la matriz. En lugar de los paréntesis cuadrados también se usan paréntesis redondos o líneas verticales dobles: $\|a_{ij}\|$.

Un *vector* es una lista de números. En álgebra de matrices se consideran como matrices con una sola fila (vector fila) o una sola columna (vector columna). A menos que se diga lo contrario se supondrá que son columnas. Generalmente se representan por letras minúsculas, frecuentemente subrayadas o en negrilla:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_i].$$

La *transpuesta* de una matriz es la matriz que se obtiene intercambiando filas y columnas. La transpuesta de A se representa por A' o A^T .

La *suma* de dos matrices es la matriz de sumas de sus elementos:

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}].$$

Obviamente A y B tienen que ser del mismo orden.

Un número, para distinguirlo de un vector o matriz, se llama un *escalar*. El producto de un escalar y una matriz se obtiene multiplicando cada uno de los elementos de la matriz por el escalar:

$$kA = k[a_{ij}] = [ka_{ij}] .$$

De aquí, la resta o diferencia de matrices es

$$A - B = A + (-1)B = [a_{ij} - b_{ij}] .$$

El *producto matricial* $AB = C$ se obtiene de la siguiente manera:

$$[c_{ij}] = \left[\sum_k a_{ik} b_{kj} \right] .$$

Es decir, el elemento ij del producto es la suma de productos de los elementos de la fila i de A con los de la columna j de B. Claramente para que el producto esté definido se necesita que el número de columnas de la primera matriz sea igual al número de filas de la segunda.

El definir el producto de esta manera es útil, por ejemplo, para manejar sistemas de ecuaciones lineales. El sistema de ecuaciones

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

se puede escribir simplemente

$$A\mathbf{x} = \mathbf{b} .$$

Una suma de cuadrados es

$$e_1^2 + e_2^2 + \cdots + e_n^2 = \sum_{i=1}^n e_i^2 = \mathbf{e}'\mathbf{e} ,$$

donde

$$\mathbf{e} = [e_1 e_2 \cdots e_n]' .$$

Aunque dos matrices tengan las dimensiones adecuadas para obtener los productos AB y BA , en general los resultados son diferentes (el producto matricial no es conmutativo). Aparte de esto, y de que las operaciones no siempre son posibles (tienen que haber ciertas correspondencias entre números de filas y columnas), la suma, resta y producto de matrices se comportan como las operaciones con escalares. Por ejemplo,

$$\begin{aligned} A(B + C) &= [a_{ij}][b_{ij} + c_{ij}] = \left[\sum_k a_{ik}(b_{kj} + c_{kj}) \right] = \left[\sum_k a_{ik}b_{kj} + \sum_k a_{ik}c_{kj} \right] \\ &= AB + AC . \end{aligned}$$

PREGUNTAS, EJERCICIOS

1. Demuestre que la suma es conmutativa, $A + B = B + A$, y asociativa, $(A + B) + C = A + (B + C)$.
2. Demuestre que $(AB)' = B'A'$.
3. Calcule AB y BA , donde

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}.$$

4. Calcule $\mathbf{x}'\mathbf{y}$ y $\mathbf{y}'\mathbf{x}$, donde

$$\mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}.$$

Nota: A menudo una matriz con un solo elemento se considera como un escalar.

5. Demuestre que $p(A + B) = pA + pB$ y $(p + q)A = pA + qA$.
6. Demuestre que el producto es asociativo: $(AB)C = A(BC)$.

Las matrices con unos en la diagonal y ceros en el resto

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

se llaman *identidad*. Actúan como el número 1 en los escalares; el multiplicar cualquier matriz por la identidad (del orden correspondiente) no la cambia:

$$IA = AI = A.$$

Falta un análogo de la división de los escalares. Tal como la resta se puede ver como la suma con un negativo, $a - b = a + (-b)$, la división se puede ver como la multiplicación por un recíproco: $a/b = a(1/b) = ab^{-1}$, donde $b^{-1}b = 1$. En matrices el análogo del recíproco es la *inversa*,

$$A^{-1}A = AA^{-1} = I.$$

Nótese que para que esto tenga sentido A debe ser cuadrada (igual número de filas que de columnas). Aún así, no todas las matrices cuadradas tienen una inversa, las que no la tienen se llaman *singulares*.

