Notas Estadístico-Sociales

LAS DISTRIBUCIONES SOCIOGRÁFICAS ASIMÉTRICAS Y SU REPRESENTACIÓN MEDIANTE CURVAS ESTADÍSTICAS

Por ÓSCAR URIBE VILLEGAS.
De la Escuela Nacional de Ciencias Políticas y Sociales de la Universidad Nacional Autónoma de México.

Los fenómenos sociales pueden someterse a un primer intento de aprehensión científica mediante descripciones estadísticas. Para que tales descripciones sean posibles, es necesario —desde luego— comenzar por medir uno o más de los caracteres del fenómeno por estudiar, mediante la comparación de los mismos con un carácter análogo que se toma como unidad. La medida de los caracteres produce un conjunto de medidas que, en una segunda etapa, será necesario ordenar, mediante la formación de series crecientes o decrecientes de valores. Una tercera etapa consiste en tomar estas medidas ya ordenadas y sujetarlas a alguna elaboración que permita mostrar que la seriación de esos valores obedece a alguna ley matemática. Una etapa final ha de consistir en interpretar, desde el ángulo sociológico, la expresión resultante.

A través de un procedimiento como el descrito, se pasa, en efecto, de los niveles más bajos de abstracción —en que, para describir el fenómeno en estudio, era necesario detenerse en todos y cada uno de los detalles de cada uno y de todos los individuos del conjunto— a un alto nivel de abstracción en que todos esos detalles individuales, múltiples y aparentemente dispares, quedan subsumidos en una expresión matemática o una curva típica, de características conocidas y fáciles de interpretar.

En esta forma, substituir el conjunto desordenado de caracteres de los individuos del conjunto por una fórmula o una curva que expresen matemáticamente dicho conjunto no sólo sirve para que nos refiramos más lacónicamente al fenómeno, sino que permite que encontremos con mayor facilidad las conexiones estructurales, funcionales y, en el mejor de los casos, causales y significativas del fenómeno.

De ahí la importancia de las descripciones estadísticas —basadas en fórmulas o en curvas— en cuanto primer paso de la aprehensión científica.

Esas descripciones estadísticas, que tan importantes son en cuanto primer paso en el proceso científico de aprehensión, dependen, en grado considerable de la existencia previa —establecida por la matemática— de ciertas fórmulas y curvas *típicas*. Gracias a esa previa existencia de tipos de fórmulas y de curvas, es posible comparar una distribución de medidas sociográficas obtenidas de la realidad, con uno de esos tipos de distribución, con el fin de elegir el más conveniente, en cuanto representativo de la distribución concreta de que se trate.

De ahí que las posibilidades de descripción estadística adecuada estén en relación directa del número más o menos amplio de fórmulas y curvas típicas de que pueda disponerse como elementos con los que comparar las distribuciones concretas.

De ahí también el interés que tiene extender, tanto como sea posible, el repertorio de expresiones matemáticas y curvas típicas de que puede disponer el estudiante o el estudioso de ciencias sociales, pues, conforme más extenso sea ese repertorio, será más fácil encontrar una expresión o curva típica que describa con máximo rigor y precisión (dentro de grados de simplicidad compatible con tal rigor y tal precisión) una distribución determinada que se le presente para su estudio.

En los cursos de estadística elemental —según puede testimoniarlo cualquier libro de texto a ellos destinado— es frecuente que las expresiones y curvas descriptivas de las distribuciones de caracteres sociográficos se reduzcan a un número mínimo. No deja de ser frecuente —en efecto— que, en el campo de las distribuciones discontinuas, se estudie única, o casi únicamente, la distribución binomial y que en el de las continuas sea la normal la única que se estudie. Si bien esto —como punto de partida— no constituye adquisición que pueda ser despreciable por el estudiante, sí, en cambio, puede determinar serias limitaciones en sus actividades de investigación, en caso de que no amplíe convenientemente el campo de sus conocimientos.

En efecto, tan pronto como el estudioso deja las ejemplificaciones más o menos artificiosas o simplificadas que una buena pedagogía aconseja se utilicen dentro de las aulas, la práctica se encarga de enfrentarlo a distribuciones surgidas de la realidad cotidiana, ajenas a la conveniente simplificación o al artificio orientado hacia un fin y, en este enfrentamiento, suele descubrir que la mayoría de las distribuciones discontínuas difiere de la binomial, en la misma forma en que es difícil reducir la mayoría de las contínuas a la distribución o curva normal.

En casos en que la desviación de las distribuciones empíricas con respecto a las distribuciones teóricas centrales a que hemos aludido (binomial y normal) es poco considerable, el estudiante puede utilizar sus tipos únicos de expresión matemática sin deformar grandemente la realidad, y las interpretaciones sociológicas a las que le lleva dicha utilización pueden llegar a tener un grado de validez tan alto como pequeña sea dicha desviación; pero, cuando dicha desviación es considerable, usar de esos tipos centrales únicos representa un serio riesgo en cuanto su uso hará que la descripción deforme la realidad, y que esa deformación repercuta en interpretaciones que se supone deben basarse en una descripción de la misma que le sea tan fiel como sea posible.

Frente a la manifiesta necesidad de extender el campo de las disponibilidades en este terreno —incluso en el nivel elementalísimo de preparación matemática en que pudimos movernos al redactar una *Técnicas estadísticas para investigadores sociales*— recurrimos a la presentación de un sistema de curvas continuas (curvas de Gram-Charlier) mediante el que es posible representar no sólo distribuciones simétricas (como ocurre con la curva normal) sino otras francamente asimétricas. El sistema utilizado entonces tiene limitaciones bien conocidas, entre las que no es la menor la que consiste en que, ocasionalmente, aparecen frecuencias negativas de difícil o imposible interpretación; pero, con todo, en su momento, tal sistema nos pareció que podía servir para llenar una laguna en la preparación de los estudiantes de estadística de la Escuela Nacional de Ciencias Políticas y Sociales a quienes estaba destinado el texto.

Lo poco satisfactorio del sistema de Gram-Charlier y el conocimiento que tenemos de que la preparación matemática media del alumno de la Escuela ha ido elevándose durante los últimos años, nos impele a poner en sus manos, a través de esta *Revista*, una sistematización de las curvas asimétricas continuas que, a más de no presentar los inconvenientes del sistema de Gram-Charlier, puede mantenerse dentro de los límites de una si no modestísima sí, por lo menos, de una mediana preparación matemática.

De acuerdo con una norma prudente, que nos impone partir de lo conocido para llegar a lo desconocido, tomaremos como punto de partida las distribuciones simétricas, ya conocidas, o sea, que usaremos la distribución binomial (discontinua) que tiende hacia la distribución normal (continua).

La distribución binomial puede expresarse en la forma siguiente:

$$f(S) = \binom{M}{S} p^S \ (1\text{-}p)^{M\text{-}S} \tag{0}$$

En esta fórmula: f(S) representa la frecuencia con la que aparece el carácter S (número de soles obtenidos, por ejemplo, al tirar al aire un cierto número de monedas); M representa el número de individuos del conjunto (número de monedas que se arrojan al aire); p, la probabilidad que hay de que salga un sol con una moneda. Consiguientemente, 1—p representa la probabilidad que hay de que no salga el carácter considerado (la probabilidad de que salga un águila) y M—S el número de individuos que, en el conjunto, no mostraron el carácter S (número de monedas que, tras haber caído, muestran en su cara visible águilas y no soles).

Como se puede recordar si se trae a la mente el estudio que se ha hecho de la distribución binomial, p y 1—p pueden ser diferentes. Esto ocurre cuando se trata de determinar la probabilidad que hay de obtener un 6 al lanzar un dado $(p = 1/6)$ frente a la que hay de obtener un número que no sea 6 al lanzar dicho dado $(1-p = 5/6)$. Sin embargo, para partir del caso más sencillo, consideraremos una distribución binomial en la que p y 1—p sean iguales (según ocurre en el caso de las probabilidades que hay de obtener un sol o un águila con una moneda, ya que, en tal caso, $p = 1/2$ y $1-p = 1 - 1/2 = 1/2$).

