



INSTITUTO DEL MAR DEL PERU

# Boletín

ISSN-0378-7699  
VOLUMEN 13(1)

*UN NUEVO METODO DE ANALISIS DE  
DISTRIBUCIONES POLIMODALES  
EN BIOLOGIA PESQUERA*

**Antonio Landa**

Publicación Nº 84 de PROCOPA  
auspiciada por la Deutsche Gesellschaft  
für Technische Zusammenarbeit (GTZ)

CALLAO - PERU, 1989

**UN NUEVO METODO DE ANALISIS DE DISTRIBUCIONES POLIMODALES  
EN BIOLOGIA PESQUERA**

Antonio Landa

Santa Isabel N° 280, Lima 21, Perú

Publicación N° 84 de PROCOPA

**CONTENIDO**

RESUMEN	....	5
ABSTRACT	....	5
EL PROBLEMA	....	6
UN CASO DE LA LITERATURA PARA MOSTRAR EL ESTADO DEL ANALISIS	....	7
DESCRIPCION Y DESARROLLO DEL NUEVO METODO	....	10
<u>El método original de Cassie</u>	....	11
<u>El presente método</u>	....	12
Explicación preliminar	....	12
El modelo propuesto	....	13
a) Un solo grupo	....	13
b) La combinación de grupos	....	14
c) Otras estipulaciones del modelo	....	15
<u>El procesamiento del análisis</u>	....	17
1) Construir una "tabla de datos"	....	17
2) Encontrar los puntos de inflexión	....	17
3) Obtener un primer estimado de los parámetros de grupo	....	20
4) Estimar los traslapes y las proporciones de mezcla	....	21
5) Estimar los parámetros definitivos de los grupos	....	23
6) Reconstruir la distribución original	....	23
7) Evaluar el número de parámetros correctamente estimados	....	25
APLICACION DEL NUEVO METODO A LAS DISTRIBUCIONES ESPECIALMENTE CONSTRUIDAS	....	25
<u>Construcción de cinco distribuciones polimodales</u>	....	25
<u>Aplicación y resultados</u>	....	26
DISCUSION Y CONCLUSIONES	....	26
AGRADECIMIENTO	....	28
REFERENCIAS	....	29
TABLAS (2) TABLAS ANEXO (5)		
FIGURAS (8)		

---

**RESUMEN**

La operación básica de encontrar los grupos de longitud que conforman una distribución de frecuencias con varios modos (usualmente bajo el supuesto de que los grupos constituyen algún tipo de cohortes) debería dar por resultado  $n \times 3 + 1$  estimaciones de otros tantos parámetros -proporción, promedio y desviación estándar de los  $n$  grupos que entran en la combinación y la distribución de las frecuencias combinadas. La aplicación de seis métodos corrientes de análisis de frecuencias multimodales a una muestra de longitudes de peces tomada de la literatura revela las limitaciones de ellos: no pueden estimar correctamente todos los parámetros, no indican cuales son los parámetros correctamente estimados y cuales no lo son, no indican si uno o más de los grupos son estimados correctamente en todos sus parámetros.

El autor ha llegado a la conclusión anterior mediante la aplicación de un nuevo método desarrollado por él mismo en base de muestras especialmente construidas alrededor de parámetros de valor conocido. Se presenta detalladamente el nuevo modelo.

Se sostiene que el presente trabajo revela la existencia de tres clases de parámetros en la combinación de grupos: aparentes, de mezcla y verdaderos.

**ABSTRACT**

The basic operation of finding the length groups which conform a polymodal frequency distribution (usually under the general assumption that the groups are some kind of cohorts) should render  $n \times 3 + 1$  estimates of that many number of parameters - proportion, mean, and standard deviation of the  $n$  combining groups, and the distribution of the combined frequencies.

The application of six current methods of analysis of polymodal frequencies to a sample of fish lengths taken from current literature reveal their limited power: they can not estimate correctly all the parameters involved, do not indicate which parameters are correctly estimated and which not, do not indicate whether one or more groups are correctly estimated in their three parameters.

The author has arrived at the above conclusion applying his own new method developed on basis of samples especially built around known parameter values. An account of the new method is given.

It is claimed that the present work reveals the existence of three kinds of parameteres of combining groups: apparent, mixing, and true ones.

## EL PROBLEMA

En su manual de métodos básicos para la evaluación de stocks de peces (GULLAND, 1983) el Dr. Gulland, al tratar la cuestión de estimar el crecimiento y la edad, escribe: "Si las longitudes de los peces en la población en un momento determinado se plotean como una distribución de frecuencias se verá modos correspondientes a los sucesivos tamaños de los peces nacidos durante la mayor intensidad de la estación de desove y que han crecido de acuerdo a la tasa promedio. Lo típico para peces de longevidad moderada es que el primero o los dos primeros modos aparezcan distintamente y que el siguiente o los dos siguientes sean detectados solo en virtud de la convicción del analista, o, tal vez, mediante cálculos más o menos complejos, que no siempre son mejores; los grupos más viejos se mezclan borrosamente (pag. 85, la traducción y el subrayado son míos). Luego en la página 135 ofrece los ejercicios 4.1: "La tabla 4.1 da la distribución de longitud de una muestra de porgy (*Taeniis tumifrons*) pescados en el Mar de China Oriental". (Ver la Tabla 2 del presente trabajo para la distribución completa). "¿Qué puede Ud. decir acerca de la composición de edades de esta muestra? ¿Cuántos grupos de edad pueden distinguirse claramente? ¿Cuáles son sus longitudes modales? ¿Cuántos individuos hay en cada grupo? ¿Puede identificarse más fácilmente los últimos modos mediante la determinación completa de la distribución de frecuencias de los primeros grupos?" En la página 207 se da respuesta a tres de las cuatro preguntas: "Se puede distinguir claramente dos grupos de edad (con modos de 10.6 y 14.9 cm) y otro menos claramente (en 19.4 cm). Los individuos de los grupos son aproximadamente 5710, 4600 y 2040".

Nótese primero que no hay modo de saber si las respuestas dadas son correctas o no y luego que la cuarta pregunta no tiene respuesta, dejando al lector la aplicación de algunos cálculos para, aparentemente, descubrir más grupos y para determinar completamente las distribuciones de los primeros grupos. Es el caso que esos cálculos apuntan a establecer completamente las distribuciones de todos los grupos detectados, esto es, a estimar sus promedios y desviaciones estándar; estos estimados, junto con los de las proporciones, permiten reconstruir cada uno de los grupos y luego su mezcla es contrastable con la distribución combinada de la muestra. De este modo, el uso de cálculos es una útil y necesaria alternativa a la técnica de la simple inspección visual puesto que los resultados pueden ser sometidos a prueba.

El presente trabajo se ocupa del problema de en que medida son contrastables los resultados de los análisis de distribuciones multimodales de frecuencias; concretamente, encontrar en que medida el análisis estima correctamente el tamaño, el promedio y la desviación estándar de los grupos componentes - si tal cosa es posible - y la distribución de frecuencias cuando los grupos son combinados, lo cual es siempre posible. Desde otro punto de

## Análisis, distribuciones polimodales

---

vista, este ejercicio puede considerarse como una primera aproximación a la cuestión considerarse la sentencia del Dr. Gulland subrayada más arriba.

A modo de justificar el presente esfuerzo, quiero llamar la atención hacia cierto paralelismo, *mutatis mutandis*, con el contenido de "The Forgotten Requirement for Age Validation in Fishing Biology" (BEAMISH and MCFARLANE, 1983) donde se señala que entre 1907 y 1980 casi la mitad de los estudios que incluían estimaciones de edad de peces no mencionan la validación de la edad ni trataron de validar la técnica usada, y, después de revisar algunas de las consecuencias negativas de tan poco exigente procedimiento, con toda justicia declaran que cuando no es posible validar todas las clases de edad en una población mediante estudios de marcación y recuperación de marcas la determinación de edad debe ser hecha por varios métodos, y debe considerarse la posibilidad de errores en los estimados.

En la primera parte aplicaré 6 métodos corrientes incluyendo mi propio nuevo al ejemplo del manual de Gulland y compararé los resultados remarcando como, por el momento, la forma en que el nuevo método fue desarrollado ofrece un camino para llegar a un análisis significativo. En la segunda parte describiré el nuevo método y en la tercera el procedimiento para aplicar el nuevo método será seguido paso a paso. Una cuarta parte demuestra como se desarrolló el método al ser aplicado a distribuciones especialmente construídas.

### UN CASO DE LA LITERATURA PARA MOSTRAR EL ESTADO DEL ANALISIS

La distribución de frecuencia de longitudes del Ejercicio 4.1 de Gulland 1983 se da completa en la primera columna de la Tabla 2 de Tanaka 1962. Está copiada en la columna 2 de mi Tabla 2; el contenido de las otras columnas se explica en la misma tabla. Todas las frecuencias en la tabla, excepto las de la muestra, se obtuvieron aplicando la fórmula

$$F_i = \sum (N_j / s_j \sqrt{2\pi}) \exp - (.5(i-.5 - \bar{y}/s_j))^2 \quad (1)$$

donde  $F_i$  es la suma de las frecuencias en el límite superior de la clase  $i$ ;  $N_j$  el tamaño del grupo componente  $j$ ;  $\bar{y}_j$  el promedio del grupo  $j$ ;  $s_j$  la desviación estándar del grupo  $j$ . A su vez,  $N_j$ ,  $\bar{y}_j$  y  $s_j$  se obtuvieron de la tabla 3 de Tanaka para los métodos de Tanaka, Cassie y Buchanan-Wollaston; para los otros tres métodos estos parámetros fueron calculados según los propios métodos. Todos ellos se muestran en la Tabla 1. En primer lugar se ve que, excepto el caso del método NORMSEP en el cual el número de grupos es una conjetura del que lo utiliza, los métodos coinciden en

estimar en 5 el número de componentes, es decir 2 más que el número supuesto a base de la observación directa, tal como era de esperarse. Luego se ve que, aunque bastante similar en general, métodos diferentes dan diferentes estimados de los parámetros de los grupos componentes, proporción, promedio y desviación estándar.