Usando la inversa podríamos escribir la solución del sistema de ecuaciones $A\mathbf{x} = \mathbf{b}$ dado anteriormente:

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

Para que esta solución exista A debe ser cuadrada ($m = n$, es decir, el número de ecuaciones debe ser igual al número de incógnitas). Además, para que A no sea singular las ecuaciones deben ser “linealmente independientes” (no deben haber ecuaciones redundantes). Existen diversos métodos para invertir matrices, uno de los más usados es la eliminación de Gauss. Esta puede usarse también para resolver sistemas de ecuaciones sin computar toda la inversa.

No es difícil comprobar las siguientes propiedades:

$$(AB)^{-1} = B^{-1}A^{-1}$$

$$(A')^{-1} = (A^{-1})'$$

Por último, podemos definir derivadas de vectores y matrices. La derivada de una matriz con respecto a un escalar y la derivada de un escalar con respecto a una matriz se definen simplemente como la matriz de las derivadas. Es fácil entonces verificar resultados como estos (A y \mathbf{a} contienen constantes):

$$\frac{dA\mathbf{x}}{dt} = A \frac{d\mathbf{x}}{dt}$$

$$\frac{d\mathbf{a}'\mathbf{x}}{d\mathbf{x}} = \mathbf{a}$$

$$\frac{d\mathbf{x}'\mathbf{x}}{d\mathbf{x}} = 2\mathbf{x}$$

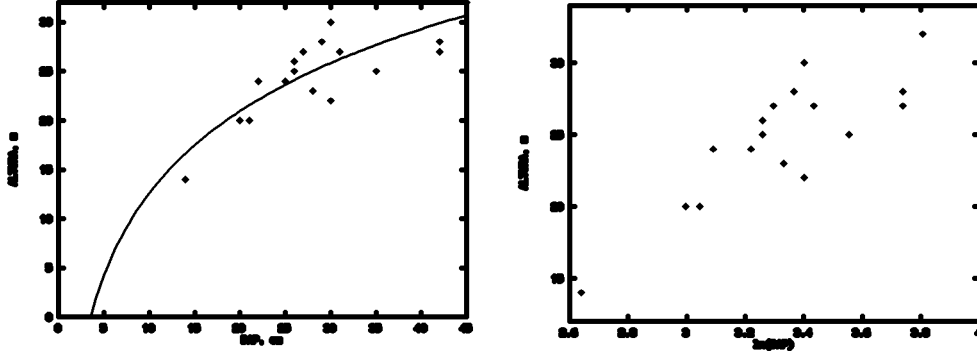
$$\frac{d\mathbf{x}'A\mathbf{x}}{d\mathbf{x}} = (A + A')\mathbf{x}$$

etc. En general los resultados son similares a los obtenidos con escalares, teniendo en cuenta la no conmutatividad de los productos.

B.2 El método de mínimos cuadrados

Muchos métodos de mensura se basan en relaciones entre una *variable dependiente* y una o más *variables independientes* o *predictores*. Interesa describir la relación entre las variables, o estimar o predecir el valor de la variable dependiente conociendo el valor de los predictores. Por ejemplo, se puede usar la relación entre alturas y diámetros para estimar la altura de un árbol conociendo su DAP, el que es más fácil de medir. O estimar el volumen conociendo su DAP y altura. O predecir el volumen de un rodal a una edad determinada.

Tomemos como ejemplo una relación entre dos variables. Es útil construir un *diagrama de dispersión*, graficando las observaciones disponibles con el predictor en las abscisas (“eje x ”) y la variable dependiente en las ordenadas (“eje y ”). El gráfico de la izquierda muestra observaciones de altura y DAP en un rodal de *Eucalyptus nitens* tomadas por el curso de 1994.