En el caso de una binomial como ésta, la expresión anterior se convierte en:

$$f(S) = \binom{M}{S} p^S \ p^{M\text{-}S} = \binom{M}{S} p^M \tag{1}$$

La expresión que acabamos de escribir da la frecuencia con la que aparecen S soles con M monedas. La expresión que da la frecuencia con la que aparecerán S+1 soles, si se utiliza el mismo número de monedas (M), es:

$$f(S+1) = \binom{M}{S+1} p^M \tag{2}$$

Si restamos de la última expresión la previa (de la 1 la 2), tendremos el incremento en las frecuencias, que corresponde al incremento de una unidad, sufrido por la variable (que ha pasado de S a S+1). Si dicho incremento se representa por delta mayúscula de f(S), tendremos:

$$\Delta\ f(S) = f(S+1) - f(S) = p^M \left[\binom{M}{S+1} - \binom{M}{S}\right] \tag{3}$$

En forma parecida, si sumamos las dos expresiones y dividimos la suma resultante entre dos, obtendremos la frecuencia correspondiente a un valor intermedio. Esto lo representaremos por f(S+.5).

$$f(S+.5) = \frac{f(S+1) + f(S)}{2} = \frac{p^M \left[\binom{M}{S+1} + \binom{M}{S}\right]}{2} \tag{4}$$

Si dividimos la expresión (3) entre la (4), obtendremos:

$$\frac{\Delta\ f(S)}{f(S+.5)} = \frac{2\ p^M}{p^M} \frac{\left[\binom{M}{S+1} - \binom{M}{S}\right]}{\left[\binom{M}{S+1} + \binom{M}{S}\right]} \tag{5}$$

En esta expresión, el 2 del numerador procede del 2 que figuraba en el denominador de f(S+.5). La fracción del segundo miembro puede simplificarse, ya que p elevado a la M figura en el numerador y en el denominador, como factor de ambos, por lo cual se reduce a la unidad. Una simplificación ulterior puede lograrse si se considera que:

$$\binom{M}{S+1} = \frac{M-S}{S+1}\binom{M}{S} \tag{6}$$

La substitución del factor combinatorio por su equivalente, nos permite escribir:

$$\frac{\Delta f(S)}{f(S+.5)} = \frac{2\binom{M}{S}\left(\frac{M\text{-}S}{S-1} - 1\right)}{\binom{M}{S}\left(\frac{M\text{-}S}{S+1} + 1\right)} \tag{7}$$

Expresión es ésta en la que, tras la substitución, se ha sacado como factor común al coeficiente combinatorio de M en S. Como dicho coeficiente figura en el numerador y en el denominador, se reduce a la unidad. Reducido este coeficiente, es posible escribir, como segundo miembro de la ecuación:

$$\frac{2\left(\dfrac{M-S-S+1}{S+1}\right)}{\left(\dfrac{M-S+S+1}{S+1}\right)} \tag{8}$$

En esta expresión, se han ejecutado las operaciones de dentro de los paréntesis de la (7). El resultado hace figurar S+1 como denominador tanto de la fracción numeradora como de la fracción denominador y, por tanto, S+1 se reduce a la unidad. Si, además, se realizan las operaciones indicadas en los numeradores de la fracción numeradora y en los de la fracción denominadora, se obtiene, como segundo miembro, la expresión (9):

$$\frac{2(M-2S+1)}{M+1} = \frac{2M-4S+2}{M+1} = \frac{4S+2M+2}{M+1} \tag{9}$$

En la última expresión, puede reconocerse la existencia de: un término variable (—4S) y la suma de dos términos constantes (2M+2), en el numerador, y dos términos constantes (M y 1) en el denominador. Si representamos por x la expresión variable (—4S), y por a y b las sumas de constantes (2M+2) y (M+1), la expresión anterior puede reducirse a:

$$\frac{x+a}{b} \tag{10}$$

Por lo que se refiere al primer miembro de la igualdad (10), que es, simultáneamente el primer miembro de la (7), ésta puede considerarse —en el límite— como la derivada de la función con respecto a la variable, dividida entre la función misma; o sea, como:

$$\frac{D_x y}{y} \tag{11}$$

Consiguientemente, toda la expresión puede escribirse como:

$$\frac{D_x y}{y} = \frac{x+a}{b} \tag{12}$$

Esta expresión corresponde, por tanto, a la forma diferencial de la curva hacia la que tiende la binomial cuando de discontinua pasa a continua, o sea, cuando, en vez de crecer "por saltos", crece por incrementos infinitamente pequeños. Escrita en su forma diferencial, es probable que el estudiante no reconozca una curva bien conocida; ello se debe a que la ha manejado siempre en su forma integral. A fin de mostrar el proceso que lleva de la forma diferencial de la ecuación a la integral, procederemos a tomar las integrales de ambos miembros de la ecuación e integraremos con respecto a x. Para expresarlo en fórmulas, antepondremos a cada miembro el operador $\int_x$

$$\int_x \frac{D_x y}{y} = \int_x \frac{x + a}{b} \tag{13}$$

La integral del primer miembro es bien conocida; se trata de la integral de la derivada de una función, dividida entre la función misma; dicha integral es igual al logaritmo de la función (o sea, que todo el primer miembro se reduce a Ly). En cuanto al segundo miembros de la (13), la integral del mismo se descompone o puede descomponerse en dos integrales: las integrales de dos fracciones (que tienen por numeradores a x y a, respectivamente, y por denominador común a b). También puede decirse que, en cuanto el denominador común de esas fracciones es una constante (b), el resultado estará constituido por la suma de las integrales de x (o sea, $x^2/2$) y de a(o sea, ax) dividida dicha suma entre el denominador común b. Lo cual puede expresarse como sigue:

$$\frac{1}{b}\left(\frac{x^2}{2} + ax\right) \tag{14}$$

Este segundo miembro corresponde a una ecuación cuyo primer miembro es, según hemos señalado, Ly (logaritmo natural de y)

$$Ly = \frac{x^2 + 2ax}{2b} + C \tag{15}$$

Si en esta fórmula se toman los antilogaritmos del primero y del segundo miembro, la expresión no se altera.

El antilogaritmo del logaritmo de y es, por supuesto, y (puesto que se ejecutan con dicha literal dos operaciones opuestas que anulan sus efectos). O sea, que y será el primer miembro de la ecuación resultante. En cuanto al antilogaritmo del segundo miembro, éste se obtiene si se toma dicho miembro como exponente de la base de los logaritmos naturales, o sea, que el resultado es el número usualmente representado por la letra e (base de dichos logaritmos) elevado al que era segundo miembro de nuestra expresión (15) de origen. Toda la igualdad se convierte, por tanto, en:

$$y = e^{\frac{x^2 + 2ax}{2b} + C} \tag{16}$$

En el segundo miembro figura e elevado a una fracción más C; o sea, que dicho miembro puede descomponerse en dos factores exponenciales, de los cuales, el primero sería e elevado a la fracción, y el segundo, e elevado a C. Si a e elevado a C lo representamos por Y', la expresión anterior puede escribirse:

$$y = y' e^{\frac{x^2 + 2ax}{2b}} \tag{17}$$

Esta forma integral de la curva correspondiente puede ser reconocida fácilmente por el estudiante, pues se trata de la curva normal.

En esta expresión existen: una variable (x) y dos parámetros (a y b) cuyos valores hay que determinar. Y' no nos interesa por el momento.

Para la determinación del valor de los parámetros, recurriremos a los momentos.

Como se recordará, un momento estadístico se define como la media de las potencias de las desviaciones o, lo que es lo mismo, como la suma de las potencias de las desviaciones dividida (dicha suma) entre el efectivo de la distribución.

Cuando el momento lo es con respecto al origen, las desviaciones son los propios datos (que hemos decidido representar por x), con lo cual la fórmula del momento (que representaremos por v) de orden n (orden que representaremos mediante la adición de un subíndice n a la literal que indica el momento) es igual a:

$$v_n = \frac{\Sigma x^n}{N} \qquad \text{(N: efectivo de la distribución)} \qquad (18)$$

Cuando la serie es de frecuencias, las potencias aparecen ponderadas (o sea, en este caso, "multiplicadas") por las frecuencias (que hemos representado aquí, generalmente, por y). O sea, que la fórmula del enésimo momento con respecto al origen será, en el caso de dicha serie de frecuencias:

$$v_n = \frac{\Sigma x^n y}{N} \qquad (19)$$

Si en vez de ser discontinua, es continua la serie de la que se trate, se determinarán las correspondientes modificaciones en la fórmula de los momentos; en ella, la sumatoria se transforma en integral (prácticamente, el operador sigma mayúscula Σ se transforma en el operador ese alongada $\int$, o integrador con respecto a x) y N se transforma en 1.

Según todo lo anterior, la fórmula del enésimo momento, para una distribución continua como la que nos interesa, resulta ser:

$$v_n = \int_x x^n y \qquad (20)$$

De acuerdo con esto, si en la expresión diferencial de la curva continua (12) quitamos denominadores (pasando b que está como denominador en el segundo al primer miembro, como factor, y pasando igualmente a y que figuraba como denominador del primer miembro al segundo, como factor) tendremos:

$$b\,D_x y = y\,(x + a) \qquad (21)$$

Si se parte de esta expresión, podrá obtenerse el enésimo momento de la distribución si ambos miembros se multiplican por la enésima potencia de la variable independiente

