Tratamiento matemático de algunas estadísticas genuinamente españolas

Por **D. César Gómez Lucía,**
Coronel del Ejército del Aire, Actuario de Seguros,
Director-Gerente de IBERIA (Compañía de Navegación Aérea).

TEMA.—*Tratamiento matemático de algunas estadísticas genuinamente españolas.*

INDICE:

PRELIMINARES

PÓRTICO.

El Instituto de Actuarios Españoles, de reciente creación, ha tenido la amabilidad de acoger mi petición de formar parte de él y me ha invitado, como a todos sus miembros, a que aporte algún trabajo científico relacionado con nuestra profesión.

Cuando en la vida se ha doblado el cabo de la juventud, no se tiene vanidad. La experiencia nos enseña a todos cuán difícil es lograr, en el complejo de inteligencia, trabajo y suerte, frutos que puedan servir para beneficio de la Humanidad. Hacer algo que valga la pena de que los demás se paren a mirarlo, y no digamos hacer algo que sea útil para todos, es verdaderamente difícil, y a ello renuncio, desde luego.

Tampoco tengo la virtud de la modestia; y como tengo, en cambio, muy desarrollada, por vocación y ejercicio, la costumbre de la obediencia, cumplo la indicación que se me hace sin detenerme a considerar si mi aportación es o no valiosa ni siquiera si será o no adecuada.

Se ha dado una orden; pues a cumplirla.

ENFOCANDO EL TEMA.

El Actuario es por antonomasia el artífice del complejo aleatorio. Su *misión* es descubrir su espíritu. Su herramienta, la matemática. La masa sobre que opera, la Estadística.

Complejo aleatorio es el conjunto de diversas causas, conocidas o no, cuyos efectos sobre él podemos medir. Estadística es la compilación ordenada y adecuada de los efectos observados.

Una misión del Actuario es diagnosticar: esto es, señalar la ley a que obedecen los efectos. Con ella podrá él, o los que utilicen sus esfuerzos, encontrar en muchos casos la trabazón y cuantía de las causas, esto es, el espíritu del complejo. Otra misión es pronosticar, o sea predecir la cuantía, límite, agrupación, etc., de los efectos. Ambas misiones son en la vida actual de tal *importancia*, que *no* hay ramo de la ciencia que no las necesite. La Biología y la Economía no han sido más que meras tinieblas hasta que se han apoyado en ese estudio de los complejos que es la esencia de la profesión actuarial.

La matemática del Actuario está forjada en el acero del cálculo de la probabilidad. La teoría del cálculo de probabilidades cabe en dos páginas de un libro, pero lleva en sí concepciones filosóficas tan extensas y profundas, tan enmarañadas y a veces contradictorias, que no ya un libro, sino bibliotecas enteras tenemos donde engolfarnos y deleitarnos. No hay ramo de la matemática que apasione más ni que haya subyugado tanto a los grandes matemáticos modernos. Y es que el cálculo de probabilidades tiene alma, mientras que el resto de la matemática es el ingente, inconmovible, pero frío monumento eterno, erigido por el hombre a la verdad física.

El tratamiento matemático de las estadísticas es una sirena. ¿Cómo evitar que se sienta atraído por su canto el neófito o el versado y aun el ajeno, si con su ayuda se han descubierto estrellas y encontrado o comprobado leyes de la físico-química y explicado fenómenos económicos y atisbado ese profundo secreto de la Biología que guarda el Creador para mostrarnos que El es Infinito? A la vocación actual por la ciencia actuarial (sin saber que cultivan tal ciencia), que es un fenómeno universal, ha cooperado esa aparente sencillez y miniatura del *cálculo de probabilidades*. Pero sus resultados son muy diferentes. En el haber de esa cuenta de resultados hay grandes descubrimientos, pero en su debe hay un vasto campo de "cementerios de números", salidos incluso de doctas Corporaciones que tienen la única misión de estudiar las estadísticas. Y es que el pincel con que trazó Velázquez su cuadro de las lanzas, puede, en otras manos no geniales, producir sólo unos chafarrinones que molesten a la vista.

Las anteriores consideraciones os comprobarán el miedo con que me aventuro en la navegación, corta, de cabotaje, que supone este trabajo, en un medio tan lleno, no de escollos, como la navegación de dos dimensiones, ni aun de formación de hielo, como en la de tres dimensiones, a la que estoy acostumbrado, sino de causas y efectos con la que, más aún que en las otras navegaciones, se echa de ver lo fácil que es zozobrar, caer en barrena o fracasar. Y ahora... Vista, suerte y al toro, al mus y a la lotería.

Indice del trabajo.

Siempre me ha dejado de mal humor el leer, aun en libros españoles, que los ejemplos con que se confirman las teorías de la probabilidad son ejemplos exóticos. He ahí la razón de presentaros este trabajo: nacionalizar, aunque sea poquito. Entre las muchísimas estadísticas de mi archivo y entre los trabajos de muchos años de vocación he seleccionado los que precisamente son más castizos, aunque no tengan en sí utilidad práctica. No se hicieron, como otros que pudiera presentar, buscando un fin de la vida real. Se hicieron precisamente por afición, por deleite, para mí mismo.

Dos de ellos los presento como caso de "prognosis"; otros dos, como caso de "diagnosis".

a) *Prognosis.*

En los primeros se trata de complejos aleatorios que obedecen a causas conocidas, en trabazón conocida, y se señalan los efectos que *matemáticamente* deben producir. Uno es el juego del mus. Causas: El que haya 10 oros, 10 copas, etc., y que haya 8 naipes que valen como reyes, 8 que valen como ases, etc... Efectos: Las distintas jugadas que se producen. La compilación de efectos la hice yo mismo en circunstancias que me dejaban tiempo para hacerlo.

El otro, nuestra Lotería Nacional, tiene como causas el número de bolas de dos urnas. Efectos: Los que consigna la "lista grande". Sobre éste tengo muchas estadísticas de comprobación; pero como también las tiene todo el mundo, se las dejo al lector, si es que ve despierta su curiosidad.

b) *Diagnosis.*

En general, en los casos de diagnosis se pretende buscar la ley de la aleatoria. Se conocen ciertos efectos, a veces demasiado pocos, y se trata de encontrar la expresión matemática de una ley que produjera los mismos efectos.

Laplace y Gauss dieron la ley de repartición de los errores, y con ella nació el tratamiento matemático de la estadística. Su fundamento es la normalidad de un hecho y su desviación de lo normal por una masa de pequeños errores en número infinito que siguen una ley. Pero las fórmulas primitivas no satisfacían a todos los casos de la vida práctica, a pesar de los aditamentos de Poisson y Lexis, y se salió del paso achacando las irregularidades de ciertos complejos a observación insuficiente. La estadística dió un paso de gigante cuando Gram demostró que la curva normal de Gauss no era más que un caso particular de un sistema general de frecuencias disimétricas que era el que en realidad aparecía en la práctica. Hoy se trabaja en las estadísticas expresando la probabilidad en una serie convergente cuya función generadora es la de la probabilidad normal o descomponiendo en curvas conocidas la curva de la frecuencia propuesta.

En mi primer trabajo presento un complejo compilado por mí, en mi mocedad, pero estudiado este mismo año, al ponerse muy en boga el peso de los toros de lidia. No pretendía nada con el estudio, pero me encariñé con el asunto al irle tratando, y lo he hecho prolijamente.

El segundo trabajo es una estadística de mi niñez. Cuando se compiló, yo no sabía escribir, pero asistía a veces a medir los efectos del complejo. Es una estadística de mi archivo, heredada. Es el resultado de la tienta de vacas madres en la ganadería de mi abuelo, D. Félix Gómez, por los años del 86 al 96. La desempolvo, porque es la más aguda o disimétrica de las que poseo y requiere tratamientos más específicos. Ya sé la objeción que me van a hacer algunos si no leen antes estas páginas, y es: Que se trata de una estadística irreal, contenida, como la que se produce en los fenómenos de economía, por el freno a la demanda, y que, de no existir este tope, sería una estadística que seguiría la Ley de Gauss. Es verdad—les podría yo contestar—. Este tipo de complejos se da, en el caso de mi estudio, porque no se dejaba tomar a las vacas más de seis puyazos, y se da, modernamente, por ejemplo, en las estadísticas de utilización de líneas de tráfico aéreo, porque los aviones no llevan más que diecisiete plazas y la demanda es mayor. Es cierto que si las vacas hubieran seguido tentándose hasta caer muertas, hubiese sido otro el resultado; como es cierto que la estadística de tráfico sería otra si el avión tuviese cuarenta plazas o si se compilase en la estadística no los billetes vendidos, sino los pedidos. Pero ello no niega que en la vida práctica no se presenten complejos tan disimétricos. Prescindiendo de los ejemplos tan abundantes de casos análogos que ofrece la Biología, tengo una estadística real, aún muy pequeña para presentarse, que ofrece las mismas características. Es la rotura de material de vuelo en los servicios regulares de tráfico aéreo en todo el Mundo. Tomando como casos posibles de damnificación del material, por accidente, el porcentaje desde el uno al cien del valor del avión, vemos que hoy día hay una pequeña frecuencia de accidentes, con roturas hasta del 5 por 100 del valor del avión. La frecuencia casi se anula en los demás casos, toma valor tangible en roturas próximas al 90 por 100 y adquiere su valor máximo en la rotura total, en la catástrofe. La media de la serie está en el valor 96 por 100. El modo de ser actual de la navegación aérea es así. Una seguridad cercana a la certeza que, cuando falla, no tiene paliativos.

PROGNOSIS

I.—El juego del mus.

El popularísimo y castizo juego del mus tiene once jugadas diferentes incompatibles entre sí. Estas jugadas se definen por combinaciones formadas por los números de los naipes de cuatro cartas que constituyen el lote de cada jugador, en una baraja de cuarenta unidades. Estas jugadas son las que se expresan en la columna (1) del cuadro resumen de la experiencia, donde están ordenadas por probabilidad de aparición, según resulta de la teoría.