Se puede usar una curva como la indicada para estimar las alturas de árboles del rodal para los que solo se conoce el DAP. Es claro que el conocer el DAP ayuda a estimar la altura, es decir, contribuye a reducir la incertidumbre sobre su valor. La curva es un “modelo” que proporciona valores de altura para ser usados en lugar de los valores desconocidos, o que puede servir como una descripción resumida de las observaciones. En cualquier caso conviene tener una ecuación para la curva con el fin de facilitar su uso, y la curva debería pasar “cerca” de las observaciones.

En algunos casos existen razones de tipo teórico que sugieren un tipo de ecuación determinada. En otros, como en este ejemplo, la ecuación es puramente empírica, elegida con criterios de conveniencia y de ajuste a los datos. En general se tendrá una clase de ecuaciones o modelos $y = f(\mathbf{x}, \mathbf{b})$, donde y es la variable dependiente, \mathbf{x} es el vector de variables independientes, y \mathbf{b} es un vector de parámetros cuyos valores se determinarán para producir un buen ajuste. Con un \mathbf{x} de dimensión 2 se tendrá una superficie en lugar de una curva, y para dimensiones mayores una hipersuperficie. Para elegir la forma de la ecuación se puede utilizar experiencias con problemas similares, ensayo y error, gráficos con transformaciones que produzcan una tendencia lineal en los datos, consideraciones sobre la forma que la curva debería tomar al tender a los valores extremos, etc. En el ejemplo se ha usado $H = f(D, b_1, b_2) = b_1 + b_2 \ln D$, viéndose en el diagrama de dispersión del lado derecho que la relación entre H y $\ln D$ es aproximadamente lineal (obsérvese de paso que la extrapolación a diámetros pequeños fuera del rango de los datos eventualmente produce alturas negativas). Siempre se podría elegir una curva que pase por cada una de las observaciones. Aunque en cierto sentido ésta describiría perfectamente los valores observados, en general curvas mucho menos irregulares, con un reducido número de parámetros, producirán mejores estimaciones para valores futuros o no observados.

Una vez decidida la forma de una ecuación a probar, hay que elegir los valores de los parámetros que resulten en un buen ajuste. Se puede suponer que para un D dado la diferencia entre el H desconocido y $f(D, b_1, b_2)$ tenderá a ser más pequeña si estas diferencias son pequeñas para los valores observados. O

sea, \mathbf{b} debe ser tal que los valores absolutos de los *desvíos*, *residuos* o “errores” $e_i = H_i - f(D_i, b_1, b_2)$ sean pequeños para todas las observaciones (D_i, H_i) . Obviamente, si tratamos de reducir un e_i más allá de cierto valor los otros e_i aumentarán, de modo que necesitamos algún criterio que considere el conjunto de todos estos. Un posible criterio sería minimizar la suma de los valores absolutos $\sum |e_i|$ (“regresión en la norma L_1 ”). Otra posibilidad sería minimizar el error más grande ($\min \max |e_i|$, el criterio *minimax*). El criterio más usado, por ser matemáticamente conveniente y por tener en algunos casos ciertas justificaciones de tipo estadístico que veremos más adelante, es el de *mínimos cuadrados* que consiste en minimizar $\sum e_i^2$.

Se tiene entonces un modelo $y = f(\mathbf{x}, \mathbf{b})$, n observaciones (y_i, \mathbf{x}_i) , $i = 1, 2, \dots, n$, y se busca un \mathbf{b} tal que haga mínimo

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n [y_i - f(\mathbf{x}_i, \mathbf{b})]^2 .$$

Equivalentemente, se minimiza el *error cuadrático medio* (ECM) $\sqrt{\frac{1}{n} \sum e_i^2}$, el que es una medida útil de la bondad del ajuste. En general, este problema de optimización no puede resolverse analíticamente y es necesario recurrir a métodos numéricos de optimización iterativos. Una excepción importante se presenta cuando el modelo es una función lineal de los parámetros \mathbf{b} . En este caso de *regresión lineal* se pueden obtener soluciones explícitas para los valores óptimos (de mínimos cuadrados) de los parámetros o *coeficientes*.