(x^n). La multiplicación de ambos miembros de la ecuación (21) por x elevada a la n produce:

$$bx^n D_x y = yx^n(x + a) \tag{22}$$

Para obtener el enésimo momento, se necesitará, además, integrar ambos miembros con respecto a x (o sea, que habrá que anteponer a ambos el operador $\int_x$):

$$\int_x bx^n D_x y = \int_x yx^n(x + a) \tag{23}$$

El integrado (o expresión afectada por el integrador) del segundo miembro, puede desarrollarse multiplicando cada término de dentro del paréntesis por los que están fuera de él; con ello, dicho segundo miembro se convierte en:

$$\int_x (yx^{n+1} + ayx^n) \tag{24}$$

Este miembro puede transformarse si se aplica en la integración una regla ya conocida para su análoga, la sumación, y que es bien conocida en matemáticas como ley distributiva: la integral de una suma (como la sumatoria de una suma) es igual a la suma de las integrales de los sumandos (como la sumatoria lo era de la suma de sumatorias de dichos sumandos). De este modo, el segundo miembro se convierte en:

$$\int_x yx^{n+1} + \int_x ayx^n \tag{25}$$

En este segundo miembro, puede reconocerse: como primer término, la integral de las potencias n+1 de la variable x, ponderadas por la función y, o sea: el momento de orden n+1 (representable por v_{n+1}) y, como segundo término, la integral de una constante (a) por el producto de ciertas variables (una de ellas potenciada), que se reduce a la constante (a) por la integral del producto, el cual, a su vez (por consideraciones análogas a las anteriores) no es sino el momento de orden n (representable por v_n). O sea, que el segundo miembro de la expresión (23) es igual a:

$$v_{n+1} + a\ v_n \tag{26}$$

En cuanto al primer miembro de la (23), si sacamos a b fuera del integrador, tendremos:

$$b \int x^n D_x y \tag{27}$$

Esta integral puede obtenerse por el llamado "método de integración por partes". La regla correspondiente a la integración por partes dice que:

La integral de un producto (integral de x^n por D_x y) es igual a:
el primer factor (x^n),
por la integral del segundo factor (integral de $D_x y = y$)
menos la integral del producto formado por
la derivada del primer factor (derivada de $x^n = nx^{n-1}$)
por la integral del segundo (integral de $D_x y = y$)
Por tanto, en el caso concreto, se tendrá, como primer miembro de la ecuación (23):

$$b \left[x^n y - \int_x nx^{n-1} y \right] \qquad (28)$$

En este primer miembro, $x^n y$ tiende a desvanecerse, por lo cual todo el miembro puede considerarse reducido a b multiplicado por el segundo término (el término integral) de dentro del paréntesis; pero, la integral de dentro del paréntesis no es sino n veces la integral de las potencias de orden n—1 de x, ponderadas por y; o sea, n veces el momento de orden n—1; por tanto, este miembro puede considerarse reducido a:

$$-bnv_{n-1} \qquad (29)$$

De acuerdo con esto, la expresión (23) equivale a:

$$-bnv_{n-1} = v_{n+1} + av_n \qquad (30)$$

Si se deja a v_{n+1} en un miembro, se tiene:

$$-v_{n+1} + av_n + bnv_{n-1} \qquad (31)$$

Esta expresión genérica nos dice que en una curva típica como la que estamos tratando, si se multiplica por el parámetro a el enésimo momento y al producto se le resta el resultado que se obtiene de multiplicar el momento inmediatamente anterior por b y por n, se obtendrá el momento siguiente.

Una expresión como la que acabamos de anotar, permite que se resuelva uno de los dos problemas siguientes:

1. Conocidos los momentos de una distribución (las v's de la fórmula) es posible determinar los parámetros (a y b) o,
2. Conocidos los parámetros (a y b) de un tipo de distribución es posible determinar los momentos (v's) de la distribución a la que corresponde.

Como es fácil comprender, los dos problemas anteriores son inverso el uno del otro, y mientras el primero interesa fundamentalmente al practicante de la estadística, el segundo es de utilidad primordial para el teórico de la misma. Esto no significa, con todo, que el segundo problema no pueda tener interés práctico o que el primero carezca de interés teórico. Con todo, de momento, en cuanto el estudiante de ciencias políticas y sociales tiene que ser, en su preparación inicial, más un práctico que un teórico de la estadística, será del primer problema del que nos ocupemos.

Si tenemos una distribución y calculamos los momentos respecto al origen y, por otra parte, tenemos la expresión genérica anterior (21), podemos determinar a qué curva específica corresponde, tan pronto como determinemos los valores de a y de b y los substituyamos en la expresión misma. Para hacerlo, debemos recordar que si tenemos dos incógnitas (los parámetros a y b) necesitaremos dos ecuaciones simultáneas que las contengan. Para obtener esas ecuaciones simultáneas, lo que haremos será especificar la expresión genérica, dando a n diferentes valores. Esos valores conviene sean pequeños (0, 1, 2...) pues los primeros momentos de una distribución están menos sujetos a error que los momentos de orden superior. Como necesitamos dos ecuaciones simultáneas, daremos a n, sucesivamente, los valores de 0 y 1. Así obtendremos:

$$av_0 + b(0)v_{-1} = -v_1 \qquad (32)$$

$$av_1 + b(1)v_0 = -v_2 \qquad (33)$$

En la primera ecuación, no nos interesa calcular el valor de v_{-1} pues todo el factor se reduce a 0, ya que en el producto figura 0 como factor. La reducción de las dos ecuaciones a sus términos más simples nos da:

$$av_0 = -v_1 \qquad (34)$$

$$av_1 + bv_0 = -v_2 \qquad (35)$$

Al despejar al parámetro de la expresión (34), se obtiene $a = -v_1/v_0$; pero, como v_0 es igual a 1 (el momento de orden cero siempre es igual a la unidad, como puede comprobarse fácilmente si se recuerda que es la media aritmética de las potencias de orden cero, de las desviaciones), se tiene:

$$a = -v_1 \qquad (36)$$

Si se considera que v_0 vale 1, la segunda ecuación se reduce a:

$$av_1 + b = -v_2 \qquad (37)$$

Al substituir en esta ecuación el valor de a, se tiene:

$$-v_1{}^2 + b = -v_2 \qquad (38)$$

Y, al despejar a b, se obtiene:

$$b = -v_2 + v_1{}^2 \qquad (39)$$

Conocidas la expresión genérica (21) y las expresiones (36) y (39) que dan los valores de los parámetros de dicha expresión genérica en términos de los momentos, es posible interpolar una curva normal a una serie cuyos dos primeros momentos con respecto al origen (v_1 y v_2) pueden calcularse aritméticamente, de acuerdo con procedimientos ya conocidos.

El recorrido que hemos hecho por terreno parcialmente conocido, nos ha permitido mostrar cuál es el método que habrá de seguirse para obtener fórmulas que sirvan para la descripción estadística de series que ya no sean simétricas como lo es la curva normal, sino asimétricas. Esta obtención abreviada y en sus términos más simples, es la que nos proponemos alcanzar en seguida.

Si realizamos un recorrido paralelo al que hemos hecho previamente, nuestro esfuerzo deberá concentrarse en:

1. Partir de la expresión de una curva discontínua apropiada al caso,
2. Determinar la relación que existe entre un incremento unitario y la frecuencia media del intervalo unitario de dicha curva.
3. Determinar, en la expresión resultante, cuáles son los términos constantes y cuáles los variables, y representarlos en la forma más simple que sea posible.
4. Pasar de la expresión diferencial a la expresión integral de la curva correspondiente, y
5. Determinar la relación entre los parámetros de la expresión obtenida y los momentos de la distribución.

Si en el caso anterior, nuestro punto de partida estuvo en una binomial, simétrica, en la que p era igual a 1—p, en este caso, en que no nos interesa obtener una curva continua simétrica, sino un sistema de curvas continuas asimétricas, tomaremos como punto de partida una binomial asimétrica (o sea, una en la que p es diferente de 1—p):

$$f(S) = \binom{M}{S} p^{M} (1-p)^{M-S} \tag{40}$$

En esta fórmula, puede suponerse —para referencia concreta— que M es el número de dados empleados en una jugada y que S es el número de los que dieron un 6, mientras M-S es el número de los que dieron una puntuación distinta de 6.

Si mediante esta fórmula se calcula el incremento de las frecuencias para una unidad de incremento en la variable, se calcula asimismo la frecuencia media del intervalo y se las relaciona, podrá obtenerse en el segundo miembro una expresión fraccionaria, en cuyo numerador figurarán variables y constantes, y en cuyo denominador (a diferencia de lo que ocurría en el caso anterior, en que sólo figuraban constantes) figuran también variables y constantes. El desarrollo que conduce a este resultado es, en términos generales, como sigue:

$$f(S) = \binom{M}{S} p^{S}(1-p)^{M-S} \tag{41}$$

$$f(S+1) = \binom{M}{S} p^{S+1}(1-p)^{M-(S+1)} \tag{42}$$

$$\text{Incremento:} \quad \Delta f(S) = \binom{M}{S} p^{S} (1-p)^{M-S} \left[1 - \frac{M-S}{S+1} p\ (1-p)\right] \tag{43}$$

$$\text{Frecuencia:} \quad f(S+.5) = \frac{1}{2}\binom{M}{S} p^{S} (1-p)^{M-S} \left[1 + \frac{M-S}{S+1} p\ (1-p)\right] \tag{44}$$

Al dividir la (44) entre la (45), el factor combinatorio, la potencia S de p y la potencia M-S de (1—p) que son comunes a los segundos miembros de ambas expresiones, se reducen a la unidad, con lo cual se tiene:

$$\frac{\Delta f(S)}{f(S+.5)} = \frac{2\,[\,1 - Mp(1-p) + S\ (1 + p(1-p))\,]}{1 + Mp(1-p) + S\ (1 - p(1-p))} \tag{45}$$

Como puede verse, hemos agrupado los términos a modo de que en el numerador de la fracción del segundo miembro sea claramente visible que existe una serie de términos constantes 1 y Mp(1—p) que, aunque unidos por diferente signo, también aparecen al principio en el denominador, así como que en la parte final del numerador y en la final del denominador aparece un término variable (un término en S) afectado por dos coeficientes distintos (ya que en el numerador, los términos de dicho coeficiente, 1 y p(1—p) están unidos por el signo más y en el denominador los une el signo menos).