Los únicos datos proporcionados por la muestra son su tamaño y las frecuencias a cada clase; en cambio, cualquiera de los métodos de análisis proporciona, además, el número de grupos así como su tamaño, promedio y desviación estándar cuyas magnitudes no pueden ser evaluadas en cuanto a su corrección porque no están explícitamente contenidas en la muestra. Estos parámetros, sin embargo, permiten la reconstrucción de las frecuencias combinadas las cuales sí pueden ser comparadas con las de la muestra mediante una prueba de Chi cuadrado o, preferentemente, una prueba G. Los resultados de la reconstrucción y prueba se muestran en la Tabla 2 (que contiene también la explicación de sus columnas). Se señala especialmente (i) que para los resultados inmediatos ninguno de los métodos reconstruye las frecuencias de modo que su total sea igual ( $\pm 0.5$ ) al de la muestra por lo cual los valores de G encontrados prácticamente no tienen significado, justificándose entonces la suavización, mediante promedios móviles de 5, de todas las frecuencias incluyendo las de la muestra (ii) los valores de la prueba G para los resultados inmediatos van de 8.3 a 612.85, a los que se añadió los valores de G correspondientes a la diferencia entre el total de la muestra y el total del análisis (iii) los valores finales de G después de la suavización van de 8.66 a 613.21, sus probabilidades (con 2 g.l., última línea dentro de la tabla) son mayores que 0.5 -indicando que el ajuste entre las frecuencias resultantes del análisis y las de la muestra es estadísticamente aceptable- solo en el caso del método del autor, de Tanaka y de Buchanan (iv) la prueba G usada es

$$G = 2 \sum F_o \log F_o / F_e$$

donde  $F_o$  es la frecuencia observada y  $F_e$  la esperada (esto es, esperada de acuerdo a la teoría o modelo del método usado); los valores G son muy cercanos a los de una estadística Chi cuadrado para  $n-1$  grados de libertad, donde  $n$  es el número de clases comparadas; la G tiene ventaja sobre la de Chi cuadrado porque es aplicable a cualquier clase incluso aquellas cuya frecuencia es menor de 5 obviando así la operación subjetiva de decidir que clases con frecuencias menores de 5 tienen que ser juntadas. (Para otros aspectos de G ver SOKAL and ROHLF, 1981). Finalmente (v) la prueba G aplicada a los totales, probablemente una novedad, se basa en el hecho que dicha prueba es posible aún solo entre dos clases, en el presente caso una de las clases es ficticia, compara 1 en la muestra con 1 en el resultado del análisis -lo

cual por supuesto da un  $G = 0$  - en la otra clase se compara 1 en la muestra con la diferencia entre totales menos 0.5 más 1. El valor final de  $G$  se ilustra a continuación para el caso del método de Bhattacharya.

- $G$  de la comparación de frecuencias (después del suavizamiento) = 70.76 con 26 g.l.
- $G$  de la diferencia entre totales (antes del suavizamiento) = 3.58 con 1 g.l.
- $G$  final = 70.76 + 3.58 = 74.34 con 26-1 = 25 g.l.

Considerando las Tablas 1 y 2, los resultados obtenidos con el método del autor serían aceptables ya que ese modelo produjo una distribución de longitudes que se ajusta muy bien a la muestra, y tal vez también los resultados de los métodos de Tanaka y de Buchanan-Wollaston puesto que la probabilidad de su ajuste está por encima de 0.5.

Hasta aquí todo está bien puesto que, siendo el total de variables y su distribución los únicos datos directamente ofrecidos por la muestra, el criterio de un buen ajuste de las frecuencias es ciertamente importante. Pero no hay modo inmediato para medir la exactitud de los estimados de los parámetros de los grupos -su tamaño, promedio y desviación estándar- encontrados por los diferentes análisis. Concretamente, la mayoría de los métodos corrientes usados han sido desarrollados más o menos de la siguiente manera: después de una breve declaración tanto de su base racional como de los algoritmos propuestos el método es aplicado a alguna distribución resultante de un muestreo, esto es, a una distribución multimodal de la cual solo se conoce el total de sus frecuencias y su distribución mientras que las características de los grupos componentes (el número de ellos, su tipo de distribución -normal o no normal- su tamaño, promedio y desviación estándar) no son conocidas, de modo que, estrictamente, el método es susceptible de ser probado solo en relación al total y a la distribución de frecuencias.

Los procedimientos empleados en el nuevo método del autor han sido desarrollados usando distribuciones polimodales obtenidas sumando las frecuencias de grupos normales construidos en base de promedios, desviaciones estándar y tamaños arbitrariamente escogidos de modo que cuando se aplicó el método a esas distribuciones fue posible medir su grado de efectividad con respecto a las frecuencias y también a los parámetros de los grupos.

La Tabla 4 del Anexo es un sumario de los resultados de aplicar el nuevo método a 5 distribuciones de frecuencia polimodales especialmente construidas, o lo que es lo mismo, a 5 combinaciones, cada una de 4 grupos normales o cuasi normales, también especialmente construidos. Los rasgos importantes son (i) la aplicación del método a cada una de las combinaciones resultó

en distribuciones que se ajustaron a las originales con probabilidades mayores de 0.995; sin embargo, (ii) los valores de G fueron de 2.55 a 8.03 mientras el número de parámetros estadísticamente bien estimados varió también, paralela pero inversamente a los valores G (iii) los valores de G fueron considerablemente menores que los requeridos para la probabilidad de 0.995 (alrededor de 25 para los 26 g.l. de los ejemplos contruïdos).

Considerando que los ejemplos contruïdos contienen 4 grupos y consecuentemente hubieron 12 valores de parámetros para estimar en cada combinación, la Tabla 5 del Anexo sugiere de inmediato una relación como la que tentativamente se ha trazado en la Fig.1 entre el valor de  $G'$  ( $G' = G \cdot 3 \cdot n/10$  -donde n es el número de grupos) y la proporción de parámetros de grupo correctamente estimados. En el caso del ejemplo de porgy, el uso de esta relación indicaría que el nuevo método puede estimar correctamente los valores de 3-4 de los 15 parámetros y cero los otros métodos.

#### Resumiendo

Método	G	G'	Num. de Par.	Corr. Est.
Nuevo	8.66	12.99	3-4	(22.25%)
Buchanan	20.47	30.70	0	0
Tanaka	19.87	29.80	0	0

nótese que no se ha calculado los valores de  $G'$  para los métodos cuya distribución no tuvo probabilidad aceptable.

#### DESCRIPCION Y DESARROLLO DEL NUEVO METODO

El nuevo método aquí presentado se basa en el mismo principio del método gráfico de Cassie (CASSIE, 1945): el conocimiento de que cuando el ploteo de los porcentos acumulados de una distribución en papel de probabilidad normal toma la forma de una curva ondulada, es posible analizar la distribución polimodal que representa y encontrar las distribuciones unimodales o componentes simples normales que entran en su formación, cada uno de los cuales será mostrado mediante el análisis como una línea oblicua recta en la cual se puede estimar el promedio y desviación estándar del componente mientras que la distancia porcentual entre los puntos de inflexión de la curva ondulada representa la proporción que cada grupo contribuye a la combinación.



Se diferencia del método de Cassie en que no requiere procedimientos gráficos y en que la estimación del promedio y desviación estándar de cada grupo es hecha directa y simultáneamente en contraste con la necesaria previa estimación de cada punto del componente. Estas diferencias, que acarrearán un tratamiento completamente objetivo que a su vez conduce a una completa computarización, dependen de conceptos no compartidos con el método de Cassie.

Se ha hecho notar al autor que una exposición como la que antecede, o similar, podría ser casi ininteligible para muchos biólogos y que una explicación completa, incluyendo algo acerca de puntos de inflexión, sería lo apropiado. Así pues, antes de la descripción del nuevo método va una breve del método de Cassie.

#### El método original de Cassie

Consiste en plotear la distribución porcentual acumulada en un papel de probabilidad normal tal como muestran los círculos de la Fig. 2 (reproducida de la Fig. 1 de Cassie 1954). Los segmentos verticales en la curva son los puntos de inflexión localizados al ojo, esto es, los puntos donde la curva, considerada como que va de izquierda a derecha, cambia de cóncava hacia arriba a cóncava hacia abajo; hay cuatro de estos puntos, al 10, 66, 91, y 98.6 por ciento respectivamente, indicando la existencia de cinco componentes normales cuyos tamaños son iguales al 10%, 66-10=56%, 91-66=25%, 98.6-91=7.6% y 100-98.6=1.4% del tamaño total de la muestra, respectivamente. El cociente de 100 dividido por el porcentaje de cada uno de estos componentes es un factor que multiplicado por el porcentaje correspondiente a un círculo de la curva dentro del rango del componente da un nuevo porcentaje, marcado por un punto negro en la figura; así los cuatro primeros círculos originan los cuatro puntos a través de los cuales se ha trazado al ojo la línea recta 0 que representa el primer componente; al pasar al primer elemento del segundo componente- quinto círculo en la ilustración- el tamaño del primer componente se sustrae de él antes de multiplicarlo por el factor del segundo componente; entonces el primer punto para trazar la línea 1, del segundo grupo, es

$$1.6 = (10.9 - 10) \times 100/56$$

este procedimiento se continúa hasta que en el 15avo círculo (en el caso ilustrado) el punto resultante, encerrado en un círculo en la reproducción de la figura original, aparece situado a la derecha de la línea oblicua indicando que hay traslape con el siguiente componente, se requiere entonces una corrección que será extensiva a los primeros puntos del tercer componente. (Nótese que al tratar con los círculos 4to y 15avo la línea respectiva todavía no ha sido trazada de modo que la decisión que el 4to punto está en la línea así como la de que el

15avo está a la derecha (o abajo) de su línea son dejadas enteramente al juicio del analista). La corrección se hace como sigue: trazar una línea paralela al eje de las X desde el 15avo círculo (64.8%) hasta la línea recta 1 y de ahí una vertical hasta encontrar el punto de 97% de la escala; en base a estas cifras en el método se declara que "el porcentaje total representado por el segundo", línea 1, es

$$54.3\% = 97 \times 56/100$$

a esto se le añade el 10% del primer componente, y el primer punto del tercer componente, línea 2, es entonces

$$2.0\% = (64.8 - 64.3) \times 100/25$$

El procedimiento correctivo se aplica círculo por círculo hasta que en el 18avo (en el ejemplo) la línea vertical encuentra el 99.9% de la escala; a partir del 19avo círculo la corrección ya no es aplicada hasta que, por supuesto, un punto del tercer componente cae a la derecha de la línea 2 (aunque la línea no ha sido aún trazada, como antes) y hay que encontrar una nueva corrección en la forma descrita. Cuando se termina el proceso las líneas rectas resultantes se usan para estimar el promedio de cada grupo -leyendo en la escala de las longitudes el nivel al cual la línea oblicua que representa el grupo cruza la línea vertical del 50%- y la desviación estándar como la diferencia absoluta entre la lectura del promedio y aquella correspondiente al nivel de cruce de la línea oblicua con la vertical del 15.87% o la del 84.13%.

Los comentarios inmediatos y más pertinentes al método se derivan de su naturaleza gráfica, pues que ella implica varias apreciaciones subjetivas: (i) la localización de los puntos de inflexión al ojo (ii) asignarles un valor en la escala de porcentajes acumulados mediante una proyección visual (iii) la localización al ojo de los puntos que se usan para (iv) estimar también al ojo las líneas oblicuas de cada grupo (v) la apreciación de cual es el primer punto de traslape en cada línea; todas estas apreciaciones subjetivas contribuyen a una acumulación no sistemática de errores de estimación.

### El presente método

#### Explicación preliminar

El papel de probabilidad y todos los procedimientos gráficos han sido eliminados en la aplicación del método pero el enfoque gráfico se usa libremente para desarrollar los algoritmos usados este cambio han seguido otros, pero antes de describir el método es conveniente unas cuantas palabras acerca de los probits. El ploteo de las frecuencias porcentuales acumuladas de una sola

distribución normal en papel de probabilidad normal da por resultado una línea recta ablicua, excepto en los extremos, en contraste con la curva sigmoidea que el mismo ploteo produce en el papel cuadrículado regular. Este resultado es debido al espaciamiento característico de las abscisas como se aprecia en las Figs. 2 y 3. SOKAL and ROHLF (1981) explican elegantemente como es que resulta ese espaciamiento mediante su Fig. 6.5-reproducida aquí en nuestra Fig. 3- y en las páginas 544 y 545 también explican que si el eje de las abscisas de la Fig. 3 fuese colocado paralelamente a una escala de unidades normales estandarizadas, esto es, los valores de la escala usual menos el promedio y divididos por la desviación estándar, el punto del 50% coincidirá con cero divisiones estándar y el de 84.13% con una desviación estándar. Ahora, lo que una unidad normal estandarizada es a una frecuencia normal una desviación normal equivalente (NED) es a una frecuencia normal acumulada. Los probits son NEDs + 5.0; entonces, abandonando a Sokal y Rohlf, en una escala de probits en la abscisa el promedio coincidirá con el probit 5.0 y el porcentaje acumulado hasta el promedio más una desviación estándar al probit 6, etc. Los probits pueden plotearse en papel gráfico regular; probit 1.2809 corresponde a 0.01 porcentaje acumulado y probit 8.7191 al 99.99%, y por supuesto probit 5 al 50%, el ploteo resultante puede manejarse de acuerdo a la geometría plana.