Nuestro ejemplo de H vs. D es un caso de regresión lineal. Se puede escribir

$$y = b_1 + b_2 x ,$$

con $y = H$, $x = \ln D$. Esta es una recta, considerándose aquí como predictor la variable x . En general tanto y como x pueden ser transformaciones dadas de las variables originales. Idealmente los datos satisfecerían el sistema de n ecuaciones

$$\begin{aligned} y_1 &= b_1 + b_2 x_1 \\ y_2 &= b_1 + b_2 x_2 \\ &\vdots \\ y_n &= b_1 + b_2 x_n \end{aligned}$$

que con notación matricial se pueden escribir como

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$\mathbf{y} = \mathbf{Xb} .$$

Si tuviéramos $n = 2$, se tendría un sistema de dos ecuaciones con dos incógnitas (b_1 y b_2), generalmente con una única solución. En términos de matrices, $\mathbf{y} = \mathbf{X}\mathbf{b}$ con \mathbf{X} cuadrada e invertible tiene la solución $\mathbf{b} = \mathbf{X}^{-1}\mathbf{y}$.

Con $n > 2$ en general no todas las observaciones son colineales, y el sistema de ecuaciones es incompatible. El objetivo es obtener un \mathbf{b} tal que la aproximación $\mathbf{y} \approx \mathbf{X}\mathbf{b}$ sea lo mejor posible, en el sentido de minimizar la longitud $|\mathbf{e}|$ del vector $\mathbf{e} = \mathbf{y} - \mathbf{X}\mathbf{b}$ calculada por la generalización del teorema de Pitágoras a n dimensiones:

$$|\mathbf{e}|^2 = \sum_{i=1}^n e_i^2 = \mathbf{e}'\mathbf{e}.$$

Existen algoritmos, basados en factorización de matrices, que permiten obtener directamente la solución de mínimos cuadrados de $\mathbf{y} \approx \mathbf{X}\mathbf{b}$. Estos son usados en los paquetes estadísticos de mejor calidad. Se utilizan también a veces *seudoinversas* o *inversas generalizadas*, \mathbf{X}^- , en términos de las cuales la solución se puede escribir como $\mathbf{b} = \mathbf{X}^-\mathbf{y}$. El lenguaje de computación APL que usaremos en las prácticas tiene un operador de inversión generalizada y de división generalizada de matrices que hace muy simple el cálculo de regresiones lineales. En la notación APL el producto matricial $\mathbf{X}\mathbf{b}$ es $X + . \times B$ (indicando que se trata de sumas de productos). Los coeficientes se pueden obtener con la inversa generalizada, $B \leftarrow (\mathbb{X}X) + . \times Y$ o, preferentemente, con la división matricial generalizada $B \leftarrow Y \mathbb{X}X$.

Antes de presentar la solución de mínimos cuadrados usada más comúnmente en los textos y en los cálculos manuales, veamos la situación más general de regresión lineal *múltiple* donde en contraposición al ejemplo anterior de regresión lineal *simple* en que había un solo predictor x , se tienen p predictores. El modelo es

$$y = b_1x_1 + b_2x_2 + \dots + b_px_p = \mathbf{b}'\mathbf{x} = \mathbf{x}'\mathbf{b}.$$

La regresión lineal simple es el caso especial $p = 2$, $\mathbf{b} = (b_1, b_2)$, $\mathbf{x} = (1, x)$. El sistema de ecuaciones, incluyendo ahora los desvíos e_i , es

$$\begin{aligned} y_1 &= b_1x_{11} + b_2x_{12} + \dots + b_px_{1p} + e_1 \\ y_2 &= b_1x_{21} + b_2x_{22} + \dots + b_px_{2p} + e_2 \\ &\vdots \\ y_n &= b_1x_{n1} + b_2x_{n2} + \dots + b_px_{np} + e_n \end{aligned}$$

o sea,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}.$$

La ecuación matricial es igual que antes, otra vez hay que minimizar $\mathbf{e}'\mathbf{e}$, y la solución directa por factorización y con APL no cambian. Casi siempre se incluye una constante en el modelo, y entonces x_1 y los x_{i1} son iguales a 1.