De acuerdo con lo anterior, si como casos anteriores, designamos por x a la variable y por a, b y c, a las constantes, podremos escribir (tras recordar, por otra parte, lo que hemos dicho anteriormente con respecto a $\Delta f(S)\ /\ f(S+.5)$):

$$\frac{D_x y}{y} = \frac{x + a}{bx + c} \tag{46}$$

En este caso, hemos igualado el segundo miembro de la expresión original a $D_x\ y/y$ por razones análogas a las que nos movieron a hacerlo en el caso anterior. Según esto, la expresión (46) es la fórmula diferencial del sistema de curvas asimétricas que buscamos.

Esta expresión, como la del caso previo, tiene posibilidad de ser integrada. Sin embargo, antes de integrar, transformaremos la expresión, a fin de que tanto ella misma como los resultados que de ella procedan sean más fáciles de manejar.

La transformación puede lograrse si, en lugar de utilizar los datos mismos (x), utilizamos las desviaciones con respecto a la media aritmética de la serie $(x - \bar{x} = d)$. En este caso, los parámetros serán distintos. Podemos representarlos por a', b' y c'. De acuerdo con esto, la expresión diferencial, cuando se utilizan desviaciones respecto a la media aritmética, será:

$$\frac{Dyd}{y} = \frac{d + a'}{b'x + c'} \tag{47}$$

Si, para simplificar aún más la expresión resultante, en vez de tomar desviaciones respecto a la media, en unidades originarias (d), tomamos desviaciones con respecto a la media en unidades sigmáticas (delta minúscula δ), los parámetros también variarán. Si representamos el parámetro correspondiente de a' por alfa minúscula; por beta minúscula subíndice cero el correspondiente a c' y por beta minúscula subíndice uno el correspondiente a b', la expresión resultante será:

$$\frac{Dyd}{y} = \frac{\delta + \alpha}{\beta_1 + \beta_0} \tag{48}$$

Esta expresión, por su parte, puede resultar más fácilmente integrable, si substituimos a delta minúscula por $D - \beta_0/\beta_1$:

$$\frac{Dy}{y} = \frac{D - \dfrac{\beta_0}{\beta_1} + \alpha}{\beta_1 \left(D - \dfrac{\beta_0}{\beta_1}\right) + \beta_0} \tag{49}$$

Es fácil ver que, en el denominador del segundo miembro, al multiplicar beta minúscula subíndice 1 por beta cero sobre beta uno, se obtendrá beta cero, con signo negativo, que habrá que sumar a la beta cero de fuera del paréntesis, con la que se anula. Con esto, el denominador quedará reducido al producto de beta uno por D.

En cuanto al numerador, los dos últimos términos son constantes y pueden representarse por un nuevo parámetro que designaremos A.

De acuerdo con lo anterior, la expresión (49) puede escribirse:

$$\frac{Dy}{y} = \frac{D + A}{\beta_1 D} = \frac{D}{\beta_1 D} + \frac{A}{\beta_1 D} = \frac{1}{\beta_1} + \frac{A}{D\beta_1} \tag{50}$$

O bien:

$$\frac{Dy}{y} = \frac{1}{\beta_1}\left(1 + \frac{A}{D}\right) \tag{51}$$

Ésta es la fórmula diferencial del sistema de ecuaciones que estamos investigando.

Para obtener la forma integral del sistema, bastará con anteponer el operador propio de la integración ($\int_x$ en el primer miembro y $\int_D$ en el segundo) a los dos miembros de la ecuación. Sin embargo, como en el segundo miembro aparece el factor constante $1/\beta_1$, el mismo puede quedar fuera del integrador que afectará, por tanto a la suma dentro del paréntesis:

$$\int_x \frac{Dy}{y} = \frac{1}{\beta_1}\int_D \left(1 + \frac{A}{D}\right) \tag{52}$$

La integral del primer miembro es conocida: se trata de la integral de la derivada de una función, dividida entre la función, la cual es igual al logaritmo de la función, o sea, a L y. En cuanto a la integral que aparece en el segundo miembro, se trata de la integral de una suma, que es igual a la integral de cada sumando sumada con la integral de todos y cada uno de los restantes (o sea, a la "suma de las integrales de los sumandos"):

$$L\,y = \frac{1}{\beta_1}\left[\int_D 1 + \int_D \frac{A}{D}\right] + C \tag{53}$$

La integral del primer término del segundo miembro es igual a D (variable de referencia del integrador); la del segundo término del segundo miembro es igual a A por el logaritmo de D, puesto que se trata de una constante por el recíproco de la función que, como se sabe, al integrarse, produce el logaritmo natural. Según esto, la expresión se nos convierte en:

$$L\,y = \frac{1}{\beta_1}\,(D + A\ LD) + C \tag{54}$$

En esta expresión es posible tomar antilogaritmos.

El antilogaritmo del logaritmo de y es y, en el primer miembro.

En el segundo miembro, el sumando C se convierte, al tomar antilogaritmos, en e (base de los logaritmos naturales) elevado a C, y este factor lo representaremos por Y.

En el mismo segundo miembro, el otro sumando se convierte en otro factor. Al tomar antilogaritmos $1/\beta_1$ que figura como factor en la expresión logarítmica, pasará a figurar como exponente de la antilogarítmica. Por su parte, los términos de dentro del paréntesis representarán, en la expresión antilogarítmica, factores en vez de sumandos: el primer factor estará constituido por e elevado a D (primer sumando de la logarítmica) y el segundo por e elevado a A L(D) (segundo sumando de la logarítmica) que, a su vez, se reduce a D elevado a la A. O sea, que al pasar de la forma logarítmica a la antilogarítmica, se obtiene:

$$y = Y\ (e^{D}\ D^{A})^{1/\beta_1} \tag{55}$$

Esta expresión (55) es importante, pues se trata de la forma integral del sistema de curvas simétricas que estamos considerando.

Antes de proceder a encontrar los valores de D y A en términos de los momentos, conviene determinar el valor de Y; para ello, integraremos los dos miembros de la ecuación entre cero e infinito. La integración, en el segundo miembro, puede hacerse dejando a Y fuera del integrador, por tratarse como un factor constante que puede entrar y salir del integrando sin alterarse o alterar la expresión.

$$\int_0^\infty y = Y \int_0^\infty e^{D/\beta_1} D^{A/\beta_1} \tag{56}$$

Con el fin de poner de relieve la estructura esencial de la integral que figura en el segundo miembro, cambiaremos variable. Para ello, haremos a D igual a $-\beta_1$ z. Cuando se cambia variable, en una integración:

1º Hay que substituir en el integrado (o expresión afectada por el integrador) la antigua variable por la nueva.

Esto hace que, en nuestro caso, D/β_1 se convierta en $-\beta_1 z/\ \beta_1 = -z$ y D a la $A/\ \beta_1$ en $-\beta_1 z$ a la $A/\ \beta_1$, o sea, en $-\beta_1$ a la $A/\ \beta_1$ por z a la $A/\ \beta_1$

2º Hay que multiplicar el integrando por la derivada de la antigua variable respecto a la nueva.

En nuestro caso, habrá que multiplicar por $D_z D = D_z(-\beta_1 z) = -\beta_1$

3º Hay que cambiar el referente del integrador.
La integración, en nuestro caso, ya no se hará con respecto a D, sino con respecto a z, y

4º En caso necesario, habrá que cambiar los límites de la integración, substituyendo los límites antiguos en la expresión que da el equivalente de la nueva variable en términos de la antigua.
Como en nuestro caso, $z = D/-\beta_1$, si substituimos por D el límite inferior, cero, obtendremos como nuevo límite inferior, cero, y si substituimos por D el límite superior, infinito, se obtendrá también infinito como límite superior.

5º Según todo lo dicho, la expresión se transforma en la siguiente:

$$\int_0^\infty y = Y \int_{z=0}^\infty e^{-z} z^{A/\beta_1} (-\beta_1)^{A/\beta_1} (-\beta_1) \tag{57}$$

Si sacamos fuera del integrador todos los factores constantes, tendremos:

$$\int_0^\infty y = Y(-\beta_1)^{(A/\beta_1+1)} \int_0^\infty e^{-z} z^{A/\beta_1} \tag{58}$$

El factor integral del segundo miembro de la expresión (58) es muy característico y bien conocido.

Se trata de:

Una integral ($\int$) / definida entre cero e infinito ($0, \infty$) / de un producto formado por un factor exponencial (e elevado a la menos z) y otro formado por una potencia (A/β_1) del simétrico del exponente (z)/en la que la potencia ($-z$) de la exponencial es el simétrico (signo menos) de la base de la potencia (z) que forma el otro factor / refiriéndose la integración a la variable de la que se trata ($z=$, en la parte baja del integrador).