Las experiencias se hicieron anotando la jugada que resultaba al repartir las cuatro primeras cartas de la baraja a un jugador, volviendo a barajar y cortar después de esta operación. Así se anotaron cuatro mil jugadas, que después se agruparon de dos maneras diferentes: una, en cuarenta series de cien pruebas cada una, y otra, en cien series de cuarenta pruebas cada una.

En la columna (5) del cuadro resumen está el número de veces que se obtuvo cada jugada.

La probabilidad de obtención de cada jugada es la que se expresa en la columna (3) del cuadro. La columna (4) es el resultado de aplicar la probabilidad correspondiente al total de los cuatro mil casos estudiados, o sea la frecuencia teórica de la presentación de las jugadas.

Para obtener esta probabilidad se ha formado la columna (2) de la manera que luego diremos, y que es el número de casos favorables a la presentación del suceso, o sea el numerador de la fracción de la probabilidad. El denominador, que es el número de casos posibles, es siempre el mismo para todas las jugadas e igual al producto de $40 \times 39 \times 38 \times 37$, o sea el producto del número de naipes que hay en la baraja al tirar cada carta. Este denominador, igual a 2.193.360, tiene que ser la suma de los números de la columna (2) y sirve de comprobación de la operación.

Cálculo de la probabilidad.

I

La jugada "ni juego ni pares" se puede dar: a) sin figuras; b) con una figura; c) con dos figuras.

El caso a) se puede dar con ases y sin ases. Luego la cifra de casos favorables será:

Con ases.—El producto de los naipes ases $= 8$; por los naipes no figuras ni ases $= 16$; por los naipes que queden de esta especie después de la salida del primero y que no sean de su clase $= 12$; por los naipes que queden de la especie después de la salida del primero y del segundo y no sean de su clase $= 8$; por los casos de aparición del as atendiendo al orden, que puede ser 4.

Sin ases.—El producto de los naipes no figuras ni ases (el cuatro, el cinco, el seis y el siete) $= 16$; por los naipes de la especie que queden después de la salida de una clase de ellos (o sea, por ejemplo, el seis, el siete y el cuatro si salió primeramente un cinco) $= 12$; por los restantes $= 8$; por los restantes $= 4$. Y así tenemos

$$a) = (8 \times 16 \times 12 \times 8 \times 4) + (16 \times 12 \times 8 \times 4) = 2^{11} \times 27 = 55296$$

El caso b) se puede dar con ases y sin ases.

Con ases.—El producto de los naipes figuras $= 16$; por el de naipes ases $= 8$; por el de naipes ni figuras ni ases $= 16$; por el de naipes que queden de esta especie y que no sean de la clase del aparecido, ya que no puede haber pares ni más de una figura $= 12$; por el número de combinaciones en que pueden aparecer estos naipes $= 12$.

Sin ases.—El producto de naipes figuras $= 16$; por el de naipes no figuras ni ases $= 16$; por los que quedan de esta especie y no sean de la clase del primero que ha salido $= 12$; por los restantes, siguiendo el mismo proceso $= 8$; por el número de casos de orden diferente de aparición de la figura $= 4$. Y así tenemos

$$b) = (16 \times 8 \times 16 \times 12 \times 12) + (16 \times 16 \times 12 \times 8 \times 4) = 2^{17} \times 3 = 396216$$

El caso c) puede darse:

Con rey y figura, con as.—El producto de naipes reyes $= 8$; por el de naipes figuras y no reyes $= 8$; por el de naipes ases $= 8$; por el de naipes no figuras ni ases $= 16$; por el de las combinaciones posibles de aparición que es factorial de $4 = 24$. O sea $= 2^{16} \times 3$.

Con rey y figura, sin as.—El producto de naipes reyes $= 8$; por el de figuras no reyes $= 8$; por el de naipes no figuras, ni ases, ni sietes (porque éstos darían forzosamente juego con el cuarto naipe) $= 12$; por el de los de esta especie y que no sean de la clase del que ha aparecido $= 8$; por el número de combinaciones de aparición $= 8$. O sea $= 2^{14} \times 3$.

Sin rey y con as.—El producto de naipes caballos $= 4$; por el de naipes sotas $= 4$; por el de ases $= 8$; por el de naipes no ases ni figuras $= 16$; por el número de combinaciones de aparición $= 24$. O sea $= 2^{14} \times 3$.

Sin rey y sin ases.—El producto de naipes caballos $= 4$; por el de naipes sotas $= 4$; por el de naipes cuatro, cinco o seis (el siete daría lugar a juego con el cuarto naipe) $= 12$; por el resto de la especie, después de eliminar la clase aparecida $= 8$; por el número de combinaciones de aparición $= 8$. O sea $= 2^{12} \times 3$. Y así tendremos

$$c) = 2^{12} \times 3 \times 25 = 307.200.$$

Resultando para la jugada "ni juego ni pares"

$$a) + b) + c) = 755.712$$

II

La jugada "pares sin juego" se puede lograr: a) por pares de reyes, entrando ases y sin entrar ases. b) por pares de figuras no reyes, entrando ases y sin entrar. c) por pares de ases, entrando dos figuras, una de ellas rey. d) por pares de ases, entrando dos figuras, ninguna de ellas rey. e) por pares de ases, entrando una figura; f) por pares de ases, no entrando ninguna figura. g) por pares de cuatros o de cincos, en los mismos casos que la pareja de ases, o sea con dos figuras, con rey y figura, con figura o sin figura. h) por pares de seises o de sietes, pero sólo en los casos de una figura, sea o no rey, y sin figura.

El cálculo es el siguiente:

a) El producto del número de reyes $= 8$; por el de reyes que queden después de salir el primero $= 7$; por el número de ases $= 8$; por el de naipes que no sean ases ni figuras $= 16$; por el de combinaciones de orden $= 12$. Todo esto en el caso de que entren ases.

Cuando no entran ases, los naipes que no son reyes no pueden ser más que el cuatro con el cinco o el seis, y, por tanto, resultará: a) $= 21\ (2^{10} + 2^{12})$.

El cálculo de los otros apartados es el siguiente:

b) $= (8 \times 3 \times 8 \times 16 \times 12) + (8 \times 3 \times 4 \times 8 \times 12) = 9\ (2^{10} + 2^{12})$

c) $= (8 \times 7 \times 8 \times 8 \times 12)$

c) $= (8 \times 7 \times 8 \times 8 \times 12)$

d) $= (8 \times 7 \times 8 \times 4 \times 6)$.

e) $= (8 \times 7 \times 16 \times 16 \times 12)$

f) $= (8 \times 7 \times 16 \times 12 \times 6)$.

g) $= (8 \times 3 \times 8 \times 8 \times 12) + (8 \times 3 \times 8 \times 4 \times 6) +$
$+ (8 \times 3 \times 16 \times 20 \times 12) + (8 \times 3 \times 8 \times 12 \times 12) +$
$+ (8 \times 3 \times 12 \times 8 \times 6)$

El primer sumando corresponde a pares de cuatros o de cincos, rey y figura. El segundo es la variante de figuras no reyes. El tercero es el caso de una sola figura. El cuarto es el caso de que entren ases sin entrar figura. El quinto es de que no entren ases ni figuras.

h) $= (8 \times 3 \times 16 \times 20 \times 12) + (8 \times 3 \times 8 \times 12 \times 12) +$
$+ (8 \times 3 \times 12 \times 8 \times 6)$

Resultando para la jugada "pares sin juego" 734.208.

III

La jugada "pares y juego" se puede lograr:

a) Con cuatro figuras, siendo los pares de reyes o de figuras:

$$(8 \times 7 \times 4 \times 4 \times 12) + (8 \times 3 \times 8 \times 4 \times 12)$$

b) Con tres figuras en las mismas variantes del anterior:

$$(8 \times 7 \times 8 \times 16 \times 12) + (8 \times 3 \times 12 \times 16 \times 12)$$

c) Con dos figuras siendo los pares de reyes, de figuras, de sietes o seises, entrando rey y figura; de sietes o seises, sin entrar rey. En los dos primeros casos, los naipes que no se apareen han de ser forzosamente el siete, el seis o el cinco, y resulta:

$$(8 \times 7 \times 4 \times 8 \times 12) + (8 \times 3 \times 4 \times 8 \times 12) +$$
$$+ (8 \times 3 \times 8 \times 8 \times 12) + (8 \times 3 \times 4 \times 4 \times 12)$$

Resultando, pues, la jugada $=$ 215040.

IV

La jugada "juego sin pares" se logra:

a) Con tres figuras, sin entrar ningún as:

$$(8 \times 4 \times 4 \times 16 \times 24)$$

b) Con dos figuras en las variantes rey y figuras; o figuras sin rey, siendo los otros naipes forzosamente el cinco, seis o siete:

$$(8 \times 8 \times 4 \times 8 \times 24) + (4 \times 4 \times 4 \times 8 \times 24)$$

Resultando, pues, la jugada $=$ 110592.

V

La jugada "treinta y una y pares" se logra:

a) Con pares de reyes, un cinco y un seis, o pares de reyes, un cuatro y un siete:

$$(8 \times 7 \times 4 \times 4 \times 12) \times 2$$

b) Con pares de figuras no reyes y las circunstancias anteriores:

$$(8 \times 3 \times 4 \times 4 \times 12) \times 2$$

c) Con pares de reyes, figura y as:

$$(8 \times 7 \times 8 \times 8 \times 12)$$

d) Con pares de figuras no reyes, rey y as, o figura y as:

$$(8 \times 3 \times 8 \times 8 \times 12) + (8 \times 3 \times 8 \times 4 \times 12)$$

Resultando, pues, la jugada = 101376.