La solución explícita más usual se obtiene como sigue. Para minimizar la suma de cuadrados $Q = \mathbf{e}'\mathbf{e}$ igualamos a cero la derivada:

$$\begin{aligned} Q &= \mathbf{e}'\mathbf{e} = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}) \\ &\quad \mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\mathbf{b} + \mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b} \\ \frac{dQ}{d\mathbf{b}} &= -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\mathbf{b} = 0, \end{aligned}$$

lo que nos da las *ecuaciones normales*:

$$\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y}.$$

Las p ecuaciones se pueden resolver numéricamente para las p incógnitas \mathbf{b} . También puede escribirse la solución en forma explícita:

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

La calidad del ajuste puede ser evaluado a través de la suma de cuadrados $\mathbf{e}'\mathbf{e}$, del error cuadrático medio $\text{ECM} = \sqrt{\mathbf{e}'\mathbf{e}/n}$, o del error standard $\text{ES} = \sqrt{\mathbf{e}'\mathbf{e}/(n-p)}$. El número de parámetros p en el ES penaliza algo la complejidad del modelo al comparar modelos alternativos, y tiene también una justificación estadística explicada en la sección siguiente.

Aunque esta expresión es útil en derivaciones teóricas, en general no es lo más recomendable desde el punto de vista numérico. Primero, el sistema de ecuaciones normales se puede resolver con menos trabajo que el requerido para computar la inversa y el producto matricial. Segundo, pueden producirse errores importantes por cancelación catastrófica, similares a los del cálculo de varianzas demostrados en la sección sobre propagación de errores. Como ya se mencionó, los procedimientos más exactos se basan en la factorización de \mathbf{X} .

Cuando el modelo incluye una constante (columna de unos en \mathbf{X}), se pueden reducir mucho los errores de cancelación de las ecuaciones normales “centrando” las variables, como en el caso de la varianza, usando desviaciones de las medias en lugar de las variables originales. En un modelo de la forma $y = b_0 + \mathbf{x}'\mathbf{b} + e$ se ve que con los parámetros de mínimos cuadrados las medias satisfacen $\bar{y} = b_0 + \bar{\mathbf{x}}'\mathbf{b}$, ya que la primera de las ecuaciones normales asegura que la suma de los residuos es cero: $0 = \mathbf{X}'(\mathbf{y} - \mathbf{X}\mathbf{b}) = \mathbf{X}'\mathbf{e}$. Restando, se tiene el modelo equivalente $y - \bar{y} = (\mathbf{x} - \bar{\mathbf{x}})'\mathbf{b} + e$. Se estima \mathbf{b} con este modelo, y la constante se obtiene de $b_0 = \bar{y} - \bar{\mathbf{x}}'\mathbf{b}$.

PREGUNTAS, EJERCICIOS

1. Estime b en el modelo $y = b + e$. ¿Reconoce el resultado?
2. Verifique que

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Use esto para obtener fórmulas para los dos parámetros en regresión lineal simple (con las variables originales).

3. Obtenga fórmulas para regresión lineal simple usando las variables centradas (desviaciones de las medias).
4. Estime por mínimos cuadrados los parámetros del modelo $y = b_1 + b_2x + b_3x^2$. Trabaje con variables centradas.
5. Investigue la solución de sistemas de ecuaciones lineales y la inversión de matrices por eliminación de Gauss.

B.3 Consideraciones estadísticas

Hemos presentado mínimos cuadrados como un método más o menos razonable y matemáticamente conveniente de “ajustar” funciones a datos observados. Bajo ciertos modelos probabilísticos de los desvíos, el criterio de los mínimos cuadrados puede también justificarse con argumentos de tipo estadístico.

Supóngase primero que las observaciones y_i son generadas de acuerdo a un modelo

$$y_i = \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i ,$$

donde los ε_i son variables aleatorias con media 0 y varianzas desconocidas σ^2 , no correlacionadas entre sí. Es decir,

$$E[\varepsilon_i] = 0, \quad E[\varepsilon_i^2] = \sigma^2, \quad E[\varepsilon_i \varepsilon_j] = 0 \text{ si } i \neq j ,$$

o, en notación matricial,

$$E[\boldsymbol{\varepsilon}] = \mathbf{0}, \quad V[\boldsymbol{\varepsilon}] = \sigma^2 \mathbf{I} ,$$

donde $V[\cdot]$ es la matriz de covarianzas. Los \mathbf{x}_i son vectores conocidos de predictores, y $\boldsymbol{\beta}$ es un vector de parámetros desconocidos que deben estimarse.