La función típica que acabamos de describir se conoce, en el cálculo integral, como la función gamma mayúscula (Γ) del exponente de la variable (A/β_1) adicionado, dicho exponente, de una unidad (o sea, que en nuestro caso, se trata de la función gamma de $(A/\beta_1 + 1)$. De acuerdo con esto, toda la expresión previa se transforma en la siguiente si, a más de todo, para facilitar la escritura, representamos $A/\beta_1 + 1$ por R:

$$\int_0^\infty y = Y (-\beta_1)^R \Gamma(R) \tag{59}$$

Basta con despejar a Y de esta ecuación para obtener:

$$Y = \frac{\int_0^\infty y}{(\beta_1)^R \Gamma(R)} \tag{60}$$

Con estas dos expresiones genéricas, se tiene andada la mitad del camino.

Para especificar las expresiones genéricas (21) y (60), será necesario que, como en el caso previo, encontremos la equivalencia de los parámetros en términos de los momentos.

En el caso anterior, utilizamos momentos con respecto al origen, pues nuestra expresión diferencial contenía x, o sea, las desviaciones respecto al origen. Como en este caso estamos tratando con desviaciones sigmáticas, nuestros momentos tendrán que serlo con respecto a la media aritmética, en unidades sigmáticas.

Nuevamente tendremos que plantearnos el problema de encontrar los parámetros alfa, beta cero y beta uno a partir de los momentos de la distribución. Como los datos son desviaciones respecto a la media, dadas en unidades sigmáticas, conviene —según ya dijimos— que los momentos también lo sean. Si representamos por gamma minúscula (afectada de diferentes subíndices que indiquen el orden del momento), dichas medidas, podremos llegar a obtener una expresión análoga a la que obtuvimos en el caso de las distribuciones simétricas, pero en la que figurará un término más que los que figuraban en aquélla (ya que, siendo tres parámetros, se necesitan tres términos y tres ecuaciones para encontrar sus valores):

$$\alpha\gamma_n + n\beta_0\,\gamma_{n-1} + (n+1)\,\beta_1\,\gamma_n = -\gamma_{n+1} \tag{61}$$

La especificación de esta expresión genérica se logra si se substituye n, sucesivamente, por 0, 1, 2 (puesto que siendo tres los parámetros por determinar, son tres las ecuaciones necesarias). Esto produce el sistema de ecuaciones:

$$\gamma_0 + 0\beta_0\,\gamma_{-1} + 1\,\beta_1\,\gamma_0 = -\gamma_1 \tag{62}$$

$$\gamma_1 + 1\beta_0\,\gamma_{\ 0} + 2\,\beta_1\,\gamma_1 = -\gamma_2 \tag{63}$$

$$\gamma_2 + 2\beta_0\,\gamma_{\ 1} + 3\,\beta_1\,\gamma_2 = -\gamma_3 \tag{64}$$

Pero, en este sistema:

β_0 es igual a 1 (pues el momento de orden cero es siempre igual a la unidad, como que es suma de potencias, cero, o sea, de unidades, entre número de datos).

β_1 es igual a 0 (pues el primer momento, *respecto a la media*, sea o no sigmático es siempre nulo), y

β_2 es igual a 1 (pues es el segundo momento sigmático o sea el segundo momento respecto a la media entre el cuadrado de la desviación cuadrática media, representada generalmente por sigma minúscula, que es igual al mismo segundo momento).

Según esto, el sistema se simplifica, reduciéndose a:

$$\alpha + 0 + \beta_1 = 0 \tag{65}$$

$$0 + \beta + 0 = -1 \tag{66}$$

$$\alpha + 0 + 3\beta_1 = -\gamma_3 \tag{67}$$

Al despejar de la primera, se obtiene:

$$\alpha = -\beta_1 \tag{68}$$

De la segunda, obtenemos:

$$\beta_0 = -1 \tag{69}$$

Al substituir en la (67) alfa minúscula por su equivalente, dado por la (68), y despejar a β_1, obtendremos:

$$\beta_1 = -\frac{2}{\gamma_3} \tag{70}$$

Este último, es un primer resultado importante, pues β_1 es uno de los parámetros que figuran en las expresiones genéricas. En dichas expresiones (21) y (60) figuran también como parámetros compuestos D y A, cuyos valores hay que determinar. D es, según sabemos, igual a la desviación con respecto a la media aritmética en unidades sigmáticas (o delta minúscula) menos beta cero sobre beta uno, pero como acabamos de encontrar que beta cero vale menos uno, al substituir tendremos:

$$D = \delta - \frac{1}{\beta_1} \tag{71}$$

En cuanto a A, es ésta una literal que nos sirvió para representar los sumandos constantes del numerador de la expresión diferencial (49), o sea que es igual a alfa menos el cociente de beta cero entre beta uno; como beta cero entre beta uno resulta ser igual a menos uno sobre beta uno, A resultará igual a alfa más uno sobre beta uno; pero como, además, alfa minúscula es igual a menos beta uno, de acuerdo con la expresión (68), tendremos que:

$$A = -\beta_1 + \frac{1}{\beta_1} \tag{72}$$

Finalmente, el parámetro R que figura en la expresión (60) y que da el valor de Y es igual a $A/\beta_1 + 1$, pero si se substituye el valor de A recién encontrado (72), toda la expresión se reduce a $-1 + 1/\beta_1^2$ (equivalente de A entre β_1) $+ 1$; o sea que:

$$R = \frac{1}{\beta_1^2} \tag{73}$$

Esto nos permite considerar como expresión para Y

$$Y = \frac{\int_0^\infty y}{(-\beta_1)^R \Gamma(R)} \tag{74}$$

y como expresión genérica del sistema:

$$y = Y(e^D D^A)^{1/\beta_1} \tag{75}$$

Las fórmulas (74) y (75) nos permiten discutir, en términos generales este sistema de curvas asimétricas.

β_1 puede ser igual a cero o diferente de cero.

Si β_1 es igual a cero, la curva pertenece al sistema de la curva normal: es simétrica y tanto más aplanada cuanto su índice de achatamiento sea menor que 3 ó tanto más picuda conforme su índice de achatamiento sea mayor que 3. Cuando dicho índice de achatamiento es igual a 3, se trata, precisamente, de la curva normal. En los otros dos casos (de índices inferiores o superiores a 3 pero de asimetría cero) se trata de curvas que, de momento, no nos interesan por estarnos ocupando exclusivamente de curvas *asimétricas.*

Si β_1 es diferente de cero, la curva es asimétrica. En tal caso, β_1 puede ser mayor que cero o menor que cero

Si β_1 es mayor que cero,

$1/\beta_1$ es mayor que cero,

$D = \delta - \dfrac{1}{\beta_1}$ crece al crecer δ ya que se le resta una constante positiva.

e^D es una exponencial creciente

$A = -\beta_1 + \dfrac{1}{\beta_1}$ es:

$$\text{negativa si } \beta_1 > \frac{1}{\beta_1}$$

$$\text{nula si } \beta_1 = \frac{1}{\beta_1}$$

$$\text{positiva si } \beta_1 < \frac{1}{\beta_1}$$

D^A es una potencial

decreciente (hiperboloide) cuando A es negativa
recta horizontal (D^0) cuando A es nula
creciente (paraboloide) cuando A es positiva.

De acuerdo con lo anterior:

Si β_1 es mayor que su recíproco la exponencial "se atempera" con la multiplicación por el hiperboloide.

Si β_1 es igual a su recíproco ($\beta_1 = 1/\beta_1$) la exponencial no cambia substancialmente.

Si β_1 es menor que su recíproco, la exponencial "se refuerza" con la multiplicación por el paraboloide.

Conforme $e^D\ D^A$ aumente, aumentará $(e^D D^A)^{1/\beta_1}$ que es, propiamente una raíz.

Sea β_1 mayor o menor que cero R, recíproco de su cuadrado, será positivo. Cuando β_1 sea igual a 1, R será igual a 1, y Γ (R) será igual a 1. Cuando β_1 sea mayor que 1, su cuadrado será también mayor que 1, el recíproco de su cuadrado será menor que 1, y Γ (R) será la función gamma de un número menos que 1. Cuando β_1 sea igual a 1, ya dijimos que Γ (R) será igual a 1. Cuando β_1 sea menor que 1, su cuadrado será menor que 1, el recíproco de su cuadrado será mayor que 1, y Γ (R) será la función gamma de un número mayor que 1. Si se examina una tabla de la función gamma para valores comprendidos entre 1 y 2, puede observarse que los valores de la función decrecen de 1 a 1.46 y aumentan de ahí hasta 2; por otra parte, debe observarse también que los factores de la función gamma mayúscula aumentan continuamente de una unidad a otra, pues en tratándose de enteros existe una relación clara entre la función gamma y los factoriales; así, la función gamma mayúscula de 1 es 1, la de 2 es 1, la de 3 es 2, la de 4 es 6, la de 5 es 24 (o sea: 0! 1! 2! 3! 4!)