VI

La jugada "treinta y una sin pares" se logra:
a) Por rey y figura; cinco y seis, o cuatro y cinco:

$$(8 \times 8 \times 4 \times 4 \times 24) \times 2$$

b) Por figuras sin rey y las circunstancias del caso anterior:

$$(4 \times 4 \times 4 \times 4 \times 24) \times 2$$

c) Por tres figuras y un as:

$$(8 \times 4 \times 4 \times 8 \times 24)$$

Resultando, pues, la jugada = 86016.

(1) Jugadas	(2) Casos favorables	(3) Probabilidad	(4) Teoría	(5) Práctica	(6) Tanteo
Ni juego ni pares	755712	0'344545	1378	1340	0
Pares sin juego	734208	0'334741	1339	1323	1
Pares y juego	215040	0'098041	392	382	3
Juego sin pares	110592	0'050421	202	205	2
31 y pares	101376	0'046219	185	187	4
31 sin pares	86016	0'039216	156	183	3
Duplex sin juego	61488	0'028034	112	112	3
Medias sin juego	55296	0'025211	101	100	2
Juego y medias	37632	0'017157	69,5	92	4
Juego y duplex	22176	0'010110	40,5	48	5
31 y medias	13824	0'006302	25,	27	5
	2193360	0'999997	4000,0	4000,0	

VII

La jugada "duplex sin juego" se presenta en los siguientes casos:

a) Ases-ases, ases-reyes, ases-los seis naipes restantes:

$$(8 \times 7 \times 6 \times 5) + (8 \times 7 \times 8 \times 7 \times 6) + (8 \times 7 \times 4 \times 3 \times 6) \times 6$$

b) Reyes-cincos o cuatros:

$$(8 \times 7 \times 4 \times 3 \times 6) \times 2$$

c) Cuatro, cinco, seis o siete-consigo mismo:

$$(4 \times 3 \times 2 \times 1) \times 4$$

d) Cuatros-con los cinco naipes restantes no ases ni reyes:

$$(4 \times 3 \times 4 \times 3 \times 6) \times 5$$

e) Cincos-seises, sietes, sotas, caballos:

$$(4 \times 3 \times 4 \times 3 \times 6) \times 4$$

f) Seises-sietes:

$$(4 \times 3 \times 4 \times 3 \times 6)$$

Resultando para la jugada "duplex sin juego" = 61488.

VIII

La jugadas "medias sin juego" se presenta en los siguientes casos:

a) De ases, con cualquier otro naipe:

$$(8 \times 7 \times 6 \times 32 \times 4)$$

b) De cuatros, cincos o seises, con cualquier otro naipe:

$$(4 \times 3 \times 2 \times 36 \times 4) \times 3$$

c) De sietes, con naipes no figura:

$$(4 \times 3 \times 2 \times 20 \times 4)$$

Resultando para la jugada "medias sin juego" = 55296.

IX

La jugada "juego y medias" se obtiene:

a) Con medias de reyes o de figuras y un naipe que no dé treinta y uno ni duplex:

$$(8 \times 7 \times 6 \times 24 \times 4) + (4 \times 3 \times 2 \times 28 \times 4) \times 2$$

Resultando para "juego y medias" = 37632.

X

La jugada "duplex y juego" puede ser:

a) De reyes o figuras-consigo misma:

$$(8 \times 7 \times 6 \times 5) + (4 \times 3 \times 2 \times 1) \times 2$$

b) De reyes-figuras y de figuras-figuras:

$$(8 \times 7 \times 8 \times 3 \times 6) + (4 \times 3 \times 4 \times 3 \times 6)$$

c) De reyes-sietes o seises y de figuras-sietes o seises:

$$(8 \times 7 \times 4 \times 3 \times 6) \times 2 + (8 \times 3 \times 4 \times 3 \times 6) \times 2$$

Resultando para "duplex y juego" = 22176.

La jugada "treinta y una y medias" se logra:

a) Con medias de figuras o reyes y un as:

$$(8 \times 7 \times 6 \times 8 \times 4) + (4 \times 3 \times 2 \times 8 \times 4) \times 2$$

b) Con medias de sietes y figura:

$$(4 \times 3 \times 2 \times 16 \times 4)$$

Resultando para "treinta y una y medias" = 13824.

Concordancia entre la teoría y la práctica.

La comparación de las columnas (4) y (5), que expresan, respectivamente, la frecuencia teórica sobre los 4.000 casos y la frecuencia obtenida, señala una concordancia altamente satisfactoria. Las disparidades más acusadas lo son en las jugadas "juego y medias", "treinta y una sin pares" y "juego y duplex", cuya desviación vamos a medir.

Para la jugada "juego y medias", la probabilidad es $p = 0{,}017157$. La probabilidad contraria es $q = 0{,}982843$. El producto $p\ q = 0{,}16711$. El cuadrado de la dispersión $= n\ p\ q$ será $= 66{,}844$, lo que da un desvío medio cuadrático de 8,1. Siendo la media teórica 69,5 y la media obtenida 92,0, la dispersión resulta de 22,50, que es un poco superior a dos veces y media el desvío medio cuadrático que, como es sabido, se ha tomado siempre como unidad de dispersión.

Los mismos razonamientos, aplicados a la jugada "treinta y una sin pares", nos dan como resultado $n\ p\ q = 153{,}6$. Desvío medio cuadrático = 12,4. Media teórica, 156. Media obtenida, 183. Desviación, 27,0, que es un poco mayor de dos unidades dispersión.

En la jugada "juego y duplex" tenemos $n\ p\ q = 39{,}6$. Desvío medio cuadrático, 6,3. Media teórica, 40,5. Media obtenida, 48. Desviación, 7,5, que es 1,25 unidades dispersión.

Los demás resultados son altamente satisfactorios, y habida en cuenta la relatividad de las jugadas dispares a las acordes, resulta la tirada de una gran ejemplaridad.

La columna (6) del cuadro estadístico, de la que aún no habíamos hablado, expresa el tanteo o premio del jugador; es, en realidad, la esperanza matemática de cada jugada. Una visual sobre la columna muestra que el tanteo no corresponde más que muy someramente a la teoría del juego. Se podía garantizar que el juego del mus no es invención de un gran matemático, como pasa con la ruleta, obra del gran Pascal. La jugada más ventajosa por su tanteo resulta la de "pares y juego", lo cual sin saber matemáticas saben perfectamente todos los jugadores de mus. El ir a un descarte de naipes con una jugada en el reparto inicial de "pares y juego" es muy desventajoso para el jugador.

Para continuar el estudio de la concordancia entre la teoría y la práctica, vamos a coger los resultados de la jugada "juego sin pares", que se ha dado 205 veces y que, como dijimos al principio, se ha agrupado en dos maneras diferentes, A) y B). La primera, de cien series de cuarenta jugadas, y la segunda de cuarenta series de cien jugadas, tal y como se expresa en el siguiente cuadro:

JUEGO SIN PARES

A) Número de veces	A) Frecuencia	A) Total	B) Número de veces	B) Frecuencia	B) Total
0	11	0	0	0	0
1	28	28	1	0	0
2	29	58	2	3	6
3	16	48	3	5	15
4	10	40	4	8	32
5	5	25	5	10	50
6	1	6	6	6	36
7	0	0	7	3	21
8	0	0	8	2	16
9	0	0	9	2	18
			10	0	0
			11	1	11
			12	0	0
			13	0	0
			14	0	0
	100	205		40	205

El cálculo de la frecuencia teórica o probabilidad le podemos hacer directamente. Siendo la probabilidad de "juego sin pares" 0,050421, en cada serie de cuarenta pruebas el suceso se podrá dar desde cuarenta veces a cero veces y la probabilidad de que se dé un número *n* de veces es

$$(0{,}050421)^{n} \times (0{,}949579)^{40-n} \times \binom{40}{n}$$

El desarrollo del binomio $(0{,}05 + 0{,}95)^{40}$ nos dará la probabilidad de ocurrencia de 40; 39; ... 0 casos en cada serie de las cien que constituyen la experiencia A). Del mismo modo, el desarrollo del binomio $(0{,}05 + 0{,}95)^{100}$ nos dará la probabilidad de ocurrencia de 100; 99; 98; ... 0; casos en cada serie de las cuarenta de la experiencia B).

CÁLCULO DE LA PROBABILIDAD PARA EL CASO A)

Veces (n)	0	1	2	3	4	5	6
$\log \binom{40}{n}$	0,00	1,602060	2,892095	3,994757	4,960899	5,818231	6,55584
$\log q^{40-n}$	$\bar{1}{,}108960$	$\bar{1}{,}131236$	$\bar{1}{,}153512$	$\bar{1}{,}175768$	$\bar{1}{,}198049$	$\bar{1}{,}220326$	$\bar{1}{,}24260$
$\log p^{n}$	$\bar{0}{,}000$	$\bar{2}{,}698970$	$\bar{3}{,}397940$	$\bar{4}{,}096910$	$\bar{6}{,}795880$	$\bar{7}{,}494850$	$\bar{8}{,}19382$
log probabilidad.	$\bar{1}{,}108960$	$\bar{1}{,}432266$	$\bar{1}{,}443547$	$\bar{1}{,}267435$	$\bar{2}{,}954828$	$\bar{2}{,}533407$	$\bar{3}{,}992260$
Probabilidad.	0,12852	0,27050	0,27760	0,18511	0,090121	0,034151	0,009823
Frecuencia obtenida en 100 tiradas.	11	28	29	16	10	5	1

La concordancia, como se ve comparando las dos últimas líneas, es completamente satisfactoria.