#### El modelo propuesto

En su aplicación práctica el método no requiere ningún ploteo pero su base conceptual, sus principios y asunciones, se describen mejor con la ayuda del modelo gráfico presentado a continuación.

#### a) Un solo grupo

1. Si para una sola distribución normal los porcentos acumulados a cada clase sucesiva de la escala de medida se plotean en papel de probabilidad producen una sola línea recta oblicua desde la izquierda inferior hasta la derecha superior del papel; en esta línea el promedio se lee al nivel de su cruce con la línea del 50% y la desviación estándar se estima sustrayendo del promedio el valor leído al nivel del cruce con la línea del 15.87% o sustrayendo el promedio de la lectura al nivel del cruce con el 84.13%. Ya que una sola distribución abarca del cero al cien porcentaje resulta que el promedio se encuentra al nivel del porcentaje acumulado

$$50 = .100 \times 0.5 \dots\dots\dots (2)$$

y el otro punto para calcular la desviación estándar en el por ciento

$$15.87 = 100 \times 0.1587 \dots\dots\dots (3)$$

o en el por ciento

$$84.13 = 100 \times 0.8413 \dots\dots\dots (4)$$

b) La combinación de grupos

2. Las Figs. 4a y 4b ilustran las partes esenciales del método, la ordenada es la escala de medidas, una escala de las abscisas es la frecuencia absoluta y la otra los probits. Las curvas a, b y c están formadas por las frecuencias de los dos grupos componentes y por la parte en que ellas se mezclan. Las curvas d están formadas por los probits correspondientes a los porcentos acumulados de las frecuencias combinadas. Las mezclas de 4a y 4b difieren marcadamente en lo que se refiere a la separación de los promedios, 10 vs. 3 unidades de medida; la razón entre los tamaños completos de los grupos, 4 a 1 y 1.5 a 1; y la forma y extensión de la porción c. Las curvas d también difieren pero son esencialmente similares porque ambas tienen 3 porciones: una central -un segmento de línea recta generado por la porción de mezcla c y dos terminales de curvatura generados por la porción de los grupos que no participan en la mezcla. Cada porción puede no estar netamente diferenciada de las otras. Pero el razgo importante es que la existencia de las dos curvaturas opuestas -una vonvexa hacia abajo y la otra convexa hacia arriba- determina automáticamente la existencia de un punto de inflexión (p.i.) entre ellas situado en la porción recta intermedia y que corresponde al punto donde los grupos se inter cruzan.
3. La Fig. 5 muestra la base geométrica para identificar el p.i. de interés, en la columna más angosta bordeada a cada lado por columnas más anchas, (un p.i. entre una porción convexa hacia arriba y otra convexa hacia abajo se mostraría en la columna más ancha bordeada por columnas más angostas), cuando el ancho de cada columna es la distancia entre probits de clases sucesivas.
4. En las ecuaciones (2), (3), y (4) sustituyamos 100 por la proporción aparente (PA) de un grupo para obtener

$$\begin{aligned} \text{ACC} &= \text{PA} \times 0.5 \dots\dots\dots (2') \\ \text{SDCC} &= \text{PA} \times 0.1587 \dots\dots\dots (3') \\ (\text{SDCC})' &= \text{PA} \times 0.8413 \dots\dots\dots (4') \end{aligned}$$

para los puntos de la curva de porcentos acumulados que proyectados en la escala de medidas estiman el promedio y los puntos situados una desviación estándar por debajo y otra por encima del promedio, respectivamente, del primer grupo; estas acuaciones también se aplican a los otros grupos si a la derecha de las igualdades se añade el valor de los p.i. anteriores.

5. Estos estimados son necesarios para estimar el valor del traslape, que a su turno se usa para modificar las proporciones aparentes y obtener las de mezcla y llegar finalmente a una estimación definitiva de los parámetros.

c) Otras estipulaciones del modelo

6. La Fig. 5 describe el modelo gráfico para encontrar el intervalo de clase donde ocurre un p.i., y dentro del intervalo el punto mismo. El ancho de cada columna es la diferencia entre el probit del límite superior de una clase y el de su límite inferior. Estos anchos disminuyen de izquierda a derecha hasta llegar a un mínimo y luego aumentar; la columna más angosta, rayada oblicuamente en la figura, está limitada arriba por la porción de la curva con inclinación más próxima a la vertical, esto es, casi no contribuye a la porción cóncava hacia arriba ni tampoco a la cóncava hacia abajo; por esta razón la columna más angosta es donde se encuentra el p.i. Ya que al lado superior de la columna rayada es practicamente un segmento recto, no tiene significado práctico repetir el análisis dentro de ella y por tanto parece correcto asumir que el p.i. se sitúa en el punto medio de la clase y tiene el correspondiente probit y valor del por ciento acumulado.
7. Si  $n$  grupos componentes se acercan mucho a la normalidad no es muy probable que al combinarse produzcan más de  $n + 1$  p.i. pero podrían producirse algunos p.i. suplementarios o falsos en la vecindad de los verdaderos si los grupos difieren notablemente de la normalidad, especialmente hacia los extremos; en tales casos algunos de los p.i. aparecerán muy cerca unos de otros. De lo que antecede resulta claro que para detectar un p.i. debe considerarse tres intervalos de clase sucesivos, de los cuales los de los extremos pertenecerán a grupos diferentes; esta es la situación más simple posible y para completarla es necesario aumentar un p.i. en el extremo inferior del primer grupo y otro en el extremo superior del segundo, como

resultado tendremos como situación mínima lo siguiente

$$3 \text{ clases/ } 3 \text{ p.i.} = 1 \text{ clase por p.i.} \dots (5)$$

si interviene un mayor número de clases esta razón permanecerá la misma o será mayor de modo que se puede postular que para el modelo "el número de clases del grupo más pequeño en la combinación es por lo menos igual a la razón entre el número total de clases y el número de p.i. encontrado más dos puntos" o, en otras palabras, "si en una combinación se encuentran grupos con un número de clases inferior al número promedio de clases por p.i. más dos puntos, esos grupos no son verdaderos sino que resultan de falsos p.i." Sin embargo; cuando el primero (o el último) p.i. verdadero separa una cola de la distribución larga pero delgada (con frecuencias muy pequeñas) como si fuera el primer (ó último) grupo, tal grupo debe ser suprimido si el número de individuos por clase con frecuencia positiva es menor que 1.00.

8. El número de traslapes es  $j \times 2 = n \times 2 - 2$  donde  $j$  es el número de p.i. y  $n$  el número de grupos; cada p.i. origina 2 áreas de traslape, una bajo el flanco descendente del grupo situado a la izquierda del p.i. y la otra correspondiente a la parte ascendente del grupo situado a la derecha, de modo que, excepto por el primero y el último, que traslapan una sola vez, cada grupo traslapa dos veces.
9. Los traslapes son usados para modificar las proporciones aparentes: a la proporción aparente se le añade las dos partes del mismo escondidas (o traslapadas) por los grupos vecinos y se le sustrae las dos partes, una por vecino, de los grupos vecinos que son traslapados por el propio grupo.
10. Para el modelo el traslape es válido cuando ocurre entre el flanco descendente o derecho de un grupo y el ascendente o izquierdo del siguiente grupo, tal vez también el que ocurre con el flanco ascendente del grupo siguiente al siguiente, pero el traslape del flanco de un grupo con los dos flancos de un grupo vecino no es permitido; esto es, si llamamos LE al terminal izquierdo y RE al derecho, cuando

$$LE_j > LE_{j+1} \quad \text{o} \quad RE_j > RE_{j+1}$$

donde el suscrito es el número del grupo, el análisis no puede continuar.

11. Al construir los algoritmos para desarrollar el modelo la unidad de la escala de medida (la escala de referencia) es igual a la unidad de la escala de probits. Todas las interpolaciones son lineales.

### El procedimiento del análisis

La declaración de cada operación o bloque de operaciones y de los "inputs" y "outputs" es seguida, cuando sea necesario, de las fórmulas usadas y de notas auxiliares que sirvan de guía para la formulación de algoritmos para la computación automática.

- 1) Construir una "tabla de datos" de 5 columnas

Col. 1; escala de referencia de las medidas desde 1 hasta NC (=número de clases)  
col. 2; frecuencia absoluta en cada clase  $F_i$   
col. 3; porcentaje acumulado de las frecuencias a cada clase,  $CP_i$ , con 4 decimales  
col. 4; probits correspondientes a cada porcentaje acumulado,  $P_i$ , con 4 decimales  
col. 5; diferencia de probits de la clase  $i$  menos la clase  $i-1$ ,  $DP_i$ , con 4 decimales, escrita al nivel de la clase  $i$ .

Notas: (a) NC,  $F_i$ , y N (=total de las frecuencias) son los inputs de la muestra; para uso posterior también debe entrarse el intervalo de clase de la muestra, CI, y el valor del límite superior de la primera clase en las unidades de medida usadas en la muestra,  $U_1$  (b) La columna 1 es la columna de la variable de cuenta, I. (c) Las frecuencias de las clases se entienden como pertenecientes al límite superior de la clase. (d) La conversión de porcentajes acumulados a probits se hace siguiendo a ABRAMOWITZ and STEGUN (1968).

- 2) Encontrar los puntos de inflexión (p.i.), el número de grupos componentes, y la proporción aparente de cada uno de ellos.

(1) Identificar los intervalos de clase que contienen p.i. identificando en la columna 5 de la tabla de datos aquellas

diferencias de probits,  $DP_i$ , tales que

$$DP_{i-2} > DP_i \leq DP_{i+1} \dots \dots \dots (7)$$

la columna 1 de los valores de los límites superiores de las clases donde se encuentran los p.i.,  $UI_i$ ; de la nota (b) se desprende que  $UI = 1$  y que  $UI_{i-1}$  es un límite inferior. El resultado de (i) es una serie de  $GG+2$  límites superiores de clases donde ocurre p.i.; pero el 2 adicionado es para las clases cero y NC (la última clase o número total de clases):

$$0, UI_1, UI_2, \dots, UI_{GG}, NC \dots \dots (8)$$

- (ii) Identificar y eliminar los grupos falsos y por consiguiente los falsos p.i. entre los  $GG+1$ , ( $=G$ ), grupos de la serie (8) usando el criterio del mínimo número de clases por grupo en la fórmula (5) y recordando que un grupo siempre está limitado por 2 p.i. (y por tanto 2 clases que contienen p.i.); esto es, cada vez que

$$UI_{i+2} - UI_i < NC/(GG+2) \dots \dots (9)$$

eliminar  $UI$  y añadir el número de clases entre  $UI_{i+1}$  y  $UI_i$  (menor que el requerido) al grupo siguiente hasta que en (9) la diferencia en el lado izquierdo más las adiciones sea igual o mayor que la diferencia en el lado derecho.