Si β_1 es menor que cero,

$1/\beta_1$ es menor que cero,

$D = \delta - \dfrac{1}{\beta_1}$ crece al crecer δ, ya que se le agrega una constante.

$A = -\beta_1 + \dfrac{1}{\beta_1}$ es:

$$\text{positiva si } \beta_1 > \frac{1}{\beta_1}$$

$$\text{negativa si } \beta_1 < \frac{1}{\beta_1}$$

$$\text{nula si } \quad \beta_1 = \frac{1}{\beta_1}$$

D^A es una potencial

creciente (paraboloide)	cuando A es positiva
recta horizontal (D^0)	cuando A es nula
decreciente (hiperboloide)	cuando A es negativa.

De acuerdo con esto:

Si β_1 es mayor que su recíproco la exponencial "se refuerza" con la multiplicación por el paraboloide.

Si β_1 es igual a su recíproco ($\beta_1 = 1/\beta_1$ la exponencial no cambia

Si β_1 es menor que su recíproco, la exponencial "se atempera" con la multiplicación por el hiperboloide.

Conforme $e^D D^A$ aumente, disminuirá $(e^D D^A)^{1/\beta_1}$ pues el exponente es, en este caso, menor que cero y una potencia negativa equivale al recíproco de la potencia positiva correspondiente.

Trataremos de recoger, en seguida, algunas de las variaciones anteriores en forma gráfica.

Tomaremos, en primer término, el caso de una β_1 menor que cero. Si $\beta_1 = -.5$, su recíproco será igual a -2 $(1/\beta_1)$; D, que es igual a delta minúscula menos dicho recíproco, será igual a $\delta-(-2) = \delta + 2$; e^D será la exponencial, creciente, marcada en la Gráfica 1; $(e^D)\ 1/\beta_1$ será igual a $(e^D)^{-2}$ y quedará representada por la exponencial decreciente, o curva de parecido hiperboloide, representada en esa misma gráfica. Por su parte, A es igual al simétrico de β_1 $(-\!-.5 = +.5)$ más el recíproco de β_1 $(1/-.5 = -2)$; o sea, que $A = .5-2 = -1.5$ D^A será una potencial decreciente y $(D^A)^{1/\beta_1}$ será igual a $(D^{-1.5})^2 = D^3$ o sea, que, conforme muestra la gráfica, se tratará de una parábola de tercer grado.

Si tomamos el caso de una β_1 mayor que cero, obtendremos los valores correspondientes y las curvas componentes que hemos representado en la gráfica 2.

La multiplicación de la exponencial por la potencial en cada uno de los dos casos anteriores nos permite trazar, como dos tipos extremos, los representados en la gráfica 3.

De acuerdo con el desarrollo que nos conduce a las fórmulas (21) y (60), es posible delinear un procedimiento de interpolación para este tipo de curvas asimétricas. El mismo consistirá, fundamentalmente, en lo siguiente:

1º Calcular el tercer momento con respecto a la media aritmética en unidades sigmáticas.

Para ello habrá que:

A. Calcular la media aritmética (x)

B. Restarla de cada dato, para obtener desviaciones (d)

C. Elevar al cuadrado las desviaciones obtenidas (d^2)
sumar los cuadrados, dividir la suma entre el efectivo y extraer la raíz cuadrada del cociente,
para obtener la desviación cuadrática media (σ)

D. Dividir cada diferencia o desviación entre la desviación cuadrática media, para obtener desviaciones sigmáticas (δ)

E. Elevar al cubo los cocientes (δ), sumar los cubos, dividir la suma entre el efectivo (γ_3)

O bien habrá que:

A. Calcular la media aritmética

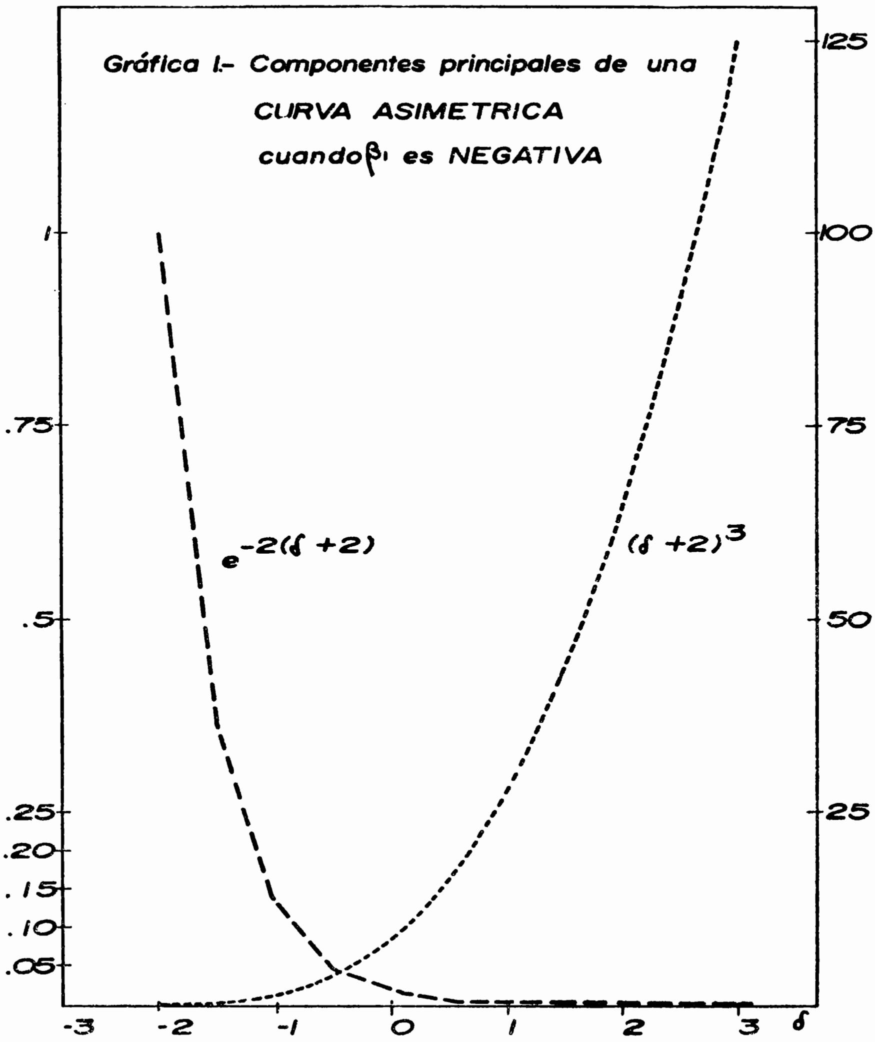
B. Restarla de cada dato

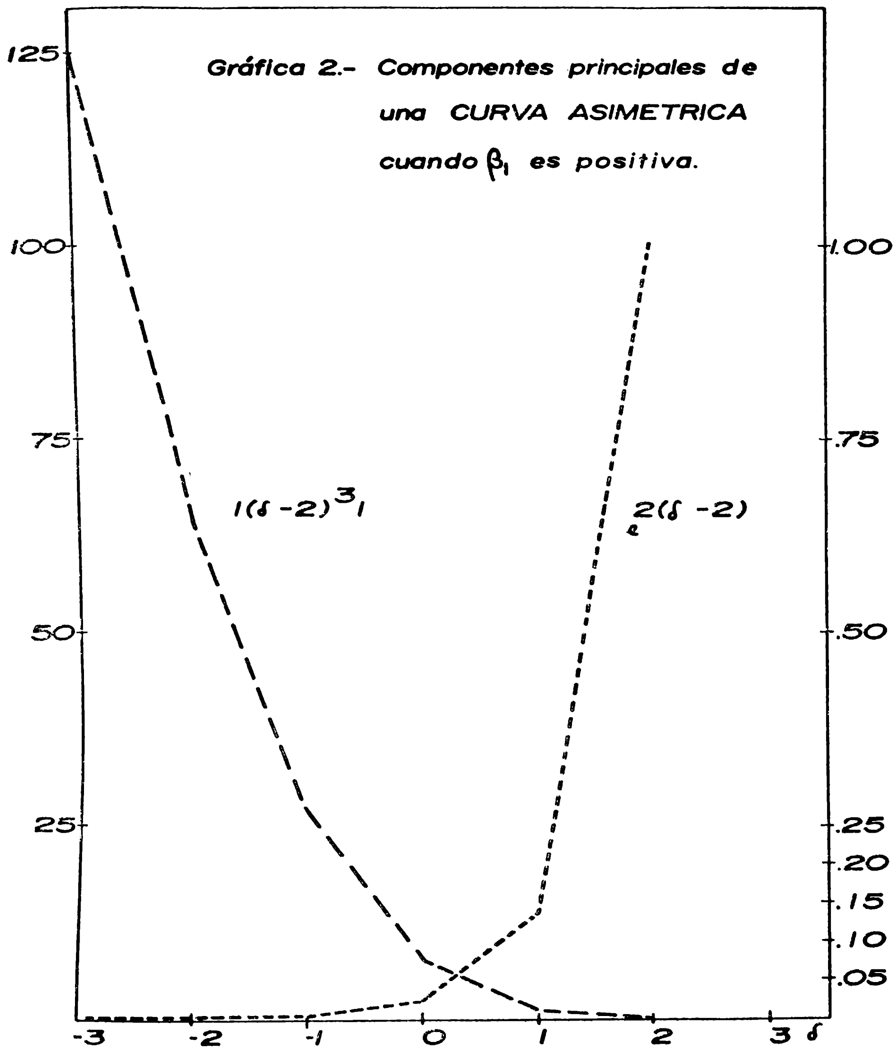
C. Obtener, a partir de las desviaciones, el segundo y el tercer momento respecto a la media:

a. elevando las desviaciones al cuadrado, sumando los cuadrados y dividiendo la suma entre el efectivo, y (u_2)

b. elevando las desviaciones al cubo, sumando los cubos y dividiendo la suma entre el efectivo (u_3)