CUADRO COMPARATIVO DE LA PROBABILIDAD Y LA FRECUENCIA OBTENIDA EN EL CASO B)

Veces (n)	Valor de la probabilidad $(0{,}95)^{100-n}_{x}(0{,}05)^{n}\binom{100}{n}$	Frecuencia teórica para las 40 pruebas	Frecuencia en la práctica
0	0,0059	0,236	0
1	0,0312	1,248	0
2	0,0812	3,248	3
3	0,1396	5,584	5
4	0,1781	7,124	8
5	0,1800	7,200	10
6	0,1500	6,000	6
7	0,1060	4,240	3
8	0,0649	2,596	2
9	0,0349	1,396	2
10	0,0167	0,668	0
11	0,0072	0,288	1
12	0,0028	0,112	0
	0,9986	39,940	40

La concordancia menos completa que en el caso anterior es, sin embargo, satisfactoria. Ha de tenerse presente que en el caso A) se estudian 100 tiradas y en el B) 40 solamente, lo que hacía prever una concordancia menos perfecta, de acuerdo con la teoría general de la probabilidad.

Como es bien sabido, no hay necesidad de calcular directamente la probabilidad exacta, por el desarrollo del binomio, si nos valemos de la fórmula de Poisson, de aproximación suficiente para la vida práctica. Las tablas de Poisson, publicadas por *Biometrika,* dan valores muy aceptables cuando se trata de series homogradas discretas, en que las observaciones son, como ahora, números siempre enteros.

A continuación exponemos, para que se juzgue la concordancia entre la teoría y la experiencia, el cuadro correspondiente a la jugada "juego y pares". Como consta en el cuadro de la experiencia, esta jugada se dió 382 veces en la serie de 4.000 tiradas, y catalogadas en 40 series de 100 experiencias, su resultado es el que figura en el cuadro siguiente, en las tres primeras columnas. La columna (4) es la probabilidad aproximada, copiada de las tablas de Poisson para el desarrollo—por su fórmula—del binomio $(0{,}1 + 0{,}9)^{100}$, muy cercano al auténtico, que sería $(0{,}098 + 0{,}902)^{100}$, ya que la probabilidad de la jugada es, según hemos deducido, 0,98041. La columna (5) nos da la frecuencia teórica multiplicando la (4) por cuarenta. La concordancia buscada se estudia por la comparación de las columnas (2) y (5).

JUEGO Y PARES

(1) Número de casos	(2) Frecuencia	(3) 1 × 2	(4) Tabla Poisson	(5) Frecuencia teórica
3	0		0076	0,30
4	1	4	0189	0,76
5	3	15	0378	1,51
6	2	12	0630	2,52
7	3	21	0901	3,60
8	5	40	1125	4,50
9	7	63	1251	5,00
10	6	60	1251	5,00
11	4	44	1137	4,54
12	3	36	0948	3,80
13	2	26	0729	2,92
14	2	28	0521	2,08
15	0	0	0347	1,38
16	1	16	0217	0,87
17	1	17	0128	0,51
18	0		0071	0,29
19	0		0037	0,14
	40	382	0,9936	39,72

PROGNOSIS

II.—La Lotería Nacional.

La Lotería Nacional no es solamente uno de los ingresos más saneados de nuestra Hacienda; es también una organización estatal de las más perfectas. Su contextura y mecanismo lo están copiando muchas naciones. Prescindiendo de la discusión de orden moral que desde su origen ha suscitado, todos coincidimos en afirmar que si de repente se suprimiese, una gran masa de españoles tendría un verdadero y profundo disgusto. La Lotería es de las cosas que más fundadamente puede apellidarse nacional.

Las probabilidades de premio en los sorteos son fácilmente calculables. Precisamente todos los cálculos de probabilidad y las teorías de la misma se han deducido desde sus primeros tiempos basándose en el juego de dados o en la extracción de bolas de una o varias urnas. La concordancia entre la teoría y la experiencia, que es precisamente el objeto de este estudio, la puede hacer fácilmente todo el mundo sin necesidad de que se presente ninguna estadística, que siempre pudiera parecer amañada. La rapidez con que se suceden los sorteos, uno cada diez días, y el registro que queda de los mismos en publicaciones oficiales, permite que sea auténticamente una "ley de grandes números" la que se comprueba con la Lotería Nacional.

Los sorteos, aunque tienen análogas características, no son todos idénticos ni lo han sido a través de los tiempos. Por esta razón, en el ejemplo que ponemos hemos cogido un número de billetes y de premios que está en concordancia con la estructura de la Lotería Nacional, aunque no coincida en el momento actual con ninguno de los tres tipos de sorteo que existen ahora.

Suponemos un sorteo entre 40.000 números de 1.600 premios. Si tomamos como disyuntiva el que se premie un número de un millar cualquiera, que es como se agrupan en la lista oficial, o el de los 39 millares restantes, la aleatoria propuesta vendrá definida por el desarrollo del binomio $\left(\frac{1}{40}+\frac{39}{40}\right)^{1600} = 1$ El millar podrá ser premiado un número de veces que va de 0 a 1.600. Lo cual podría ser si el número premiado se volviese a introducir en el bombo o urna de los números; pero desde el momento en que el número extraído no vuelve a la urna, *sólo podrá* ser premiado el millar de 0 a 1.000 veces.

En la lista oficial del resultado de la Lotería se agrupan los premios por millares. La comparación entre la teoría y la experiencia se va a hacer por el número de premios que tienen los millares de un sorteo o los de varios.

CÁLCULO DE LA PROBABILIDAD.

a) El valor medio del número de premios que tiene cada millar es igual a 40.

b) Para medir la dispersión de la serie tomamos como incremento de desvío el aumento o disminución de un premio en el número de los premios de cada millar. Así formaríamos la serie natural de los números desde 0 hasta 1.000 como serie de los sucesos que pueden presentarse o eje natural de los desvíos. En vez de referir el origen de los desvíos al valor 0, le referimos al valor 40 y formamos las columnas (1) del cuadro de la probabilidad. En la primera, figuran los números 41 al 60, que son los premios que tienen los millares en el intervalo que consideramos y son los desvíos positivos. En la columna (1) de la derecha figuran los premios de 39 a 20 y son los desvíos negativos. La pequeñez de la probabilidad para números fuera del grupo 20 al 60, aconseja no tomarlos en cuenta.

c) El desvío medio cuadrático aplicando la fórmula del caso en que la bola extraída se retira de la urna, es

$$\bar{0}^2 = 1600 \times \frac{40-1}{1600}\left(1-\frac{1600-1}{40000-1}\right) = 39(1-0{,}04) = 37{,}44\ \bar{0} = 6{,}1282$$

Tomamos como unidad de desvío y la llamamos dispersión a este valor encontrado y formamos la columna (2), que tiene los valores de z para cada uno de los casos. Así, por ejemplo, el caso de 45 premios o su simétrico en desvío el de 35 premios, tendrán una dispersión de 5 dividido por $\bar{0}$ igual a 0,82. La columna (2), aunque está toda ella con signo positivo, lleva el signo positivo para los desvíos positivos y el negativo para los negativos.

d) La probabilidad que corresponde para cada valor de z está expresada en la columna (3), obtenida según la fórmula muy aproximada de La Place. Esta fórmula solamente vale para el caso de que la probabilidad sea igual a $\frac{1}{2}$ o próximo a ella; es decir, al caso en que p sea igual a q. Es una función simétrica y, por lo tanto, la columna (3) vale para los números de la misma línea de las dos columnas (1). Los valores que en ella figuran están copiados de las tablas con seis decimales publicadas por *Biometrika* y se ha prescindido de las dos últimas cifras. La comprobación puede hacerse por la suma total de la probabilidad, que debe ser igual al desvío medio cuadrático, ya que le hemos tomado por unidad. Efectivamente, la suma de 0,40, probabilidad del caso "cuarenta premios", más el doble de la columna (3), da la cifra de 6,1372.

(1) Premios	(2) z	6 (p) 3+4+5	(3) φ_0 (z)	(4) $-C_3\,\varphi_3$ (z)	(5) $C_4\,\varphi_4$ (z)	6 (n) 3+4+5	(1) Premios
41	+0,16	+0,3902	+0,3938	—0,0047	+0,0011	+0,3996	39
42	33	3695	3776	90	9	3875	38
43	49	3423	3537	119	5	3661	37
44	66	3076	3209	135	2	3346	36
45	82	2776	2912	134	—0,0002	3044	35
46	98	2341	2468	122	5	2585	34
47	1,14	1976	2083	101	6	2178	33
48	30	1634	1714	73	7	1780	32
49	46	1323	1374	44	7	1411	31
50	63	1038	1059	15	6	1068	30
51	79	802	805	+0,0002	5	798	29
52	96	604	584	23	3	558	28
53	2,12	455	423	34	2	387	27
54	29	327	290	37	0	253	26
55	45	236	198	38	+0,0000	160	25
56	62	161	128	32	1	97	24
57	78	112	84	27	1	58	23
58	95	73	51	21	1	31	22
59	3,10	48	33	14	1	20	21
60	25	33	20	12	1	9	20
		2,8035	2,8686			2,9315	

Prueba del reparto de la probabilidad.

$\overline{0} = 6{,}1282$
$\Sigma\, \varphi_z\,(2) = 2 \times (3) + 0{,}40 = 6{,}1372$
Σ total $= 6\,(p) + 6\,(n) + 0{,}4001 = 6{,}1351$

La probabilidad corregida se va aproximando a la verdadera, no habiéndose obtenido la exacta $= 6{,}1282$, por no haber empleado en la corrección más que tres términos del desarrollo en serie.

e) Cuando, como en el caso actual, no es aplicable la fórmula de La Place, se puede seguir el método de Charlier, que expresa la probabilidad en función de la fórmula generadora de La Place y sus distintas derivadas, que dan una serie muy convergente. O sea, que el desarrollo de $(p + q)^{s} \overset{s}{=} 1$ puede expresarse en la forma $\Sigma\, c_i\ \varphi_i\ (2)$ para valores $i = 0; 1; 2; \ldots$, que queda reducida después de la transformación de origen y unidad que hemos hecho en $\varphi_0\,(z) + C_3\,\varphi_3\,(z) + C_4\ \varphi_4\,(z) + \ldots + C_n\ \varphi_n\,(z) + \ldots$

En la cual $\varphi_3\,(z)$; $\varphi_4\,(z)$; ... son las tercera, cuarta, etc., derivadas de la función normal, cuyos valores se encuentran en tablas de cuatro cifras decimales publicadas por el Instituto danés de Actuarios y por otros editores de Suecia y Alemania.