- Notas: (e) Eliminar aquí significa  $UI_i = 0$  y una reducción de 1 en el contador de p.i., ( $GG$ ). (f). El proceso comienza en la parte más baja de la tabla (clase NC), lo cual puede necesitar el uso de pasos negativos en el contador; el proceso de adicionar clases más pequeñas que lo requerido debe ser performedo a cada paso.

Nótese que el proceso (ii) puede también ser comenzado por la clase 0 tal vez con diferentes resultados diferentes resultados. El orden aquí adoptado comenzado por la parte baja de la tabla



o de la derecha a la izquierda en la combinación, está de acuerdo con el hecho de que un e.i. corresponde al cruce de dos grupos, el de la izquierda puede existir por sí solo pero la existencia del de la derecha o segundo presupone el primero y por tanto determina también la existencia de un p.i. De acuerdo con esto nótese también que (16) más adelante tiene aplicación solo cuando el proceso (ii) ha sido comenzado en la clase 0 lo cual a su vez se justificará solo como consecuencia de que (15) haya ocurrido en el curso normal de (ii), esto es, cuando se comience de NC. (La ocurrencia de (15) significa que la porción después del último p.i. no es un grupo sino unos cuantos individuos).

- (iii) Estimar el valor de p.i. en porcentaje acumulado y la proporción aparente, o tamaño en porcentaje, de cada grupo usando los porcentajes acumulados en los límites superior e inferior de las clases que contienen p.i.; esto es, para p.i.

$$IP_J = (CP(UI(J)) + CP(UI(J)-1))/2 \dots (10)$$

y para la proporción aparente

$$AP(J) = IP(J) - IP(J-1) \dots (11)$$

donde el contador J va de 1 a G

- (iv) Inspeccionar el primero y último grupos para ver si son completos

- a. para el primer grupo, si

$$AP(1) \times N / 100 / (UI(1) + 1 - ZE) < 1.00 \dots (12)$$

eliminarlo, poniendo

$$G = G - 1 \dots (13)$$

$$AP(1) = 0 \dots (14)$$

b. para el último grupo, si

$$AP(G) \times N / 100 / (NC - UI(G) + 1 - ZE) < 1.00 \quad (15)$$

eliminarlo y poner

$$AP(G) = 0 \text{ y } G = G - 1 \quad \dots\dots\dots (16)$$

Notas: (f) ZE es el número de clases con frecuencia cero. Sugerión para encontrar ZE: Z=0, para I=1 hasta UI(1), si F(I)=0 entonces sea ZE=ZE+1 para el caso a. y las correspondientes variantes para el caso b.

3. Obtener un primer estimado de los parámetros de grupo

(1) Promedio y desviación estándar de cada grupo en términos de puntos en la curva de porcentos acumulados aplicando

$$M(J) = AP(J) \times 0.5 + IP(J-1) \quad \dots (17)$$

para los promedios y

$$ST(J) = AP(J) \times 0.1587 + IP(J-1) \quad \dots (18)$$

$$TS(J) = AP(J) \times 0.8413 + IP(J-1) \quad \dots (19)$$

para los puntos situados a 1 desviación estándar debajo y por encima del punto del promedio, respectivamente

Notas: (g) Los términos del lado izquierdo de (17), (18) y (19) se convierten en porcentos acumulados a valores de la escala referencial de medida mediante interpolación lineal en la tabla de datos sin cambiar nombres, entonces se estima dos desviaciones estándar

$$ST = M(J) - ST(J) \quad \dots\dots\dots (18')$$

$$TS = TS(J) - M(J) \quad \dots\dots\dots (19')$$

y finalmente la definitiva

$$SD(J) = (ST(J) + TS(J)) / 2 \quad \dots\dots (20)$$

(ii) Encontrar los puntos terminales de cada grupo aplicando

$$UE(J) = M(J) + SD(J) \times 3.7191 \dots (21)$$

$$LE(J) = M(J) - SD(J) \times 3.7191 \dots (22)$$

para los terminales superiores e inferiores, respectivamente.

**Notas:** (h) Las fórmulas (21) y (22) se originan (ver Fig. 8) de la similaridad de los triángulos rectángulos en los que  $a$ ,  $b$  y  $a'$  tienen valores conocidos que permiten el cálculo de  $b'$  (o  $b''$ ) como el producto de la desviación estándar por la mitad del ancho de la escala de probits.

(iii) Inspeccionar los valores de los puntos terminales y detener el análisis si

$$UE(J+1) - UE(J) < -0.5 \dots \dots \dots (23)$$

o

$$LE(J+1) - LE(J) < -0.5 \dots \dots \dots (23')$$

4. Estimar los traslapes y las proporciones de mezcla.

(1) Estimar los traslapes en términos de probits

En las Figs. 6 y 7 las líneas gruesas que terminan en los márgenes son traslapes y su numeración corresponde a una combinación de 4 grupos pero el mismo esquema puede ser aplicado a combinaciones de cualquier número de grupos.

El cálculo de los traslapes en términos de probits se realiza mediante

$$OV = (OP-M)/SD + 5 \dots \dots \dots (24)$$

$$OV = (OP-M)/SD - 5 \dots \dots \dots (25)$$

**Notas:** (1)  $(OP-M)/SD$  en (24) es  $a'$  de la Fig 7 y en (25) es  $a''$  de la misma figura;  $a'$  y  $a''$  se encuentran por la similaridad

de triángulos rectángulos. (j) Si la línea oblicua de margen a margen representa digamos el grupo 2 entonces M y SD son el promedio y la desviación estándar del grupo 2, tanto en (24) como en (25) pero OP es el 3er p.i. en (24) y el 2do p.i. en (25), en términos de la escala referencial de medidas. (Ver el párrafo numerado 9 en la sección del Modelo usado).

(ii) Convertir los traslapes de probits a porcentos acumulados por interpolación lineal en la tabla de datos; el OV de (24) se sustrae de 100 para obtener el segmento grueso de la Fig. 8 ya que este corresponde a un flanco del lado derecho. Estos porcentos se multiplican por la proporción del grupo dividida por 100 para obtener el traslape en términos

de porcentaje parcial acumulado en base al tamaño total de la combinación, para ser usado en el paso siguiente

(iii) Estimar las proporciones de mezcla

De acuerdo con el modelo (ver arriba) y siguiendo la Fig. 6 el esquema para sumar y restar traslapes a fin de obtener proporciones de mezcla, para una combinación de 4 grupos pero aplicable a combinaciones de cualquier número de grupos, es

traslapes

grupo 1;	+4	-1				
2;	+1	+5	-4	-2		
3;	+2	+6	-5	-3	.....	(26)
4;	+3	-6				

Notas: (k) La conformación de algoritmos para el cálculo automatizado de (24), (25) y (26) no es simple y directa. Se sugiere

$$OV(J) = (OP(J) - M(J+1)) / SD(J+1) \dots (27)$$

para los traslapes, OV, y donde OP es p.i. en términos de la escala referencial, M el promedio y SD la desviación estándar; J va de 1 a G+2 y G

es el número de grupos. Sin embargo, antes de aplicar (27) es necesario replicar, con otro índice,  $OP(J)$ ,  $M(J)$  y  $SD(J)$ , para  $J = 1$  hasta  $G-1$ , esto es,

$$\begin{aligned} OP(J) &= OP(G+H) \\ M(J) &= M(G+H) \dots\dots\dots (28) \\ SD(J) &= SD(G+H) \end{aligned}$$

donde  $H$  va de 0 a  $G/2$

Para añadir los traslapes a las proporciones aparentes a fin de obtener las proporciones de mezcla,  $MP$ , las sugerencias son para todos los grupos con excepción del primero y del último, esto es,  $J$  va de 2 a  $G-1$ ,

$$MP(J) = AP(J) + OV(G+J-1) - OV(J) + OV(J-1) - OV(J-2) \dots\dots\dots (29)$$

para el primer grupo

$$MP(1) = AP(1) + OV(G+J-1) - OV(J) \dots (30)$$

y para el último grupo

$$MP(G) = AP(G) + OV(G-1) - OV(G+2) \dots (31)$$

5. Estimar los parámetros definitivos de los grupos

Las proporciones de mezcla están dadas por (29), (30) y (31). Para ellas debe haber nuevos p.i. que se calculan como sigue

Proporciones de mezcla de gp.	1	2	3	4
p.i. correspondientes	0	(1)	(2+1)	(3+2+1)

Para los promedios, desviaciones estándar y puntos terminales repetir los pasos en 3) arriba y luego pasar a 6) si es que ningún terminal  $n$  es mayor que su similar  $n+1$  ( $n$  es el número del grupo).

6. Reconstruir la distribución original en base de los parámetros de grupo encontrados en el análisis y compararla con la de la muestra.

- (i) Usando las proporciones de mezcla encontradas en 4) y los promedios y

desviaciones estándar en 5) calcular la distribución combinada por medio de la fórmula (1).

Notas: (1) En vez de usar  $N_1$ , directamente parece conveniente usar  $N$ , el tamaño total de la muestra, y la proporción de mezcla del grupo.

(11) Usando la distribución reconstruida como la esperada o teórica y la de la distribución de la muestra como la observada se performa una prueba G por medio de

$$G = 2 \sum F_o \log E_o/F_o \dots\dots\dots (32)$$

donde  $F_o$  y  $F_e$  son las frecuencias observadas y esperadas, respectivamente, (de SOKAL and ROHLF, 1981).

Notas: (m) Se prefiere la prueba G a la de Chi cuadrado principalmente porque para asignar una probabilidad a la prueba G es necesario sustraer 1 grado de libertad del número total de clases de la muestra analizada, independientemente del número de parámetros estimados así como del valor de la frecuencia de cada clase, no importa cuan pequeña. (n) Si una clase tiene frecuencia cero en ambas distribuciones, ambas deben ser convertidas a 1 para no perder un grado de libertad; si una de las clases tiene frecuencia cero y la otra un valor positivo, el cero se convierte en 1 y añade un 1 a la otra frecuencia. (o) Tanto la prueba G como la de Chi cuadrado requieren que las distribuciones comparadas sean del mismo tamaño  $\alpha = 0.5$ . Sin embargo, ver nota a la tercera línea al pie de la Tabla 2 así como (v) de la sección un caso de la literatura, etc.

7. Evaluar el número de parámetros correctamente estimados.

obtener

$$G' = G_{nx3}/10 \dots\dots\dots (33)$$

donde n es el número de grupos, y entrarlo como abscisa en la Fig. 1, la ordenada correspondiente determinada por la línea oblicua da, en por ciento sobre el total de parámetros que deben ser estimados, el número de estimaciones cuyos límites al nivel del 95% de confianza para los promedios, y del 90% para las proporciones y las varianzas tienen la mayor probabilidad de contener los valores verdaderos de los parámetros.

**APLICACION DEL NUEVO METODO A LAS DISTRIBUCIONES ESPECIALMENTE CONSTRUIDAS**

Aparentemente la mayor parte de los métodos de análisis han sido desarrollados alrededor de alguna distribución polimodal tomada, o imitada, de la naturaleza asumiendo solamente que es posible detectar en ella algún flanco expuesto que permite deducir algunas características de los grupos; esta afirmación se aplica por lo menos en el caso de los métodos de Cassie, Bhattacharya y Tanaka. El presente método ha sido desarrollado alrededor de distribuciones resultantes de combinaciones específicas de grupos individuales especialmente contruidos a partir de tamaños, promedios y desviaciones estándar arbitrariamente determinados. Esta nueva estrategia parece conveniente a fin de tener algún medio de comprobar los resultados no solamente con referencia a las frecuencias sino también en lo que se refiere a las proporciones, promedios y desviaciones estándar de los componentes.