D. Dividir el tercer momento entre la raíz cuadrada del segundo momento, elevada al cubo (γ_3)


Gráfica I.- Componentes principales de una
CURVA ASIMETRICA
cuando β_1 es NEGATIVA
$e^{-2(\delta+2)}$
$(\delta+2)^3$
1
.75
.5
.25
.20
.15
.10
.05
125
100
75
50
25
-3
-2
-1
0
1
2
3
δ


Gráfica 2.- Componentes principales de
una CURVA ASIMETRICA
cuando β_1 es positiva.
125
100
75
50
25
1.00
.75
.50
.25
.20
.15
.10
.05
$|(\delta-2)^3|$
$2(\delta-2)$
-3
-2
-1
0
1
2
3 δ

2º Dividir el tercer momento sigmático entre 2 y cambiarle signo (β_1)

3º Tomar el recíproco de β_1 y restarlo de δ para obtener las (D)

4º Tomar β_1, cambiarle el signo y sumarle el recíproco de β_1 para tener (A)

5º Tomar el recíproco de β_1 y dividirlo entre β_1 para obtener el recíproco del cuadrado o sea (R)

6º Obtener el valor de Y,

A. Obteniendo la función gamma de R (en tablas).
B. Elevando $-\beta_1$ a la R.
C. Multiplicando los dos resultados anteriores.
D. Dividiendo el efectivo entre el producto así obtenido.

7º Substituir el valor de los diferentes parámetros en la fórmula genérica del sistema de curvas que nos ocupa.

8º Dar valores a δ, en dicha ecuación (ya especificada mediante la substitución de los valores de los parámetros) con el fin de obtener los valores teóricos de las frecuencias.

El procedimiento que hemos mostrado aquí en forma detallada, puede sintetizarse como sigue:

Primero. Hay que calcular el tercer momento sigmático.
Segundo. Substituir su valor en la fórmula de beta uno.
Tercero. Substituir el valor de beta uno en las fórmulas de D, A, R.
Cuarto. Substituir los valores de beta uno y R en la fórmula de Y.
Quinto. Substituir los valores encontrados de los parámetros en la fórmula genérica del sistema de curvas asimétricas.

Trataremos, en seguida, de hacer una aplicación del procedimiento anterior a un caso concreto (que Elderton propone y al que aplica fórmulas un tanto distintas de las nuestras en forma menos pormenorizada que la que nos permitimos con propósito didáctico). Los datos son los siguientes: el segundo momento, respecto de la media aritmética, es igual a 1.441787, y el tercer momento respecto a la misma media aritmética es igual a 3.606622. Es decir:

DATOS

$$u_2 = 1.441787 \qquad u_3 = 3.606622$$

Para calcular el tercer momento sigmático, habrá que dividir el tercer momento no sigmático (3.606622) entre la raíz cuadrada del segundo momento ($-\sqrt{1.441787} =$

1.20074) elevada al cubo ($1.20074^3 = 1.20074^2 \times 1.20074 = 1.441787 \times 1.20074 = 1.73120$). Según esto, el tercer momento sigmático será:

$$\gamma_3 = \frac{u_3}{u_2^{3/2}} = \frac{3.606622}{1.73120} = 2.08331 \tag{76}$$

Obtenido este tercer momento, su división entre 2 y el cambio de signo del cociente nos dará β_1

$$\beta_1 = -\frac{2.08331}{2} = -1.0417 \tag{77}$$

El recíproco de beta uno se obtiene dividiendo 1 entre —1.0417, lo cual nos produce:

$$\frac{1}{\beta_1} = \frac{1}{-1.0417} = -0.959969 \tag{78}$$

De acuerdo con esto, si restamos este recíproco de delta minúscula (δ) obtendremos D:

$$D = \delta - \frac{1}{\beta_1} = \delta - (-0.959969) = \delta + 0.959969 \tag{80}$$

Para obtener el valor de A, cambiaremos el signo de β_1 ($- -1.0417 = +1.0417$) y le sumaremos el recíproco de β_1 (—0.959969)

$$A = -\beta_1 + \frac{1}{\beta_1} = 1.0417 - 0.959969 = 0.081731 \tag{81}$$

El valor de R se obtendrá dividiendo el recíproco de β_1 (—0.959969) entre β_1 ($= -1.0417$):

$$R = \frac{1}{\beta_1} = \frac{1/\beta_1}{\beta_1} = \frac{-0.959969}{-1.0417} = 0.92157 \tag{82}$$

El valor de Y se obtendrá substituyendo N, el efectivo, por 251, que es uno de los datos del problema, menos beta uno por $-(-.0417)$ o sea por $+1.0417$ y R por 0.92157.

$$Y = \frac{N}{(-\beta_1)^R \Gamma(R)} = \frac{251}{1.0417^{0.92157}\Gamma(0.92157)} \tag{83}$$

Para el cálculo, habrá que obtener la función gamma mayúscula de 0.92157. Al recurrir a la tabla de funciones gamma, puede observarse, sin embargo, que no está tabulado el valor correspondiente a 0.92157 pues la tabla se extiende de 1 a 2. Sin embargo, dicho valor puede calcularse si se tiene en cuenta lo siguiente:

La función gamma mayúscula de n+1 es igual a n veces la función gamma de n, o bien: la función gamma de n—1 es igual a la función gamma de n entre n—1:

$$\Gamma(R+1) = R\,\Gamma(R) \tag{84}$$

$$\Gamma(R-1) = \frac{\Gamma(R)}{R-1} \tag{85}$$

Si en la primera de estas fórmulas substituimos R por su valor, tendremos:

$$\Gamma(0.92157 + 1) = 0.92157\,\Gamma(0.92157) \tag{86}$$

$$\Gamma(1.92157) = 0.92157\,\Gamma(0.92157) \tag{87}$$

Al despejar a gamma de 0.92157, se obtiene:

$$\Gamma(0.92157) = \frac{\Gamma(1.92157)}{0.92157} \tag{88}$$

La tabla de la función gamma da, como valor de gamma de 1.92 (no tomamos sino dos decimales pues nuestro propósito es mostrar en concreto y en sus varias peripecias un procedimiento, más que obtener los resultados precisos de un caso substantivo), 0.96878. De acuerdo con esto:

$$\Gamma(0.92) = \frac{0.96878}{0.92} = 1.05302 \tag{89}$$

Substituido este valor en el de Y, se tendrá:

$$Y = \frac{251}{1.04^{0.92}\,\Gamma(0.92)} = \frac{251}{1.04^{0.92} \times 1.05}$$

Para obtener el valor de 1.04 a la 0.92 (simplificación de 1.0417 a la 0.92154) recurrimos a los logaritmos:

log. 1.04 = 0.01703

×.92 = 0.01567

antilog. 0.01567 = 1.04 resultado que era de esperar en cuanto el exponente 0.92 es tan cercano a 1.

$$Y - \frac{251}{1.04 \times 1.05} = \frac{251}{1.09} = 230.2$$

Si se considera que el valor de sigma para esta serie es aproximadamente igual a 1.2, el resultado se convierte en 230.2/ 1.2 = 191.8 = 192.

Obtenidos los valores de todos estos parámetros, puede tabulárseles, en su forma simplificada, como sigue:

$$D = \delta + 0.96$$
$$A = 0.08$$
$$\frac{1}{\beta_1} = -0.96$$
$$Y = 192.$$

Es decir, que la serie que hemos considerado representada por un segundo momento central igual a 1.441787 (o, prácticamente igual a 1.44) por un tercer momento central igual a 3.606622 (o, prácticamente, igual a 3.60) y cuyo efectivo era 251, puede expresarse mediante la siguiente fórmula:

$$y = 192 \; [\, e^{\delta + .96} (\delta + .96)^{0.08}]^{-.96} \qquad (92)$$

Al dar valores a δ en la fórmula (92) se obtienen los correspondientes valores de y, o sean las frecuencias con las que aparecen dichas deltas minúsculas, o desviaciones sigmáticas. Los cálculos correspondientes se consignan en el Cuadro 3. Si se quiere saber a qué valores originarios corresponden esas frecuencias, bastará con multiplicar la correspondiente desviación sigmática (δ) por la desviación cuadrática media o sigma minúscula (σ, igual a la raíz cuadrada del segundo momento, o sea a 1.20074) agregando al producto el valor de la media aritmética (que un cálculo previo, que no hemos reproducido por no ser indispensable para nuestros propósitos actuales, mostró era igual a 2.334661).