Los coeficientes, limitándonos a los tres primeros términos de la serie, son los conocidos ya por el mundo estadístico con los nombres de disimetría, el que se aplica a la tercera derivada; y exceso el que se aplica a la cuarta derivada, y resultan ser:

$$c_3 = -\frac{q - p}{6\,\overline{0}} = -\frac{\frac{39}{40} - \frac{1}{40}}{6 \times 6{,}12} = -\frac{38}{1468{,}8} = -\,0{,}0258$$

$$c_4 = +\frac{1}{24} \times \frac{1 - 6\,pq}{\overline{0}^2} = \frac{0{,}86}{898{,}56} = +\,0{,}00095$$

Con ellos y las derivadas de que hemos hablado antes se han formado las columnas (4) y (5). La suma de las columnas (3), (4) y (5), con los signos que las corresponden, forman las columnas (6), que dan la probabilidad de los casos que figuran en su línea correspondiente.

Los signos de la columna (4) corresponden a los desvíos positivos y son contrarios a los que corresponden a los desvíos negativos. Los signos de la columna (5) son válidos para los desvíos positivos y negativos. Esto es consecuencia de que las derivadas de orden impar afectan a la disimetría de la curva de probabilidad, mientras que las derivadas de orden par sólo afectan a la probabilidad en todos los casos en su cuantía. Cuando p es igual a q, los coeficientes de orden impar se anulan y la probabilidad es simétrica.

Las cifras de la columna (6) han de dividirse por 6,1282, que es el valor de la dispersión, para obtener la cifra en tanto por ciento. Como comprobación de la operación se tiene: que la suma de las dos columnas (6), más 0,4001, que es igual a 6,1351, debía ser igual al valor de la dispersión.

Así, pues, la frecuencia teórica de los millares premiados con 45 premios en la serie de cinco sorteos, sería: $200 \times 0{,}2776 : 6{,}1282 = 9$, que puede compararse con la frecuencia práctica.

Concordancia de la teoría con la práctica.

La concordancia puede establecerse reuniendo los resultados de varios sorteos idénticos en cuanto a billetes y premios. Si la concordancia se estudiase para un único sorteo, ya que se trata solamente de 40 experiencias en el caso del ejemplo o de un número un poco mayor o menor en el del sorteo elegido, se deben agrupar los casos "número de premios", en lugar de considerarlos uno a uno, y así se puede comparar mejor con el cuadro de la probabilidad. Por ejemplo: la frecuencia teórica de los premios de 46 a 50 sería la suma de las cinco probabilidades de los cinco casos expresados en la columna (6), divididas por la unidad dispersión y multiplicadas por 40, que es el número de las experiencias: $0{,}312 : 6{,}1282 \times 40 = 5{,}42$. O sea que una concordancia perfecta nos daría, en el sorteo, cinco o seis millares con un número de premios desde 46 a 50 (ambos incluídos), los cuales podrían ser, por ejemplo: cinco casos de 47 premios y ninguno de 46, 48, 49 y 50; un caso de 46, otro de 47, ninguno de 48, tres de 49 y uno de 50, etc., etc.

La mejor manera de estudiar someramente la concordancia es agrupar los casos de tres en tres o de cinco en cinco y hallar la probabilidad de cada grupo, que será la suma de la de cada componente del grupo, como hemos dicho en el párrafo anterior, o agruparlos en múltiplos de unidades dispersión. Así, si tomamos el grupo 0,5 unidades dispersión, resultaría en nuestro ejemplo 41, 42, 43; 44, 45, 46; y así sucesivamente de tres en tres. Esta agrupación en función del valor de $\overline{0}$ "unidad de dispersión" tiene la ventaja de que se pueden utilizar tablas que contienen la probabilidad ya calculada para submúltiplos de $\overline{0}$ y se hace la concordancia en un golpe de vista.

Naturalmente que lo primero que hay que establecer es cuál es la media del sorteo y cuál es su desvío cuadrático. Un sorteo de los de tipo actual es el de 46.000 números, con 1.630 premios, para el que resulta la media = 35,434 y $\overline{0}$ = 5,783.

El grupo "media dispersión" es para este caso "de 32,543 premios a 38,325", que tiene una probabilidad total de 0,43, que nos da, en consecuencia, una frecuencia teórica de 0,43 × 46 = 19,78. O sea que habrá 20 millares con premios entre 32 y 38 (ambos inclusive) o entre 33 y 39 (ambos inclusive) en un sorteo único.

El grupo "una dispersión" es desde 29,651 a 41,217, cuya frecuencia única es 31.

El grupo "dos dispersiones" es desde 23 a 48 y la frecuencia única de los que quedan fuera de ese grupo, o sea "22 premios o menos y 49 premios o más", es de dos para un único sorteo.

Disimetría.

Como es natural, dado que p no es igual a q, las probabilidades para los desvíos simétricos no son iguales. La probabilidad de obtener 39 premios es mayor que la de obtener 41, y así en todos los desvíos negativos, hasta llegar al desvío 1,75 unidades dispersión. A partir de aquí, la probabilidad del desvío positivo es mayor que la del negativo simétrico, llegando en las proximidades de 2,5 unidades dispersión a ser la probabilidad de uno doble de la del otro.

Puede, por tanto, hacerse la siguiente predicción:

El número total de millares con menos premios de la media será mayor que el de los de más de 40 premios; sobre todo en los grupos cercanos a la media.

El millar que haya obtenido más premios tendrá un exceso de éstos sobre la media, mayor que la diferencia con la media del que haya obtenido menos premios.

Las desigualdades de los premios por millar en los sorteos, notadas por todo el mundo y que han producido a veces alarmas en personas que no tenían obligación de saber la existencia de la disimetría, prueban precisamente la perfección del sistema de nuestra Lotería Nacional.

DIAGNOSIS

I.—El peso de los toros de lidia.

El complejo aleatorio.

Es el peso de los toros de lidia después de muertos en el ruedo. La estadística se compiló en 1911. Se refiere a 1.000 toros lidiados aquel año en las principales plazas de España. Aunque el número total fué mayor, se prescindió de las plazas de menor categoría, que siempre admiten ganado más pequeño, para no tergiversar la estadística. También se se prescindió de tres corridas observadas para que el número fuese precisamente de 1.000, lo cual tiene muchas ventajas de estudio y de tipificación. Las corridas fueron catorce de ocho toros y ciento cuarenta y seis de seis. En alguna de ellas hubo toros enviados al corral por defectuosos, pero que fueron sustituídos por los sobreros respectivos. En la estadística están contados dos ejemplares enviados al corral después de los avisos correspondientes al matador, y que fueron muertos y pesados "ex circensi".

El peso de los toros se expresa en unidades arroba, apreciándose por fracciones de $^1/_4$, que es, por tanto, el que pudiéramos llamar elemento diferencial de la estadística si queremos que el complejo sea función continua. Se han agrupado de tres en tres elementos y se ha dado al grupo el valor del mediano. Así, cuando se expresa que hubo 96 toros que pesaron 23 arrobas, es el resultado de los grupos 22 $^3/_4$, 23, 23 $^1/_4$ arrobas—peso que fueron precisamente 27, 33 y 36 toros. Los grupos finales comprenden todos los que pesaron menos o más de los valores correspondientes. Para facilidades del cálculo, y puesto que los valores-peso están ya escalados, sustituímos por la denominación de los grupos en su orden natural, como se indica aquí:

0	1	2	3	4	5	6	7	8	9	10	11	(grupos)
$<21^1/_4$	$21^1/_2$	$22^1/_4$	23	$23^3/_4$	$24^1/_2$	$25^1/_4$	26	$26^3/_4$	$27^1/_2$	$28^1/_4$	$>28^1/_2$	(arrobas)
14	31	76	96	108	120	141	155	132	92	28	7	Toros

CONSIDERACIONES PREVIAS.

El Actuario, antes de tratar un complejo, debe formarse su opinión de cómo debe ser éste, resultado de su experiencia, y muchas veces podrá deducirse que la estadística no está bien compilada o no es adecuada. Así, por ejemplo, si se le presenta la "retirada diaria de fondos del Banco X", ya puede suponer que habrá días anormales (pago de quincenas o de sueldos), lo que producirá una protuberancia en la frecuencia y que la estadística estará influída por el movimiento estacional de todos los negocios. En el caso actual, podemos afirmar a priori que la estadística *tenderá* a la normal de Gauss, por ser un hecho biológico libre; pero que no se *aproximará* a ella, por ser resultado de observaciones heterogéneas que proceden: *a)* de las distintas castas españolas; *b)* de la época distinta de la lidia dentro del año. La *a)* nos indica que el complejo va a ser hiponormal. Las series de "gran moda" sólo se dan en selecciones. La extracción de bolas de urnas de composición diferente sabemos que produce achatamiento en las frecuencias. Matemáticamente diríamos que el "exceso" va a ser hiponormal.

La *b)* podría corregirse con el tratamiento del movimiento histórico de estadística, nombre ampuloso en este caso, que consistiría en hallar la expresión del aumento de peso del toro a través del año, para corregir, aumentando o disminuyendo, el que dió en la romana el día de su lidia. Es como si dijésemos peso del toro al 1.º de julio, por ejemplo, si ese día era el que había correspondido al centroide de la curva. Esta corrección sería inútil sabiendo que algunos ganaderos daban pienso especial para ciertas corridas; sería pretender pesar leña en balanza de precisión. Nos habremos de encontrar, pues, con un complejo de disimetría negativa, o sea, con mayor desvío por debajo de la media, como sucede en todas las series de la naturaleza y en las económicas de rendimiento decreciente. "El enano se da con más facilidad que el gigante", si bien en este caso el enano no entra en la estadística, porque lo retira el público al corral, aunque con las desigualdades propias del error con que juzga la masa de espectadores.