Construcción de cinco distribuciones polimodales

Los cuatro grupos componentes de la distribución número 1 fueron conformados de acuerdo a la fórmula

$$f_1 = N \times \exp -(0.5 \times (1-0.5-X)/s)^2 \dots\dots\dots (34)$$

mientras que para las distribuciones número 2 la fórmula fue modificada con la introducción de componentes aleatorios

$$f_1 = N \times \exp -(0.5 \times ((1-0.5-X) + \dots) / (s + \dots))^2 \dots\dots\dots (34')$$

donde  $f_i$  es la frecuencia en la clase  $i$ ,  $N$  el tamaño del componente,  $X$  el promedio, y  $s$  la desviación estándar, y  $\pm$  son números al azar con un signo + o uno - también escogido al azar. El intervalo de clase fue 1 de modo que  $i$  fue 1, 2, 3, ...; para todos los grupos  $N$  fue designado arbitrariamente, lo mismo que  $X$  y  $s$ , excepto que  $X$  para los grupos componentes 2, 3 y 4 dependieron de algún modo de la posición de la  $X$  del grupo anterior. Sin embargo, los valores de  $N$ ,  $X$  y  $s$  usados en los cálculos posteriores fueron los que resultaron de aplicar las frecuencias calculadas, no necesariamente exactamente iguales a los escogidos arbitrariamente al comienzo. Para otros detalles ver las Tablas 1 y 2 del Anexo.

### Aplicación y resultados

Como se ha mencionado, las características de las cuatro distribuciones se dan en la Tabla 1 del Anexo mientras que los detalles del método se dan solamente para la distribución número 2 (ver las Tablas 1 y 2 del Anexo y la sección sobre el procedimiento de análisis). Los resultados del análisis se dan en las Tablas 3 y 4 del Anexo.

El paso 7, cuando se le aplica a las distribuciones especialmente construidas, resulta en la curva de la Fig. 1 la cual, en opinión del autor, proporciona un medio para poder discernir el poder y las limitaciones de un análisis de distribuciones polimodales, cuando se aplica a muestras naturales tal como se explica al comienzo del artículo.

### DISCUSION Y CONCLUSIONES

Como se muestra en la aplicación a la muestra de porgies, un resultado importante del presente ejercicio es aumentar en principio la fuerza de un análisis de frecuencias hasta hacerlo capaz de evaluar los estimados de proporción, promedio y desviación estándar de los grupos componentes.

Esta fuerza aumentada se aplica por el momento a través del uso directo de un gráfico; una limitación que es debida en parte al hecho de que en el ejercicio entró solo mi propio método y se usó cinco combinaciones de cuatro grupos; se espera que un aumento en el número de puntos ploteados en la Fig. 1 obtenido mediante un aumento del número de combinaciones, el número de grupos combinados y el número de métodos de análisis empleados conduciría a una estimación más estrecha de la fórmula usada para la relación entre  $G'$  y el porcentaje de parámetros estimados correctamente.

Una segunda limitación, independiente del número de puntos ploteados, es que por el momento no hay modo de identificar



-----

cuales son los estimados que incluyen los parámetros originales y cuales no los incluyen; tampoco para identificar el tipo de parámetros para que el estimado es correcto o no; igualmente, tampoco para identificar a que grupo componente en particular pertenecen los buenos y los malos estimados. Creo

que un más detallado análisis de los diversos pasos de cálculo en mi propio método ayudará en el tratamiento de esta limitación y al mismo tiempo espero que otros métodos de análisis ofrecerán también, cuando se les considere en detalle, medios para esclarecer estas cuestiones.

El presente estudio es empírico y un resultado empírico, hasta ahora, es que no todos los parámetros de grupo que entran en una combinación pueden ser correctamente estimados (se entiende aquí que un estimado es correcto cuando sus límites de confianza comprenden el verdadero valor de parámetros). Sospecho, sin embargo, que este resultado podría ser deducido teóricamente si los aspectos teóricos del modelo usado para la concepción y desarrollo del autor fuesen considerados en detalle y espero que los mismo se aplique a otros métodos.

Aparte de sus limitaciones e independientemente de que ellas sean superadas o reducidas me parece que el presente ejercicio ha mostrado que la proporción, promedio y desviación estándar de los grupos componentes de una distribución polimodal son de tres clases: aparentes, de mezcla y originales. Estos últimos serían aquellos que tendrían los grupos componentes si se considerasen por sí mismos y no en una mezcla o combinación de grupos; tan luego como esta combinación ocurre, algunos de los parámetros originales (la proporción y/o el promedio y/o la desviación estándar) son modificados como resultado del traslape de grupos, es necesario tomar en cuenta parcialmente estos traslapes a fin de estimar los parámetros de mezcla; nótese que si se tomase en cuenta completamente los traslapes la combinación no podría existir. El nuevo método desarrollado aquí distingue claramente entre los parámetros aparentes y de mezcla; de primera intención parece que otros métodos no hacen esta distinción.

Debe mencionarse que el reconocimiento de la existencia de algunos parámetros originales no reconocibles a través de la combinación puede proponer problemas muy interesantes en biología pesquera; por otro lado un trabajo más detallado en algunos de los procedimientos del método del autor podría llevar a una explícita determinación de los parámetros originales. En todo caso, parece que la distinción entre parámetros aparentes y de mezcla -ninguno de ellos ni ambos necesariamente iguales a los originales- puede desde ya ser un terreno fértil para otras investigaciones.

Finalmente, cuando en las combinaciones entran grupos normales o modificados por la introducción de elementos al azar o de simples sesgos las frecuencias resultantes del presente análisis tienen una probabilidad muy alta sin necesidad de transformar los datos, pero para obtener probabilidades parecidas en el caso de la muestra de porgy fue necesario aplicar a los datos, tanto de la muestra como del análisis, una transformación de suavización; esto apunta a la probable existencia de complicados procesos de perturbación en el muestreo mismo. Por otro lado, la combinación 2b -que incluye el promedio de varias variantes para cada grupo componente - da una estadística G mucho menor que la combinación 2a que comprende solo una variante por cada grupo. Estas dos observaciones enfatizan la necesidad de una buena revisión de las técnicas de muestreo para edad y crecimiento, para casos tipo porgy por lo menos.

#### AGRADECIMIENTOS

Deseo expresar mi gratitud al Dr. I. Barrett del Southwest Fisheries Center, California y al Dr. D. Pauly del International Center California y al Dr. D. Pauly del International Center for Living Aquatic Resources Managements, Manila y sus cuerpos de especialistas, por la lectura crítica del manuscrito. Al Dr. J. Gulland del Renewable Resources Assessment Group, Londres, y al Dr. R. J. Beamish de la Pacific Biological Station, Nanaimo, Canada, por sus consejos conducentes a futuras proyecciones. Al Proyecto Cooperativo Peruano Aelman (PROCOPA) en sus CO-Directores Drs. H. Salzweder y P. Muck por su ayuda material, incluyendole la necesaria para esta publicación. Ninguna de estas personas es reponsable de los errores y puntos debiles de este trabajo.

## REFERENCIAS BIBLIOGRAFICAS

- ABRAMOWITZ and STEGUN. Handbook of Mathematical Functions. National Bureau of Standards.
- BEAMISH, R.J. and G.A. MCFARLANE. 1983. The forgotten requirement for age validation in fishery biology. Transactions of the American Fisheries Society. Vol. 112, Nº 6: 735-743.
- CASSIE, R.M. 1954. Some use of probability paper for the graphical analysis of polymodal frequency distributions. Aust. J. Mar. Freshwat. Res., 5:513-22.
- GULLAND, J.A. Fish Stock Assessment, A Manual of Basic Methods. John Wiley & Sons. 1983, 223 pp.
- PAULY, D. and J.F. CADDY. 1985. A modification of Bhattacharya's method for the analysis of mixtures of normal distributions. FAO Fish Circ., (781): 16 p.
- SOKAL, R.R. and F.J. ROHLF. 1981 Biometry, 2nd. ed. San Francisco, W.H. Freeman and Company.
- TANAKA, S. 1962. A method of analysing a Polymodal Frequency Distribution and its application to the Length Distribution of the Porgy, Tilapia tilapia (TandS.)J. Fish. Res. Bd. Canada, 19 (6), 1143-1159.

Tabla 1. Proporción, promedio y desviación estándar de cada grupo componente de una muestra multimodal de longitudes de porgy (TANAKA, 1962) que resultan de la aplicación de 7 métodos de análisis.

P	Método	Total	GRUPO. 1			GRUPO. 2			GRUPO. 3			GRUPO. 4			GRUPO. 5		
			Prop.	Prom.	Des.est.	Prop.	Prom.	Des.est.	Prop.	Prom.	Des.est.	Prop.	Prom.	Des.est.	Prop.	Prom.	Des
?	Petersen	14054	40.63	3.6	—	32.73	7.9	—	14.51	12.4	—	—	—	—	—	—	—
0.00	Normsep	15034	33.8898	4.4099	0.7470	23.9200	7.4857	2.1589	14.7118	9.5000	1.0159	23.3255	13.0927	2.1084	4.1848	20.00	1.70
0.00	Bhattacharya	14048	41.8844	4.46	1.036	32.2049	9.06	1.264	17.5622	13.55	1.455	6.6160	17.52	1.620	1.6336	20.99	1.2
0.34	Cassie	13992	40.4803	4.02	0.76	31.6395	8.33	1.15	17.8816	12.65	1.32	6.9325	16.50	1.29	3.0660	19.92	1.5
0.76	Buchanan	14195	40.7326	4.05	0.844	31.1025	8.32	1.161	18.5981	12.85	1.412	6.4177	16.58	1.212	3.1631	19.82	1.4
0.79	Tanaka	14043	40.0698	3.99	0.8	31.9376	8.26	1.2	18.7282	12.84	1.4	5.9816	16.50	1.2	3.2828	19.82	1.4
0.99	Landa	14054	41.6894	4.0586	0.8713	28.5027	8.2422	1.0643	28.2815	13.1300	2.1614	4.5704	19.1368	1.9002	0.0360	25.2378	2.4

P: Probabilidad de que el valor de la estadística G de la comparación de las frecuencias de la muestra con las frecuencias calculadas con estos parámetros no difiera de 0 excepto por casualidad (ver Tabla 2).

Método: El método de Petersen es exclusivamente visual-identificar los modos al ojo. El NORMSEP es una rutina computarizada desarrollada por Tomlinson, aquí se usó una versión que requiere solamente una conjetura para cada promedio, esto es, también una conjetura del número de grupos. El de Bhattacharya es originalmente gráfico, aquí se usó una versión computarizada incorporada en el método ELEFAN (ICLARM 1987). Los de TANAKA y Cassie son gráficos, los resultados que aquí se dan están tomados de Tanaka 1962. El de Buchanan-Wollaston no son gráficos los resultados presentados han sido tomados también de Tanaka. El presente o de Landa está basado en el de Cassie, pero es completamente computarizable y no requiere ninguna conjetura al introducir los datos.