Se busca un estimador $\hat{\boldsymbol{\beta}} = \mathbf{b}$ que sea insesgado, o sea $E[\mathbf{b}] = \boldsymbol{\beta}$, y cuya varianzas sea lo más pequeña posible. Limitémosnos además a estimadores que sean funciones lineales de las observaciones, $\mathbf{b} = \mathbf{A}\mathbf{y}$ para alguna matriz \mathbf{A} . Entonces, el teorema de Gauss-Markov dice que el estimador lineal insesgado de varianzas mínima tiene $\mathbf{A} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. Este es el estimador de mínimos cuadrados.

La restricción a estimadores que sean funciones lineales de las observaciones puede parecer algo artificial. Si agregamos el supuesto de que los desvíos tienen una distribución normal, se obtiene el criterio de mínimos cuadrados por otro camino. Sea el modelo, no necesariamente lineal,

$$y_i = f(\mathbf{x}_i, \boldsymbol{\beta}) + \varepsilon_i$$

con los ε_i normales de media 0, varianzas σ^2 , e independientes. Es decir,

$$\mathbf{y} = \mathbf{f}(\mathbf{X}, \boldsymbol{\beta}) + \boldsymbol{\varepsilon} ,$$

$$\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) .$$

La *función de verosimilitud* es la probabilidad de que el modelo genere datos como los observados. El método de estimación de *máxima verosimilitud* (MV) consiste en estimar los parámetros desconocidos como los valores que maximizan esa función. Aparte de ser intuitivamente razonables, los estimadores de MV poseen una serie de propiedades estadísticas deseables, especialmente en muestras grandes.

Aquí la función de verosimilitud es igual a la función de densidad de probabilidad conjunta de los y_i , vista como una función de $\boldsymbol{\beta}$ y σ^2 . Por el supuesto de independencia, la densidad conjunta es el producto de las densidades (normales) de cada y_i :

$$L = f_1(y_1)f_2(y_2) \cdots f_n(y_n) ,$$

con

$$f_i(y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\{y_i - f(\mathbf{x}_i, \boldsymbol{\beta})\}^2}{2\sigma^2}\right] .$$

La verosimilitud es entonces

$$L = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{\sum \varepsilon_i^2}{2\sigma^2}\right] .$$

Claramente, el $\boldsymbol{\beta}$ que maximiza L es aquel que hace mínima la suma $\sum \varepsilon_i^2$. Concluimos que bajo este modelo el estimador de MV de $\boldsymbol{\beta}$ es el estimador de mínimos cuadrados.

Se encuentra también, derivando L con respecto a σ^2 e igualando a cero, que el estimador de MV de σ^2 es $\sum \hat{\varepsilon}_i^2/n = \sum e_i^2/n$, el cuadrado del ECM. El valor esperado de $\sum e_i^2$, para modelos lineales, resulta ser $(n - p)\sigma^2$, de modo que ECM^2 es sesgado. Se acostumbra usar el estimador insesgado ES^2 para la varianza residual σ^2 , y el error standard ES como estimador de σ .

Otro indicador de bondad de ajuste que se usa a menudo, incorrectamente, es el coeficiente de determinación $R^2 = 1 - \text{ECM}^2/S_y^2$, donde $S_y^2 = \sum (y_i - \bar{y})^2/n$ es la varianza de las observaciones y_i ignorando los predictores. Para comparar modelos con los mismos datos el R^2 proporciona la misma información que el ECM. Con distintos conjuntos de datos, sin embargo, un valor de R^2 cercano a uno no implica necesariamente una relación estrecha o un buen modelo. Entre otras cosas, la varianza total depende de cómo ha sido seleccionada la muestra, y a menos que ésta pueda considerarse una muestra aleatoria de una distribución multivariada no refleja una característica de la población.