El polígono de frecuencia en cuanto se le dibuja y el cálculo de la disimetría y del exceso nos confirman lo que habíamos sospechado, resultando, finalmente, que el toro de lidia español presenta circunstàncias muy tipificadas, como si se tratase de una sola casta. En la época actual, que ha conocido tantos cruces de ganaderías, la tipificación sería aún mayor; hecho cuya demostración puede hacer el que lo desee por la comparación de la estadística de 1911 con la de 1944.

Tratamiento matemático.

Se ha seguido el de Gram. Este demostró que la función-probabilidad se podía expresar así: $\Sigma\, c_i\, \varphi\ (\varphi)$ para $i = 0, 1, ...,$ en la cual φ_i es la derivada de orden i de la función de Laplace de la probabilidad normal, o sea la de $p = q = 0{,}50$ y C unos coeficientes que se han llamado, por su orden: media, dispersión, disimetría, exceso; sin que hayan tenido aún homologación oficial los quinto y sexto, que yo los llamaré, ya que los empleo, agudeza y mochez, sin intención de que se recoja mi idea, y cuyos nombres, lo mismo que los tercero y cuarto, indican algo de su sentido, ya que los términos de orden impar corrigen las desviaciones disimétricas y los de orden par las desviaciones de techo en las frecuencias.

Los valores de la probabilidad normal y los de sus derivadas, hasta la de sexto orden, correspondientes a los desvíos de los diferentes grupos del complejo, se exponen en los cuadros que siguen y están copiados de las "tablas para biométricos".

Para hallar los coeficientes de cada derivada he seguido los tres métodos más importantes: el de momentos, el de semi-invariante de Tihele y el de resolver el sistema de ecuaciones por el método de Gauss por mínimos cuadrados, comprobando la ventaja enorme de tiempo de cálculo que lleva el segundo sobre los otros.

Los resultados, o sea la frecuencia teórica, se expresan separadamente para los casos de emplear la fórmula de Laplace; tomar tres términos del desarrollo de la serie o tomar cinco términos. Así se puede apreciar la eficiencia de la teoría de Gram, verdadero padre de la "Estadística matemática". Un gráfico de resultados hace más patente lo que acabo de decir.

Cuadro núm. 1.

(1)	(d)	(f)	f × d	f d²	f d³	f d⁴	f d⁵	f d⁶
0	— 6	14	— 84	+ 504	— 3.024	+ 18.144	— 108.864	+ 653.184
1	— 5	31	155	775	3.875	19.375	96.875	484.375
2	— 4	76	304	1.216	4.864	19.456	77.824	311.296
3	— 3	96	288	864	2.592	7.776	23 328	69.984
4	— 2	108	216	432	864	1.728	3.456	6.912
5	— 1	120	120	120	120	120	120	120
6	— 0	141	0	0	0	0	0	0
7	+ 1	155	+ 155	+ 155	+ 155	+ 155	+ 155	+ 155
8	+ 2	132	264	528	1.056	2.112	4.224	8.448
9	+ 3	92	276	828	2.484	7.352	22.056	66.168
10	+ 4	28	112	448	1.792	7.168	28.672	114.688
11	+ 5	7	35	175	875	4.375	21.875	109.375
		1.000	— 1.167	+ 6.045	— 15.339	+ 87.761	— 310.467	1.824.705
			+ 842		+ 6.362		+ 76.982	
			— 325		— 8.977		— 233.485	
		S_0	S_1	S_2	S_3	S_4	S_5	S_6

Cuadro núm. 1.

Se forma suponiendo que la media del complejo corresponde al grupo 6 que tomamos como origen y cuya hipótesis corregiremos luego para ajustarla exactamente. La columna (d) es la de los desvíos de cada grupo contados desde el origen que acabamos de tomar. La columna (f) es la de la frecuencia de la experiencia, y las demás columnas nos sirven para hallar los parámetros del complejo: son sumas-potencias o momentos cuya fórmula general es $S_r = \Sigma \varphi^r \cdot F(\varphi)$ para $r = 0, 1, \ldots 6$. No olvidemos que nuestra intención es encontrar la sustitución de $F(\varphi$ por $\int_{n-1/2a}^{n+1/2a} \varphi(\varphi)\, d\varphi$ siendo a el intervalo de observación, en este caso $\frac{1}{4}$ de arroba.—Este valor de $\varphi(\varphi)$ demostró Gram que se puede desarrollar así $\varphi(\varphi) = \Sigma c_i \varphi_i(\varphi)$ para $i = 0; 1; \ldots n \ldots$ La serie es rápidamente convergente.—Las φ_i son las derivadas de orden i de la función $\varphi_0(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$ en la que $z = (\varphi - M) : \overline{0}$ o sea los desvíos de la variable sobre la media (M), expresados en unidades dispersión o «desvío medio cuadrático ($\overline{0}$)».—Los C_i son coeficientes que podemos deducir de la experiencia.—La fórmula general es $C_r = \frac{(-1)^r}{\underline{|r}\, \sigma^r} \lambda_r$ para $r=0; \ldots i \ldots$ Los parámetros λ_i se deducen de las observaciones y se llaman semi-invariantes de Thiele que fué quien los introdujo.—He aquí su cálculo:

Cálculo de las semi-invariantes de Thiele.

Recordemos que

$$
\begin{aligned}
s_1 &= \lambda_1 s_0 \\
s_2 &= \lambda_1 s_1 + \lambda_2 s_0 \\
s_3 &= \lambda_1 s_2 + 2\lambda_2 s_1 + \lambda_3 s_0 \\
s_4 &= \lambda_1 s_3 + 3\lambda_2 s_2 + 3\lambda_3 s_1 + \lambda_4 s_0 \\
s_5 &= \lambda_1 s_4 + 4\lambda_2 s_3 + 6\lambda_3 s_2 + 4\lambda_4 s_1 + \lambda_5 s_0 \\
s_6 &= \lambda_1 s_5 + 5\lambda_2 s_4 + 10\lambda_3 s_3 + 10\lambda_4 s_2 + 5\lambda_5 s_1 + \lambda_6 s_0
\end{aligned}
$$

De donde deducimos

a) $\lambda_1 = s_1 : s_0 = -0{,}325.$ $\lambda_1^2 = 0{,}105625$ $\lambda_1^3 = 0{,}034328$ $\lambda_1^4 = 0{,}011025$
$\lambda_2 = (s_2 s_0 - s_1^2) : s_0^2$

b) $\begin{cases} \sigma^2 = 6{,}045 - \lambda_1^2 = 5{,}939375; & \sigma^4 = 35{,}2836 \quad \sigma^6 = 208{,}7289 \\ \sigma = 2{,}4370 & \sigma^3 = 14{,}5342 \quad \sigma^5 = 86{,}3028 \end{cases}$

c) $\lambda_3 = (s_3 s_0^2 - 3 s_2 s_1 s_0 + 2 s_1^3) : s_0^3$

$$\begin{array}{rrl}
 & -8{,}977 & \ldots\ldots\ m_3 \\
 & +5{,}894 & \ldots\ldots\ 3 m_2 m_1 \\
 & -0{,}069 & \ldots\ldots\ 2 m_1^3 \\
\hline
\lambda_3 = & -3{,}152 &
\end{array}$$

d) $\lambda_4 = (s_4 s_0^3 - 4 s_3 s_1 s_0^2 - 3 s_2^2 s_0^2 + 12 s_2 s_1^2 s_0 - 6 s_1^4) : s_0^4$

$$\begin{array}{rrl}
 & +\ 87{,}761 & \ldots\ldots\ m_4 \\
 & -\ 7{,}858 & \ldots\ldots\ 4 m_3 m_1 \\
 & -\ 109{,}444 & \ldots\ldots\ 3 m^2 \\
 & +\ 7{,}653 & \ldots\ldots\ 12 m_2 m_1^2 \\
 & -\ 0{,}066 & \ldots\ldots\ 6 m_1^4 \\
\hline
\lambda_4 = & -\ 21{,}954 &
\end{array}$$

e) $\lambda_5 = \Big($expresado en función de m_i siendo $m_i = s_r : s_0\Big) = m_5 - 5 m_4 m_1 -$
$- 10 m_3 m_2 + 20 m_3 m_1^2 + 30 m_2^2 m_1 - 60 m_2 m_1^3 + 24 m_1^5$

$$\begin{array}{rrl}
 & -233{,}485 & \ldots\ldots\ m_6 \\
 & +142{,}611 & \ldots\ldots\ 5 m_4 m_1 \\
 & +542{,}211 & \ldots\ldots\ \vdots \\
 & -\ 18{,}852 & \ldots\ldots\ \vdots \\
 & -356{,}041 & \ldots\ldots\ \vdots \\
 & +\ 12{,}358 & \ldots\ldots\ 60 m_2 m_1^3 \\
 & -\ 0{,}073 & \ldots\ldots\ 24 m_1^5 \\
\hline
\lambda_5 = & +\ 88{,}729 &
\end{array}$$

f) $\lambda_6 = m_6 - 6 m_5 m_1 - 15 m_4 m_2 + 30 m_4 m_1^2 + 120 m_3 m_2 m_1 - 120 m_3 m_1^3 - 10 m_3^2 +$
$+ 360 m_2 m_1^4 - 270 m_2^2 m_1^2 + 30 m_2^3 - 124 m_1^6$

$$\begin{array}{rrl}
+ & 1824 & \ldots\ldots\ m_6 \\
- & 466 & \ldots\ldots\ 6 m_5 m_1 \\
- & 7893 & \ldots\ldots\ » \\
+ & 277 & \ldots\ldots\ » \\
+ & 2113 & \ldots\ldots\ » \\
- & 36 & \ldots\ldots\ » \\
- & 792 & \ldots\ldots\ » \\
+ & 24 & \ldots\ldots\ » \\
- & 1020 & \ldots\ldots\ » \\
+ & 6564 & \ldots\ldots\ +30 m_2^3 \\
 & 0 & \ldots\ldots\ 124 m_1^6 \\
\hline
+ & 595 &
\end{array}$$

Cálculo de los coeficientes $C_i = \frac{(-1)^r}{\underline{|r}\ \sigma^r} \lambda_r$

$$C_0 = 1 \quad C_1 = 0 \quad C_2 = 0 \quad C_3 = \frac{-3,152}{14,534 \times 6} = +0,03605$$

$$C_4 = \frac{21,954}{35,2836 \times 24} = -0,0261 \quad C_5 = -\frac{+88,729}{86,3028 \times 120} = -0,0091$$

$$C_6 = \frac{595}{208,7289 \times 720} = 0,0041$$

Cuadro número 2.