Total: La suma de todas las frecuencias resultantes del análisis (redondeadas a enteros). En el caso del método de Petersen el total usado es el de la muestra. El total debe ser igual al total de la muestra (generalmente + o - 0.5), de otro modo una prueba G o una de Chi cuadrado no es estrictamente aceptable. Aquí (ver Tabla 2) se aplica un procedimiento, explicado en otra parte de este trabajo, para subsanar la diferencia de totales.

Para cada grupo: Proporción: las que resultan del análisis, basadas en el total de la muestra antes del análisis. Promedio: los que resultan del análisis. En el caso del método de Petersen es el modo a la vista. Los promedios se dan en términos de una escala referencial, que comienza en 1 y procede por pasos de 1. Desviación estándar: las que resultan del análisis. Como las frecuencias son agrupadas, la desviación estándar podría ser corregida con la corrección de Sheppard o no dependiendo del modelo. En el presente método la corrección no se incluye. Para los otros métodos no hay indicación si la corrección es incluida o no.

Agradecimiento: Agradezco a J. Mendo de PROCOPA por proveer programas para NORMSEP y Bhattacharya.

Tabla 2. Distribución de frecuencias de longitud (datos al natural y suavizados) de una muestra de porov (ANAKA, 1962) como dada y como reconstruida a partir de proporciones, promedios y desviaciones estándar de los grupos componentes siguiendo 6 métodos diferentes de análisis (ver tabla 1). Para cada método, el valor de la prueba G y su probabilidad.

Sample	Morssen		Bhattacharva		Lassie		Bucanan		Janara		Landa			
	kav	smooth	kav	smooth	kav	smooth	kav	smooth	kav	smooth	kav	smooth		
1	7	119.00	3.31	31.53	6.13	168.02	9.06	83.19	9.39	103.82	0.21	103.59	0.70	128.92
2	79	567.00	14.59	296.16	106.33	560.09	12.10	533.31	28.47	343.06	22.10	368.72	38.42	562.72
3	509	1025.20	139.99	848.43	127.04	977.27	402.38	1045.22	306.15	1013.36	493.27	1028.84	363.40	1028.76
4	2240	1128.40	1324.40	1104.36	1928.63	1162.46	2352.91	1149.70	2210.47	1132.37	2326.69	1146.47	2273.03	1170.87
5	2341	1267.60	2764.11	1276.38	2084.65	1255.92	2441.57	1237.04	2377.67	1242.77	2301.07	1245.83	2331.69	1252.89
6	623	1382.00	1285.27	1349.70	931.45	1279.49	520.77	1493.65	704.19	1376.73	378.43	1390.30	711.83	1380.20
7	476	1221.80	625.39	1307.67	473.34	1171.79	447.46	1227.21	484.48	1257.28	529.06	1219.12	498.73	1242.32
8	1230	937.80	756.18	1027.20	944.72	989.90	1189.90	927.01	1183.40	964.84	1220.76	992.30	1200.43	957.73
9	1439	902.80	1114.03	975.31	1420.54	928.60	1523.39	903.57	1305.44	903.97	1467.62	916.12	1524.19	909.56
10	921	910.00	1360.36	888.86	1173.87	928.15	945.48	912.67	999.81	908.62	917.64	912.91	906.81	921.20
11	448	807.80	1025.57	967.45	625.16	863.40	413.93	822.52	446.98	818.27	446.52	815.11	470.79	806.90
12	512	634.60	693.09	869.32	471.11	716.77	482.77	633.30	307.93	633.37	315.02	638.09	306.97	631.37
13	719	539.40	649.07	697.56	631.16	581.92	734.65	633.39	726.69	633.87	731.64	662.87	678.53	557.06
14	673	538.00	622.86	528.18	677.96	521.30	689.99	335.29	682.86	532.96	683.00	537.91	643.91	537.06
15	445	497.60	500.71	456.56	502.09	479.18	436.61	504.03	443.73	498.63	440.79	497.67	537.47	481.86
16	341	399.40	327.95	350.10	322.30	399.49	324.73	407.98	331.19	406.76	321.72	398.07	368.25	391.33
17	310	298.40	184.51	248.76	260.62	299.13	325.95	302.22	339.41	304.80	311.74	292.31	230.78	300.11
18	228	237.40	116.22	176.20	233.01	224.27	235.98	240.12	262.18	243.83	233.36	232.22	174.07	227.56
19	168	192.00	115.68	136.84	176.51	180.15	162.91	196.23	167.44	200.63	154.07	191.45	186.04	176.97
20	140	142.80	136.63	118.14	128.13	140.39	127.15	144.09	137.39	146.96	140.26	141.91	174.44	140.63
21	114	101.60	131.65	104.67	101.82	98.44	105.97	98.15	112.69	99.48	117.83	99.42	115.11	107.91
22	64	68.00	92.94	84.90	62.99	63.94	63.82	66.96	63.11	67.31	63.99	69.43	50.01	71.36
23	22	40.40	46.79	38.46	22.39	38.38	27.32	41.71	22.13	40.22	21.03	41.49	14.33	36.70
24	0	18.00	16.75	32.27	4.14	18.00	7.45	20.49	4.60	17.95	4.15	17.49	3.14	13.87
25	2	5.20	4.26	13.72	0.39	3.39	1.33	7.27	0.64	5.46	0.49	5.14	1.08	4.03
26	2	0.80	0.77	4.37	0.02	0.91	0.16	1.79	0.03	1.09	0.03	0.53	0.83	1.27
27	0	1.00	0.10	1.03	0.00	0.08	0.01	0.30	0.00	0.14	0.00	0.01	0.73	0.71
28	0	0.60	0.01	0.10	0.00	0.00	0.00	0.03	0.00	0.01	0.00	0.01	0.54	0.49
29	1	0.20	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.32
tot.	14054	14035.00	14023.30	14034.97	14048.52	14035.00	13991.82	14034.79	14176.93	14034.98	14042.81	14036.44	14033.32	14034.50
Dif. con muestra			0.70	0.30	5.48	0.00	62.18	0.41	142.93	0.02	11.19	0.26	0.68	0.30
G de la distribución			2989.00	612.65	529.84	70.76	433.26	20.37	134.52	10.34	184.23	14.93	175.36	6.33
G de los totales			0.36		3.38		8.78		9.93		4.92			0.33
G final				613.21		74.34		26.65		20.47		19.87		6.66
P			0	0	0	0	0	0.337	0	0.763	0	0.788	0	0.995

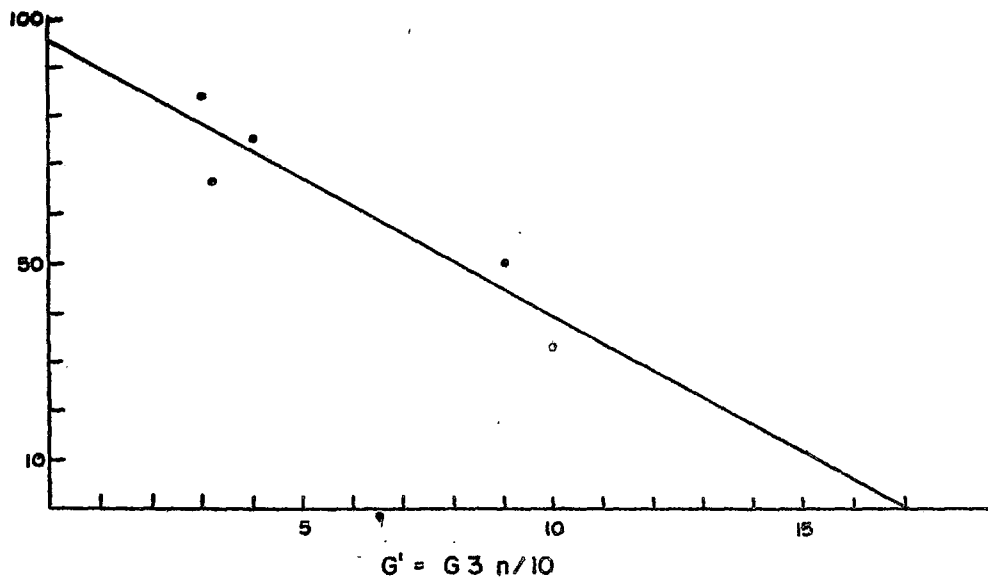


Fig.1. Relación entre valores G' y los parámetros correctamente estimados (en % del número total de parámetros). De la aplicación del método de Landa a 5 distribuciones especialmente construídas de 4 grupos cada una. Ver tablas anexo 4 y 5.

TABLA ANEXO 1. Parámetros de los 4 grupos originales y procedimiento para construir las 5 combinaciones especiales.

A. Grupos originales				
	N	Proc.	Prom.	Desv. est.
1	30	11.5385	6.5	1.75
2	110	42.3077	13.5	3.50
3	80	30.7692	24.5	5.00
4	40	15.3846	37.5	3.00

B. Procedimiento para construir las 5 combinaciones especiales	
Combinación 1:	Los parámetros de A se usaron en la fórmula (1) para obtener la distribución de frecuencias de los 4 grupos combinados.
Combinación 2a:	Los parámetros de cada grupo en A se usaron una sola vez en la fórmula (34') y las distribuciones de frecuencia resultantes fueron sumadas a través de cada clase.
Combinación 2b:	Las 4 distribuciones de grupos que fueron combinadas resultaron de promediar cinco distribuciones por cada grupo, cada una de las cuales a su vez resultó de la aplicación de la fórmula (34') con los respectivos parámetros en A.
Combinación 3:	Los parámetros de A fueron usados en la fórmula (34) -distinta de (34') en que no contiene elementos aleatorios. Luego se introdujo un sesgo: la adición de 0.5 a cada frecuencia de clases de los flancos ascendentes de los grupos 1 y 3 y de los descendentes de los grupos 2 y 4; de igual manera, sustracción de 0.5 en las clases de los flancos ascendentes de los grupos 2 y 4 y descendentes de los grupos 1 y 3. De este modo el sesgo de un grupo fue reforzado por el sesgo del siguiente grupo.
Combinación 4:	Construida como la combinación 3, pero el sesgo introducido fue adición de 0.5 a cada frecuencia de clases de los flancos ascendentes de los grupos 1 y 2 y descendentes de los grupos 3 y 4; de igual manera, sustracción de 0.5 en los flancos ascendentes de los grupos 3 y 4 y descendentes de los grupos 1 y 2. De modo que el sesgo de un grupo fue compensado por el del siguiente.

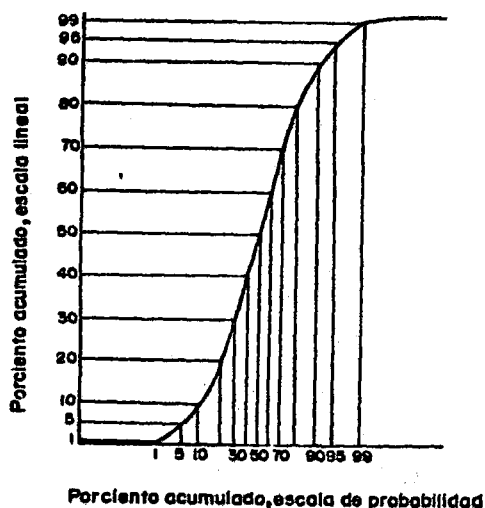


Fig. 3. Relación entre la escala lineal y la escala de probabilidades. Reproducida de la Fig. 6.5 de SOKAL and ROHLF 1961.

TABLA ANEXO 2. Distribución de frecuencias de las 5 combinaciones especiales.