PREGUNTAS, EJERCICIOS

1. Calcule una regresión lineal entre y y x con los siguientes datos:

x	1	2	3	4	5	6	7	8	9	10
y	1	4	9	16	25	36	49	64	81	100

2. Calcule el R^2 .
3. Grafique los datos y la recta de regresión.

♡ Se observa que para la regresión lineal,

$$E[\mathbf{b}] = (X'X)^{-1}X'E[\mathbf{y}] = (X'X)^{-1}X'X\boldsymbol{\beta} = \boldsymbol{\beta} ,$$

de modo que \mathbf{b} es un estimador insesgado. Lo mismo ocurre con cualquier función lineal de los parámetros \mathbf{y} , en particular, el valor esperado de la predicción $\hat{y}(\mathbf{x}) = \mathbf{x}'\mathbf{b}$ es igual a $y(\mathbf{x}) = \mathbf{x}'\boldsymbol{\beta}$ para cualquier \mathbf{x} .

Como la matriz de covarianza $V[A\mathbf{z}]$ de una transformación lineal es $AV[\mathbf{z}]A'$, se encuentra que

$$V[\mathbf{b}] = \sigma^2(X'X)^{-1} .$$

Si ε es normal, esto y el hecho de que cualquier transformación lineal de un vector normal es normal permite obtener intervalos de confianza y pruebas de hipótesis para funciones lineales de b .

Obviamente, en el mundo real no se puede esperar que estos modelos estadísticos se cumplan con exactitud. Pero se puede esperar que cuanto más nos acerquemos a los supuestos mejores serán los estimadores. Por ejemplo, si se observa que la dispersión de los residuos no es muy uniforme (*heterocedasticidad*) convendría introducir alguna transformación que cambie esta situación. La presencia de autocorrelación (correlación entre mediciones sucesivas) es otro posible problema. En especial, las pruebas de hipótesis están supeditadas a la plausibilidad del modelo estadístico.

♡ **Mínimos cuadrados generalizados** Supóngase que en el modelo lineal la matriz de covarianza de ε es de la forma σ^2W con una matriz conocida $W \neq I$. Manteniendo los otros supuestos se encuentra entonces que tanto el estimador lineal insesgado de varianza mínima como el de MV se obtienen minimizando $\mathbf{e}'W^{-1}\mathbf{e}$. La solución es $\mathbf{b} = (X'W^{-1}X)^{-1}X'W^{-1}\mathbf{y}$.

Una buena introducción a la inferencia estadística se encuentra en el Capítulo 2 de Graybill, del cual hay una traducción entre los materiales del curso. Un texto general con un buen tratamiento de regresión lineal es el de D. Peña Sánchez de Rivera, "Estadística, Modelos y Métodos" (2 Vols.), Alianza Editorial, Madrid, 1992.

Índice General

1	Introducción	1
2	Trozas y productos forestales	4
2.1	Longitud	4
2.2	Diámetros	5
2.3	Cubicación de trozas	9
2.4	Madera arrumada	14
2.5	Medición por peso	17
2.6	Madera aserrada	18
2.7	Cambio de unidades	22
3	Arboles	23
3.1	Diámetros	23
3.2	Alturas	24
3.3	Cubicación de árboles	25
3.4	Análisis fustal	27
3.5	Funciones (tablas) de volumen	29
3.6	Cuocientes y factores de forma, etc.	31
3.7	Funciones (curvas) de ahusamiento	32
3.8	Funciones de corteza	34
4	Rodales	36
4.1	Diámetro, área basal	36
4.2	Alturas	37
4.2.1	Curvas altura-diámetro	37
4.2.2	Altura dominante	38
4.3	Cubicación de rodales/parcelas	39
4.4	Funciones de volumen	41
4.5	Tablas de rodal, distribuciones	42
A	Errores	46
A.1	Límites de error	46
A.2	Cifras significativas	48

B	Regresión	53
B.1	Matrices	53
B.2	El método de mínimos cuadrados	56
B.3	Consideraciones estadísticas	61