La columna (1) son los grupos de observación. La (2) son los desvíos sobre la media exacta, que en el cuadro número 1 habíamos supuesto provisionalmente que era el grupo 6, y que ya conocemos al calcular la semi-invariante λ_1. La columna (3) es el valor de $z = \frac{\varphi - M}{\overline{0}}$ que es el desvío expresado en unidades dispersión, se obtiene dividiendo los valores de la columna (2) por $\overline{0} = 2,4370$. Las columnas (4), (5) y (6) se obtienen de las tablas "biométricas" y se exponen con su signo. Como comprobación tenemos la suma de la (4), que por ser la probabilidad total, debe ser igual a la unidad, o sea $\overline{0} = 2,4370$, que le hemos tomado por unidad. Las (7) y (8) son las (5) y (6) (funciones derivadas), multiplicadas por sus respectivos coeficientes, ya encontrados. La (9) es la suma de las (4), (7), (8), o sea de los tres primeros términos del desarrollo en serie de $\varphi(\varphi)$ y expresa la probabilidad para cada grupo. Nos da una comprobación con su suma que debe ser igual a la unidad adoptada, puesto que es la probabilidad total. La columna (10) "con corrección (CC)" es la frecuencia teórica de cada caso: Se obtiene prorrateando 1.000, que es el total observado entre las distintas probabilidades. La columna (12) es la frecuencia teórica si solamente hubiésemos cogido un término de la serie, y por eso se le titula "SC" sin corrección. Se obtiene prorrateando 1.000 entre los valores de probabilidad expresados en la columna (4), primer término de la serie. La columna (11) es la frecuencia práctica, o sea la obtenida en la experiencia. En el gráfico se puede ver mejor aún la corrección y la experiencia.

Cuadro núm. 2.

(1) φ	(2) $\varphi - \lambda$	(3) z	(4) $\varphi_0(z)$	(5) $\varphi_3(z)$	(6) $\varphi_4(z)$	(7) $C_3\ \varphi_3$	(8) $C_4\ \varphi_4$	(9) Σ	(10) C C	(11) Práctica	(12) S C
0	— 5,675	— 2,331	+ 0,265	+ 1,493	— 0,031	+ 52	+ 1	318	13	14	10
1	— 4,675	— 1,920	0,632	+ 0,828	— 3,392	+ 29	+ 85	746	34	31	25
2	— 3,675	— 1,509	1,295	— 1,457	— 7,042	— 51	+ 176	1,420	57	76	53
3	— 2,675	— 1,098	2,179	— 4,290	— 6,091	— 153	+ 152	2,178	89	96	89
4	— 1,675	— 0,687	3,151	— 5,466	+ 1,247	— 197	— 31	2,923	119	108	130
5	— 0,675	— 0,277	3,840	— 3,090	+ 9,793	— 111	— 245	3,484	141	120	157
6	+ 0,325	+ 0,133	0,3953	+ 0,1313	+ 1,1434	+ 46	— 286	3,707	151	141	162
7	+ 1,325	0,544	3,433	+ 5,064	+ 0,4976	+ 181	— 124	3,490	141	155	141
8	+ 2,325	0,955	2,529	+ 5,040	— 4,985	+ 180	+ 125	2,834	115	132	104
9	+ 3,325	1,366	1,573	+ 2,436	— 7,387	+ 87	+ 184	1,854	75	92	64
10	+ 4,325	1,777	0,827	— 0,214	— 4,932	— 7	+ 123	943	38	28	34
11	+ 5,325	2,188	0,680	— 1,424	— 1,020	— 50	+ 25	655	27	7	28
			2,4357					2,4533	1,000	1,000	997

Cuadro núm. 3.

(1) X	(2) Σ	(3) φ_5	(4) $C_5\,\varphi_5$	(5) φ_6	(6) $C_6\,\varphi_6$	(7) Σ	(8)
0	318	— 6.044	+ 54	— 13.918	— 55	317	13
1	746	— 9.977	+ 89	— 1.668	— 6	651	27
2	1.420	— 4.998	+ 44	+ 27.384	+ 109	1.573	65
3	2.178	+ 10.539	— 94	+ 41.908	+ 167	2.151	88
4	2.923	+ 22.800	— 205	+ 8.855	+ 33	2.753	113
5	3.465	+ 1.520	— 136	— 44.631	— 178	3.151	129
6	3.707	— 7.796	+ 69	— 56.077	— 224	3.552	146
7	3.490	— 22.696	+ 203	— 8.677	— 34	3.659	150
8	2.834	— 16.200	+ 145	+ 36.179	+ 144	3.123	128
9	1.854	+ 200	— 0	+ 36.422	+ 145	1.999	82
10	943	+ 9.600	— 86	+ 7.532	+ 30	887	37
11	655	+ 7.806	— 70	— 12.238	— 49	536	22
	2.4533					2.4352	1.000

Cuadro número 3.

Tiene por objeto mostrar los resultados si se toman para el desarrollo en serie cinco términos. Las columnas (1) y (2) están tomadas del cuadro número 2, y las demás han quedado explicadas en sus similares del cuadro número 2. La (7) nos da la probabilidad y la comprobación de los cálculos con su suma, y la (8) es la frecuencia teórica, cuyo resultado se lleva también al gráfico, que resulta sobradamente expresivo. En él se comprueba la teoría de Gram al ver cómo se acerca a la curva de frecuencia arbitraria (de la experiencia) la curva de la probabilidad normal de Laplace con el solo empleo de parámetros deducidos de la propia experiencia.

Otro procedimiento.

En el mismo método de desarrollo en serie de la función de la probabilidad mediante las derivadas de distinto orden de la función de la probabilidad normal, podemos emplear otro procedimiento para hallar los coeficientes de cada uno de los términos de la serie del desarrollo, en vez del que hemos usado de las "semi-invariantes". Este procedimiento es el clásico de "los mínimos cuadrados". Comenzamos hallando los dos primeros parámetros: media y dispersión que nos son indispensables para hallar los valores de z y encontrar los de la función normal y sus derivadas de distinto orden para cada valor de z. Nótese que al tomar la media M como origen para los distintos valores de X y al tomar como unidad el desvío medio cuadrático $\bar{0}$ o "dispersión", lo que hacemos es una transformación $Z = f(x) = aX + b$, siendo $a = \frac{1}{\bar{0}}$; $b = \frac{-M}{\bar{0}}$, de donde $z = \frac{X - M}{0}$. Con el fin de hacerlo todo nuevamente formamos el *cuadro número 4,* para hallar los dos primeros parámetros sin necesidad de suponer, como hicimos en el cuadro número 1, que la media correspondía al grupo 6, y corregir después. Este procedimiento de "momentos crudos" nos da para los dos primeros parámetros los mismos valores que en el cuadro número 1. Para

Cuadro núm. 4.