COMBINACIONES

RS	1	2a	2b	3	4
1	0.07	0.29	0.43	0.50	0.50
2	.31	.91	.93	.75	1.31
3	1.06	.91	2.14	1.42	2.06
4	2.78	3.76	3.94	2.96	3.77
5	5.40	6.02	5.38	5.40	6.40
6	7.83	6.38	6.75	7.82	8.82
7	8.81	7.18	8.35	7.79	8.79
8	8.41	7.70	8.35	7.39	8.39
9	8.00	9.35	7.51	7.50	7.95
10	8.62	7.97	8.14	8.11	8.52
11	10.13	10.36	8.93	9.88	10.21
12	11.77	11.24	10.25	11.72	11.94
13	12.87	12.78	10.93	12.56	12.91
14	13.11	12.88	11.26	14.11	12.11
15	12.49	10.98	10.77	13.49	11.49
16	11.22	10.66	10.14	12.21	10.21
17	9.68	9.81	9.07	10.67	8.67
18	8.23	8.71	6.61	9.23	7.23
19	7.13	8.16	3.97	8.13	6.13
20	6.49	7.93	5.06	7.49	5.49
21	6.26	5.77	5.59	7.26	5.26
22	6.29	5.61	5.53	7.29	5.29
23	6.42	6.78	6.47	7.41	5.60
24	6.49	7.31	6.80	6.85	5.85
25	6.41	7.25	6.79	5.85	6.85
26	6.13	6.44	6.28	5.60	6.60
27	5.65	5.93	6.12	5.13	6.13
28	5.03	4.47	5.06	4.50	5.50
29	4.35	4.55	4.66	3.76	4.76
30	3.72	3.42	3.56	2.98	3.98
31	3.25	2.59	3.13	2.25	3.25
32	3.06	2.03	2.16	2.06	3.06
33	3.23	1.36	2.71	2.23	3.23
34	3.74	2.91	2.72	2.74	3.74
35	4.46	2.81	4.62	3.46	4.46
36	5.15	4.46	5.03	4.19	5.14
37	5.53	6.43	5.61	4.74	5.52
38	5.41	6.32	5.69	5.74	5.74
39	4.79	6.05	4.77	5.19	5.19
40	3.81	2.80	3.35	4.26	4.26
41	2.72	1.90	2.29	3.19	3.19
42	1.74	1.72	.95	2.23	2.23
43	1.00	.42	.44	1.49	1.49
44	.51	.08	.17	1.01	1.01
45	.23	.17	.06	.73	.73
46	.10	<u>254.16</u>	<u>239.47</u>	<u>261.87</u>	<u>261.60</u>
47	<u>.03</u>				
	<u>259.92</u>				

RS= Reference soals

TABLA ANEXO 3. Ejemplo: Hoja de trabajo para análisis (Combinación especial 2b).

(1)	(2)	(2')	(3)	(4)	(5)	(6)	(7)	(7')
1	45	59	1796	2.0877	12.0377			
2	53	87	5679	2.4681	3804			
3	214	166	1.4616	2.8192	3511			
4	3.94	3.42	3.1089	3.1343	.3191			
5	6.58	6.40	6.3832	3.3681	.2338			
6	6.75	7.24	8.1722	3.6032	.2181 (23.9457)			
7	8.35	8.62	11.6531	3.6077	.2015			
8	6.35	6.83	13.1459	3.9599	1621			
9	7.91	8.25	18.2820	4.0994	1256			
10	6.14	7.59	21.6812	4.2172	1218			
11	6.93	7.77	25.4103	4.3367	1215 (23.5457)*			
12	10.25	9.15	29.9906	4.4671	1284			
13	10.93	11.18	34.2549	4.5949	1278			
14	11.28	12.66	38.8569	4.7200	1281			
15	10.77	12.94	43.4543	4.8355	1185 (31.8198)			
16	12.14	13.78	47.9568	4.9422	1067			
17	9.07	8.18	51.4732	5.0369	0947			
18	6.61	6.68	54.2284	5.1081	0892			
19	3.97	4.53	55.9543	5.1460	0419 (63.0653)*			
20	3.06	4.27	58.0073	5.2017	0357			
21	3.59	4.78	60.3416	5.2618	0601			
22	3.53	5.56	62.6508	5.3222	0604			
23	6.47	6.35	65.3526	5.3944	0722			
24	6.80	6.69	68.1922	5.4726	0782 (28.4441)			
25	6.79	7.03	71.0277	5.5538	0812			
26	6.28	6.75	73.6504	5.6323	0785			
27	6.12	6.10	76.2088	5.7127	0804			
28	5.06	5.18	78.3188	5.7828	0701			
29	4.66	4.16	80.2647	5.8509	0661			
30	3.58	3.21	81.7314	5.9058	0549			
31	3.13	2.90	83.0384	5.9564	0506			
32	2.16	2.29	83.9804	5.9928	0364 (83.9094)			
33	2.71	2.46	85.0921	6.0404	0476			
34	2.72	3.24	86.2279	6.0906	0502			
35	4.62	4.39	88.1972	6.1830	0924			
36	3.03	3.41	90.2378	6.2066	1155			
37	6.63	5.94	92.6003	6.4469	1504			
38	5.69	5.67	94.9784	6.6429	1560 (16.4906)			
39	4.77	4.69	96.9683	6.8768	2337			
40	3.33	3.35	98.5672	7.1567	2601			
41	2.29	2.07	98.3235	7.4699	3332			
42	93	1.10	99.7202	7.7709	3010			
43	44	51	99.9039	8.1023	3314			
44	17	20	99.9749	8.4799	3778			
45	09	07	99.9999	8.7532	1.2763			

(7')	(7'')
(10)	11
11	8
19	15
28	6
32	13
32	3
(45)	

(9)	(10)
1)	23.5467; 31.5198; 28.4441; 16.4906
(10)	∅; 23.9457; 55.0693; 83.8094

(11)	(12)	(13)	(14)	(15)	
1)	11.7728	7.0326	2.7523	2.4166	2.5844
	3.7367	4.2803			
32	19.8080	9.4492			
2)	39.3059	14.0775	2.3443	2.5494	2.4469
	28.6478	11.7330			
	50.0631	16.6268			
3)	69.2873	24.3962	3.7127	3.9614	3.8370
	59.5794	20.6789			
	78.9853	29.3478			
4)	91.7147	36.6390	2.7284	2.6574	2.6929
	86.1264	33.9106			
	97.3829	38.2964			
(16)	10.5	18.5	31.5		

(17)	(18)	(19)	(20)	(21)
04	6.3417	90.8616	9.0384	2.1261
01	3.8379	7.2896	7.2896	2.2976
05	6.8074	96.3785	3.6215	1.1415
02	3.4659	6.3590	6.3590	1.8088
08	6.8540	96.7737	3.2243	9.171
03	3.0916	2.8839	2.8839	4.756

(9)	(9')	(10')					
1	23.5467	+2.1261	-2.2976		23.3762	∅	
2	31.5198	+1.1415	+2.2976	-2.1261	-1.8088	31.0218	23.3762
3	28.4441	+ .9171	-1.8088	-1.1415	-4.756	28.5529	84.3980
4	16.4906	+4.756	- .9171			16.0491	83.9509

(11')	(12')	(13')	(14')	(15')	
1)	11.6891	7.0093	2.7399	2.3990	2.5694
	3.7098	4.2884			
	19.6864	9.4073			
2)	39.8871	13.9851	2.3102	2.4865	2.3983
	28.2993	11.6749			
	49.4748	16.4716			
3)	69.1744	24.3463	3.8834	4.1377	4.0105
	59.0880	20.4630			
	79.2608	28.4841			
4)	91.9764	36.7332	2.5933	2.6133	2.6033
	86.4979	34.1399			
	27.4530	39.3463			

Total 239.47 239.15 (7) 45/7 = 6.42  
 (8) 23.5457 = 2.3947 = 56.36 (8') 16.4906 2.3947 = 39.49  
 56.36 / (11 - 1) = 6.12 > 100 39.49 / 45 = 87.31 > 100  
 (22) 239.47; 23.3762 % 0.083, 2.5694; 31.0218, 13.9851, 2.3983; 29.5529,  
 24.3464; 4.0105; 16.0491, 36.7532, 2.6033

(1) Escala de referencia; (2) frecuencia original; (2') frecuencia del análisis; (3) distribución acumulada de frecuencias, %; (4) Probits; (5) diferencias de probits; (6) Puntos de inflexión y proporciones aparentes; (7) Número mínimo de clases por grupo; (7') Puntos de inflexión en términos de la escala de referencia; (7'') Número de clases entre puntos de inflexión vecinos; (8) Número de individuos por clase en el primer grupo; (8') Número de individuos por clase en el último grupo; (9) Proporciones aparentes; (9') Proporciones aparentes corregidas por traslapes; (10) Puntos de inflexión; (10') Puntos de inflexión corregidos por traslapes; (11) Puntos en la curva de porcentajes acumulados correspondientes al promedio, promedio -1 desviación estándar y +1 desviación estándar; (12') lo mismo que (12) pero corregido por traslapes; (13, 14) desviación estándar por debajo y por encima del promedio; (15) Desviación estándar; (13', 14', 15') Lo mismo que (13, 14, 15) pero corregido por traslapes; (16) Puntos de inflexión en escala de referencia; (17) identificación de zonas de traslapes; (18) Traslapes en probits; (19, 20) traslapes en porcentajes acumulados de grupo; (21) traslapes en base del total; (22) inputs para fórmula (1); (24, 24') Límites; (25) Parámetros originales.

10% Límite de proporciones

(24)	(25)	(24')
16.9780	11.8385	29.7744
24.4439	42.3077	37.6000
* 22.9894	30.7692	36.1174
* 10.1428	15.3946	21.9356

5% Límite de promedios

(24)	(25)	(24')
* 6.3201	6.50	7.6965
* 13.4308	13.50	14.3384
* 23.3958	24.50	23.2970
* 35.4616	37.50	38.0048

10% Límites de variancia

(24)	(25)	(24')
4.2343	3.0625	11.4249
3.9087	12.2800	9.2544
* 10.7678	25.0000	29.9286
* 4.0367	9.0000	13.8921



TABLA ANEXO 4: Resultados del análisis de distribuciones especiales

(1)	Gp. 1			Gp. 2			Gp. 3			Gp. 4			
	(2)	(3)	(4)	(2)	(3)	(4)	(2)	(3)	(4)	(2)	(3)	(4)	(5)
1	16.2304	6.4135	1.8221	43.6164	13.9771	3.3656	23.0943	14.8888	3.4901	17.0589	36.6691	3.1525	159.92
2a.	25.1641	7.8478	2.8056	58.0068	15.3326	3.1871	21.5584	25.1943	3.0551	15.2707	37.0085	2.5708	254.16
2b.	25.3782	7.0083	2.5694	31.0218	13.9881	2.3983	29.5529	24.3464	4.0105	16.0491	36.7332	2.6033	239.47
3	29.3580	8.5880	3.1452	31.4696	13.5475	2.4429	22.8223	24.1657	3.3586	16.5501	37.4826	3.3185	261.87
4	18.1998	6.1071	2.0118	40.6510	13.7594	3.3177	22.5958	15.2623	3.3990	18.5534	37.0306	3.4121	261.56

(1) Número de la combinación; (2) Proporción; (3) Promedio; (4) Desviación estándar (5), tamaño total.