CUADRO 1. Componentes y Forma de la Curva Asimétrica cuando β_1 es Negativa
$\beta_1 = -.5$

δ	$\delta + 2$	$-2(\delta + 2)$	$e^{-2(\delta +)}$	$(\delta + 2)^3$	Producto	y
—2	0	0	1.000000	0	0.00000	0.0000
—1.5	0.5	—1	0.367879	0.125	0.04590	0.1224
—1	1	—2	0.13534	1.000	0.13534	0.35424
—0.5	1.5	—3	0.049787	3.375	0.16803	0.44808
0	2	—4	0.018320	8.000	0.14655	0.39082
0.5	2.5	—5	0.006738	15.625	0.10528	0.28074
1	3	—6	0.002480	27.000	0.06696	0.17796
1.5	3.5	—7	0.000911	42.875	0.03910	0.10426
2	4	—8	0.000340	64.000	0.02176	0.05802
2.5	4.5	—9	0.000123	91.125	0.01124	0.02997
3	5	—10	0.000050	125.000	0.00625	0.01666
3.5						

CUADRO 2. Componentes y Forma de la Curva Asimétrica cuando β_1 es Positiva $\beta_1 = .5$

δ	$\delta + 2$	$+ (\delta - 2)$	$e^{2(\delta - 2)}$	$(\delta - 2)^3$	Producto	y
—3	—5	—10	0.00005	—125	—0.00625	0.01669
—2.5						
—2	—4	—8	0.00034	—64	—0.02176	0.05870
—1.5	—					
—1	—3	—6	0.00248	—27	—0.06696	0.17878
—0.5						
0	—2	—4	0.01832	—8	—0.14656	0.39131
0.5						
1	1	—2	0.13534	—1	—0.13534	0.36136
1.5						
2	0	0	1.00000	0	0.00000	0.00000
2.5						
3	—1	2	7.3891	1	7.3891	

CUADRO 3. Graduación de la Serie de Frecuencias Ejemplificada, de acuerdo con el Sistema de Curvas Asimétricas presentado en este Artículo

3.1/— Elaboración Previa de los Datos

DATOS DE LA SERIE		CÁLCULO DE LAS DESVIACIONES		TRANSFORMACIÓN
x	f	d	δ	$\delta + 0.96$
1	44	—1.3346612	—1.11	—0.15
2	135	—0.3336612	—0.28	0.68
3	45	0.6653388	0.55	1.51
4	12	1.6653388	1.39	2.35
5	8	2.6653388	2.22	3.18
6	3	3.6653388	3.05	4.01
7	1	4.6653388	3.88	4.84
8	3	5.6653388	4.72	5.68
	251			

CUADRO 3.2. Uso de la fórmula de la Curva Asimétrica para la Graduación Transformación Logaritmica de la fórmula:

$$\log. y = \log. 192 - 0.96\ [\delta + 0.96 \log. e + .08 \log. (\delta + 0.96)]$$

$$D = \delta + 0.96$$

D	Columna I D log. e	log. D	Columna II .08 log. D	Columna III .96 (I + II)	log. 192 — (III)	y
0.68	0.2953172	1.83251	—0.0143992	0.269681280	2.01462	103
1.51	0.6557779	0.17898	0.0143184	0.643292448	1.64001	44
2.35	1.0205815	0.37107	0.0296856	1.008255416	1.27504	19
3.18	1.3810422	0.50245	0.0401944	1.364387136	0.91891	8
4.01	1.7415029	0.60314	0.0482512	1.718163936	0.56514	4
4.84	2.1019636	0.68485	0.0547880	2.071481536	0.21282	2
5.68	2.4667672	0.75435	0.0603480	2.426030592	—0.14373	1
					1.85627	

log. 192 + 2.28330

CUADRO 4.1. Población de la República Mexicana por Grupos Decenales de Edad, según el Censo de 1960

Grupos de Edad Años	Individuos de esa edad en miles	d'	d'f	d'²f	d'³f
De 0 a 9	11 094	—2	—22 188	44 376	—88 752
10 a 19	7 894	—1	—7 894	7 894	—7 894
20 a 29	5 452	0	5 589	0	0
30 a 39	3 972	1	0	3 972	3 972
40 a 49	2 595	2	3 972	10 380	20 760
50 a 59	1 863	3	5 190	16 767	50 301
60 a 69	1 159	4	4 636	18 544	74 176
70 a 79	521	5	2 605	13 025	65 125
80 a 89	781	6	4 696	28 116	168 696
SUMAS	35 331		—3 404	143 074	286 384
MEDIAS			—0.0963	4.0495	8.1057
MOMENTOS			v_1'	v_2'	v_3'

Especificación de la fórmula del sistema de curvas asimétricas para la distribución de la población de la República Mexicana por grupos de edad de acuerdo con el Censo de 1960.

$v_1 = -0.9634$
$u_2 = 404.0250$
$u_3 = 9275.3500$

$$\sqrt{u_2} = 20.1006$$

$$\sqrt{u_2^3} = 404.0250 \times 20.1006 = 8\,121.1449$$

$$\gamma_3 = \frac{u_3}{\sqrt{u_2^3}} = \frac{9\,275.3500}{8\,121.1449} = 1.14$$

$$\beta_1 = -\frac{\gamma_3}{2} = -\frac{1.14}{2} = -.57$$

$$\frac{1}{\beta_1} = \frac{1}{-0.57} = -1.75$$

$$D = \delta - \frac{1}{\beta_1} = \delta - (-1.75) = 0.57 + 1.75 = -1.18$$

$$A = -\beta_1 + \frac{1}{\beta_1} = -(-.57) + (1.75) = .57 - 1.75 = -1.18$$

$$R = \frac{1}{\beta_1{}^2} = (-1.75)^2 = 3.06$$

$$Y' = \frac{35\,331}{0.57^{3.06}\,\Gamma\,(3.06)}$$

$\Gamma(3.06) = 2.06\,\Gamma(2.06) = 2.06 \times 1.06\ \Gamma(1.06) = 2.06 \times 1.06 \times 0.96874 =$
$2.1836 \times 0.97 = 2.118092 = 2.12$

Cálculo de: $0.57^{3.06}$
log. $0.57 - \bar{1} + .7559$
$\times 3.06 \quad -3.06 + 2.3131 = -0.7469 = 9.2531 - 10 = \bar{1}.2531$

$$\text{antilog. } \overline{1}.2531 = 0.179 \doteq 0.18 = 0.57^{3.06}$$

$$Y' = \frac{35\,331}{0.18 \times 2.12} = \frac{35\,331}{0.3816} = 92\,566$$

Como el valor de sigma minúscula es de 20.1006 $\left(= \sqrt{u_2} \right)$ y el intervalo es de 10, la desviación cuadrática media en unidades del intervalo es $20.1006/10 = 2.01$. Por lo mismo, podemos dar como valor de Y:

$$Y = \frac{92\,586}{2.01} = 46\,063$$

De acuerdo con todo esto, la fórmula ya especificada será:

$$y = 46\,063\ (e^{\delta\ +\ 1.75}\ (\delta + 1.75)^{-1.18})^{-1.75}$$

O bien:

$$y = 46\,063\ \ (e^{-1.75\ (\delta\ +1.75)}\ \ (+\ 1.75)^{2.07})$$

CUADRO 4.2. Cálculo de las Frecuencias Teóricas correspondientes a la Curva Asimétrica interpolada a los datos de población de la República Mexicana clasificados por grupos decenales de edad, según el Censo de 1960

Media aritmética: $25 - 0.0963 = 24.9037 \doteq 25$
Desviación cuadrática media: $20.1006 \doteq 20$

Puntos medios de los grupos m Años	m-x	$\frac{m-x}{\sigma}$	$\delta + 1.75$	$-1.75\,D$	$e^{-1.75D}$	$D^{2.07}$	Producto
5	—20	—1	0.75	—1.3125	0.269820	0.5625	0.1512
10	—10	—0.5	1.25	—2.1875	0.111917	1.5625	0.1747
25	— 5	0	1.75	—3.0625	0.046888	3.0625	0.1530
35	+10	+0.5	2.25	—3.9375	0.019644	5.0625	0.1012
45	+20	+1	2.75	—4.8125	0.008148	7.5625	0.0605
55	+30	+1.5	3.25	—5.6875	0.003346	10.5625	0.0317
65	+40	+2	3.75	—6.5625	0.001430	14.0625	0.0141
75	+50	+2.5	4.25	—7.4375	0.000611	18.0625	0.0108
85	+60	+3	4.75	—8.3125	0.000249	22.5625	0.0045

CUADRO 4.3 Frecuencias Teóricas de la Población de México en 1960 de acuerdo con esta graduación

Puntos medios	Frecuencia teórica
5	6 964 726
15	8 047 206
25	7 047 639
35	4 661 576
45	2 786 811
55	1 460 197
65	646 488
75	497 480
85	207 283
	32 322 406