TABLA ANEXO 5: Numero de parámetros de grupo estimados correctamente en la 5 distribuciones especiales

(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	2	4	3	9	2.5468	3.0582
2b.	2	4	1	7	2.6513	3.1816
4	2	4	3	9	3.3232	3.9878
2a.	2	2	2	6	7.8728	9.2074
3	1	2	1	4	8.0306	9.9878

(1) Número de la combinación; (2) Proporciones; (3) Promedios; (4) Desviaciones estándar; (5) Numero total de parámetros correctamente estimados; (6) Valor  $b_1$ ; (7) Valor  $b_1' = b_1 n \times 3/10$  ( $n$  = número de grupos).

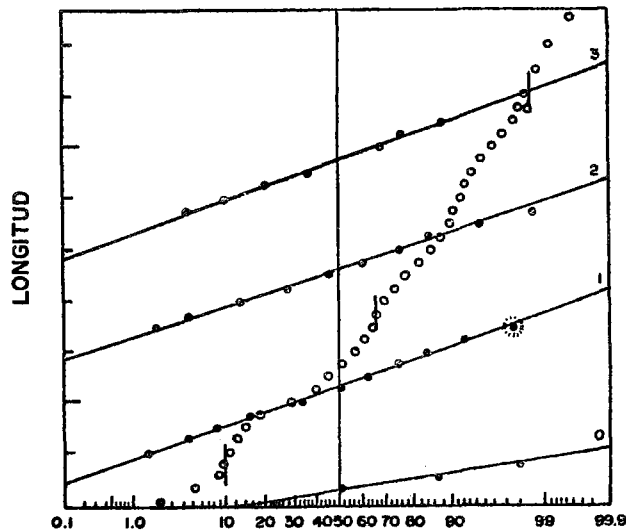
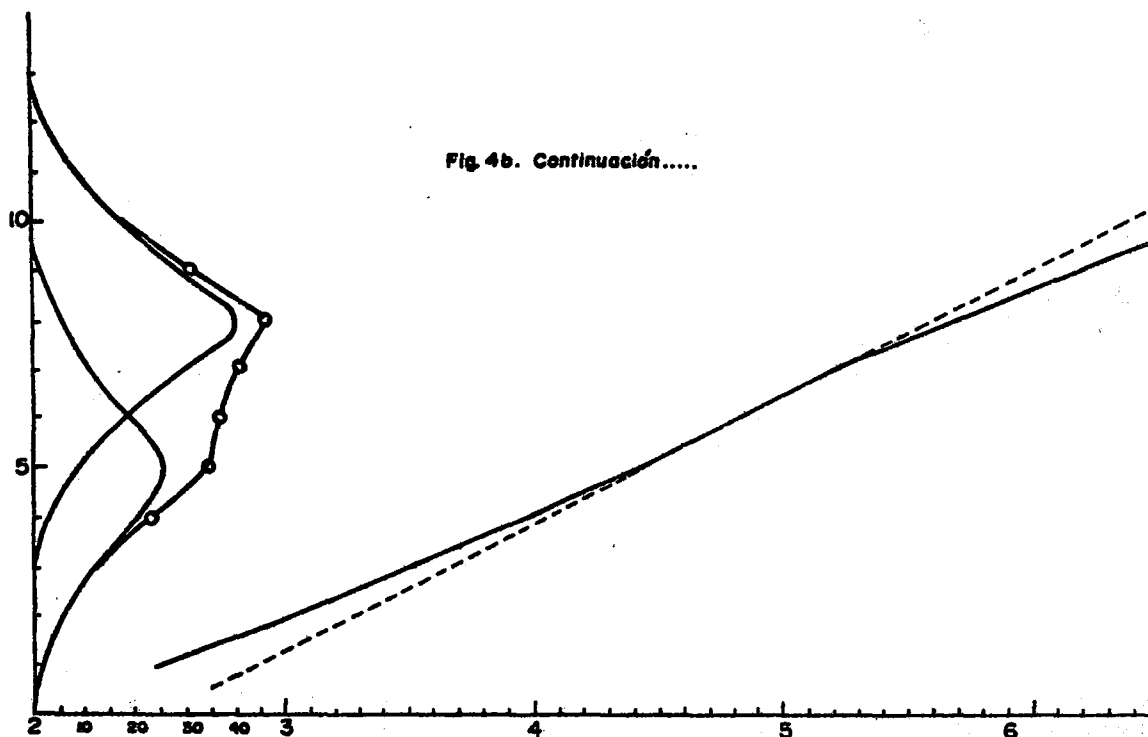
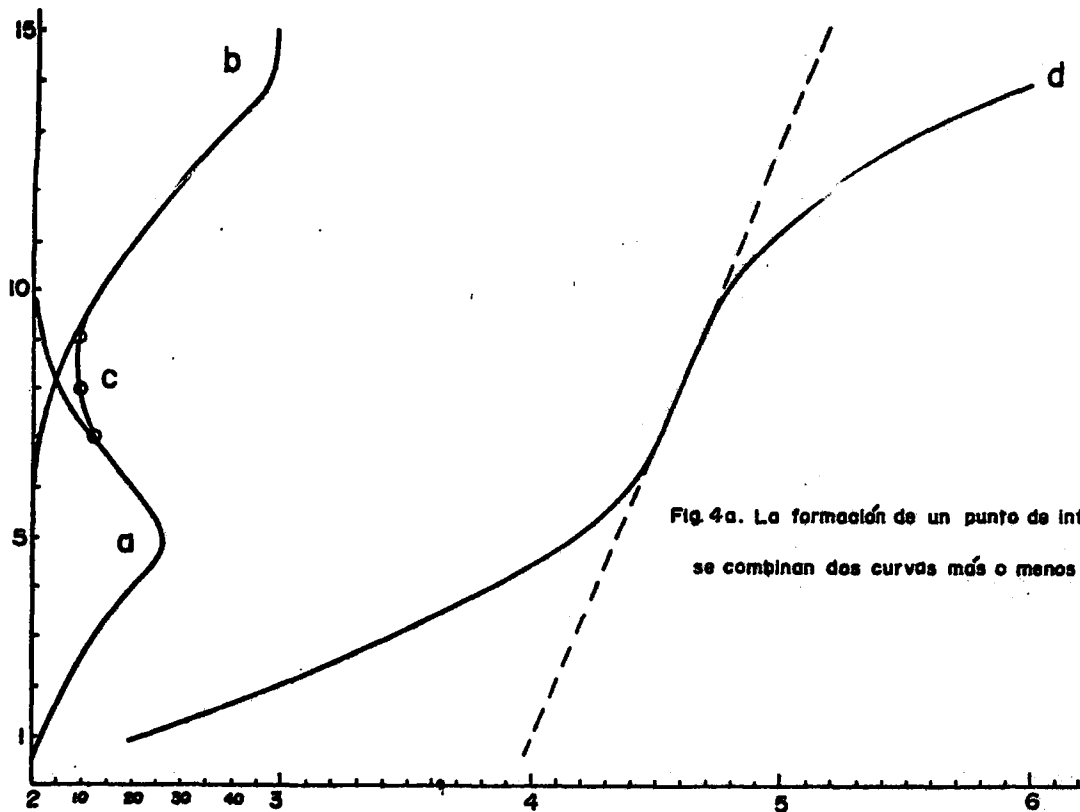
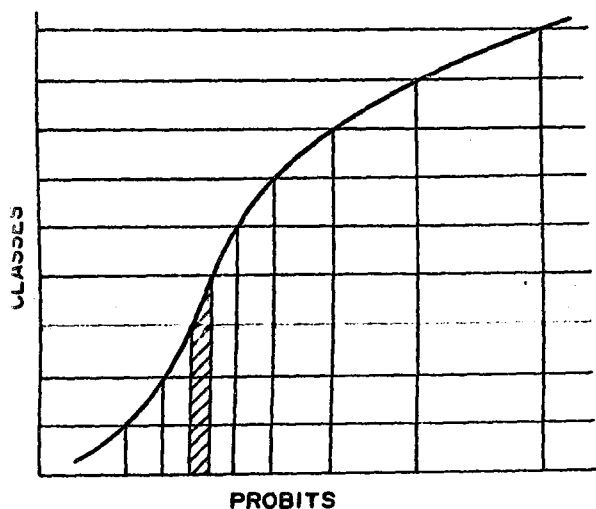


Fig. 2. Puntos de una distribución porcentual acumulada de longitudes de una muestra de mezcla, círculos abiertos, y puntos de los grupos componentes según el análisis con el método de Cassie, puntos negros (reproducida de la Fig. 1 de Cassie, 1954)





5. Modelo gráfico para encontrar el intervalo de clase donde un punto de inflexión, por medio de la diferencia entre probits de clases sucesivas.

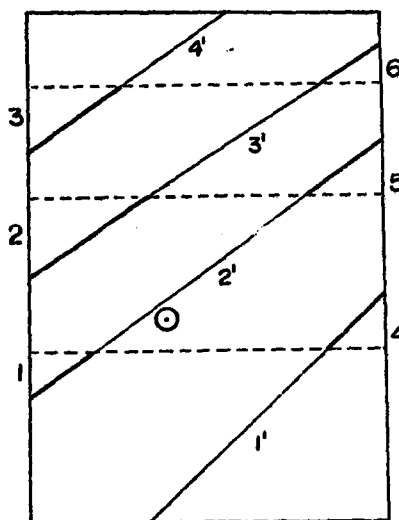
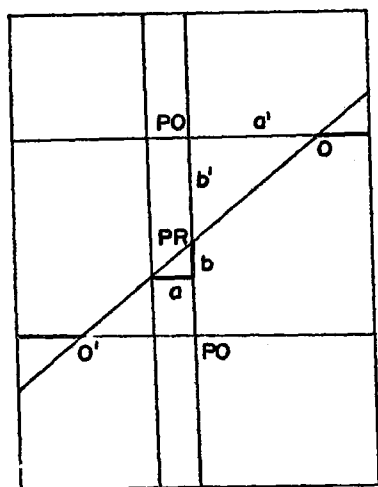


Fig. 6. Modelo gráfico que muestra la diferencia entre proporciones aparentes (segmentos delgados) y proporciones de mezcla, segmentos delgados + y - los traslapes (segmentos gruesos).



3. 7. Modelo gráfico para estimar traslapes, segmentos gruesos, en términos de probits mediante la similitud de triángulos rectángulos. (La unidad de la escala de referencia es la clase la misma que la unidad de la escala de probits).

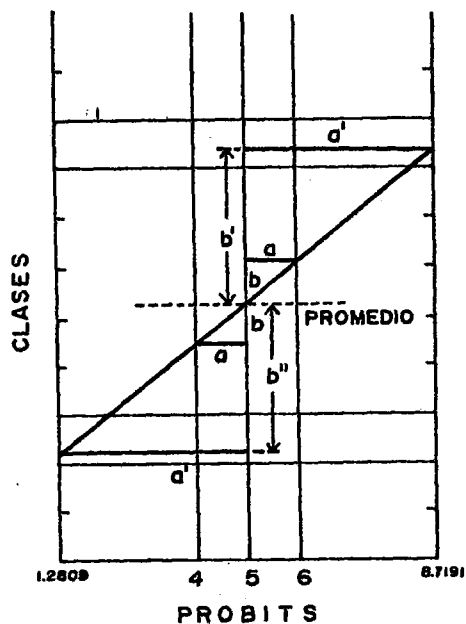


Fig. 8. Modelo gráfico para estimar los puntos finales de las líneas que representan los grupos en términos de la escala de referencia. (La unidad de la escala de referencia de las clases es la misma que la unidad de la escala de probits).