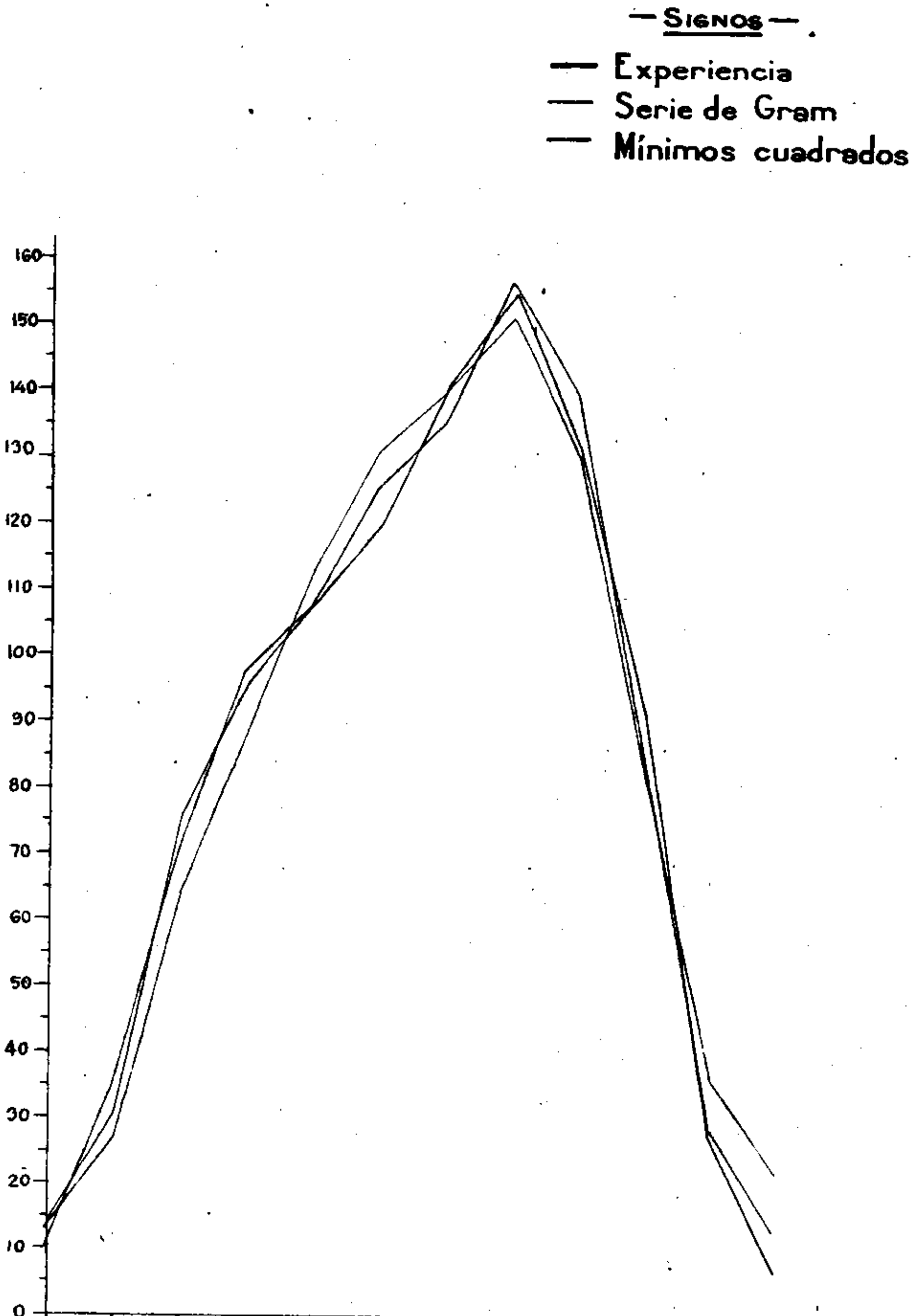
(	(2)	(1) × (2)	(1)² × (2)	(1)³ × (2)
1 —	14	14	14	14
2 —	31	62	124	248
3 —	76	228	684	2052
4 —	96	384	1536	6144
5 —	108	540	2700	13500
6 —	120	720	4320	25920
7 —	141	987	6909	48363
8 —	155	1240	9920	79360
9 —	132	1188	10692	96228
10 —	92	920	9200	92000
11 —	28	308	3388	37268
12 —	7	84	1008	12096
	1000	6675	50495	413193
	S_0	S_1	S_2	S_3

PARÁMETROS

$$M = \frac{S_1}{S_0} - 1 = 5.675$$

$$\bar{0}^2 = \frac{S_2}{S_0} - \frac{S_1^2}{S_0^2} = 50.495 - 44.555625 = 5.939375$$

$$D = \frac{S_3}{S_0} - 3\,\frac{S_2 \times S_1 \times S_0}{S_0 \times S_0 \times S_0} + 2\frac{S_1^3}{S_0^3} = 413.193 - 1011.162 + 594.817 = 3.152$$


— Signos —
Experiencia
Serie de Gram
Mínimos cuadrados
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
10
11

fines de comprobación, se ha hallado también el valor de la disimetría que resulte igual al encontrado en el cuadro número 1. Formados los valores de z para cada observación del complejo, el asunto queda enfocado en hallar los valores de los coeficientes (5) que satisfagan las ecuaciones (12) que nos proporciona la experiencia. El método de Gauss, que es el que seguimos, es, en realidad, breve y cómodo. Asignamos a los coeficientes unos valores arbitrarios que, para mayor comodidad, son múltiplos de 10, y vamos a buscar las correcciones que en ellos hay que hacer al encontrar los valores más plausibles de los mismos.

Sean:

$$C_0 = 1.000 \quad C_3 = 20 \quad C_4 = -20 \quad C_5 = -10 \quad C_6 = +10$$

Multiplicado cada uno de ellos por las derivadas, del mismo orden que los subíndices, de la función normal, que nos dan las tablas, formamos las "ecuaciones de observación". En ellas, para ventaja del cálculo, se han multiplicado todos los términos por 10. La columna A corresponde a la función normal. La B, a la tercera derivada. La C, a la cuarta derivada, etc., etc., y la O a los resultados de la experiencia, o sea a la frecuencia observada. Las diversas columnas S S_1 S_2 ... sirven de contraste de las operaciones y se forman así: S es la suma algebraica de A B C D E O. S_1 es la suma de B C D E O, y así sucesivamente hasta S_{IV}, suma de E O.

Se forman después las sumas-productos AA, AB ...; BB, BC, ...; etcétera, etc., y se pasa a formar la ecuación normal. En las líneas tituladas (a), (b), (c), (d) y (e) del cuadro "ecuación normal" están copiados los resultados de las diversas sumas-productos que hemos formado y contrastado por las columnas S^I. La línea titulada (5) se obtiene dividiendo las cifras de la línea (a), por 68997313, primer término de esta línea. La línea titulada (1) la forman los productos de las cifras de la línea (a) por el factor (— 0,00039), primer término de la línea (5). La línea (2) es el producto de las cifras de la línea (a) por el factor (— 0,01452), segundo término de la (5), y así sucesivamente formamos las líneas (3) y (4).

Ahora se resta la línea (1) de la línea (b), la línea (2) de la línea (c) y así sucesivamente, llevando los resultados a la "primera ecuación reducida". Esta se le somete a las mismas operaciones que la ecuación normal y dan la segunda ecuación reducida. Así llegamos a la cuarta ecuación reducida y, con ellas, al cálculo de los factores auxiliares V_0 V_3 V_4 V_5 V_6 que, multiplicados por los coeficientes arbitrarios elegidos, nos darán los que buscamos o plausibles:

$$C_0 = 1000\ V_0 = 403{,}73 \quad C_3 = 20\ V_3 = 9{,}2842 \quad C_4 = (-20)\ V_4 = -2{,}0288$$
$$C_5 = (-10)\ V_5 = -8{,}3780 \quad C_6 = 10\ V_6 = 5{,}1393$$

Multiplicados los coeficientes por las diversas derivadas, formamos el cuadro número 5, cuya última columna, que es la suma de las anteriores, nos da la frecuencia teórica, que está expresada en números enteros en la primera columna.

Salta a la vista la gran diferencia de trabajo aritmético que hay entre uno y otro procedimiento.

ECUACIONES DE OBSERVACION

	A	B	C	D	E	O	S		SI	SII	SIII	SIV
0	265	+ 29	+ 0	+ 60	— 139	— 140	+ 75	+ 354 — 279	— 190	— 219	— 219	— 279
1	632	+ 16	+ 67	+ 100	— 16	— 310	+ 489	+ 815 — 326	— 143	— 159	— 226	— 326
2	1295	— 29	+ 140	+ 50	+ 274	— 760	+ 970	+ 1759 — 789	— 325	— 296	— 436	— 486
3	2179	— 85	+ 121	— 105	+ 419	— 960	+ 1569	+ 2719 — 1150	— 610	— 525	— 646	— 541
4	3151	— 109	— 24	— 228	+ 88	— 1080	+ 1798	+ 3239 — 1441	— 1353	— 1244	— 1220	— 992
5	3840	— 61	— 195	— 15	— 446	— 1200	+ 1923	+ 3840 — 1917	— 1917	— 1856	— 1661	— 1646
6	3953	+ 26	— 228	+ 78	— 560	— 1410	+ 1859	+ 4057 — 2198	— 2094	— 2120	— 1892	— 1970
7	3433	+ 101	— 99	+ 226	— 86	— 1550	+ 2025	+ 3760 — 1735	— 1408	— 1509	— 1410	— 1636
8	2529	+ 100	+ 99	+ 162	+ 361	— 1320	+ 1931	+ 3251 — 1320	— 598	— 698	— 797	— 959
9	1573	+ 48	+ 147	— 2	+ 364	— 920	+ 1210	+ 2132 — 922	— 363	— 411	— 558	— 556
10	827	— 4	+ 98	— 96	+ 75	— 280	+ 620	1000 — 380	— 207	— 203	— 301	— 205
11	680	— 28	+ 20	— 78	— 122	— 70	+ 402	700 — 298	— 278	— 250	— 270	— 192
	+ 24357	+ 320	+ 692	+ 676	+ 1581	— 10000	+ 14871	24871				
		— 316	— 546	— 524	— 1369							
		+ 4	+ 146	+ 152	+ 212							

S = suma algebraica de A B C D E O
S^{I} .... B C D E O
S^{II} .... C D E O
S^{III} .... D E O
S^{IV} .... E O

A A	A B	A C	A D	A E	A O	A S
70.225	+ 7.685	+ 0	+ 15.900	— 36.855	— 37.100	19.875
399.424	+ 10.112	+ 42.344	+ 63.200	— 10.112	— 195.920	309.048
1.677.025	— 37.555	+ 181.300	+ 64.750	+ 354.830	— 984.200	1.256.150
4.748.041	— 185.215	+ 263.659	— 228.795	+ 913.001	— 2.091.840	3.418.851
9.928.801	— 343.459	— 75.624	— 718.428	+ 277.288	— 3.403.080	5.665.498
14.745.600	— 234.240	— 748.800	— 57.600	— 1.712.640	— 4.608.000	7.384.320
15.626.209	+ 102 778	— 901.284	+ 308.334	— 2.213.680	— 5.573.730	7.348.627
11.785.489	+ 346.733	— 339.867	+ 775.858	— 295.238	— 5.321.150	6.951.825
6.395.841	+ 252.900	+ 250.371	+ 409.698	+ 912.969	— 3.338.280	4.883.499
2.474.329	+ 75.504	+ 231.231	— 3.146	+ 572.572	— 1.447.160	1.905.330
683.929	— 3.308	+ 81.046	— 79 392	+ 62.025	— 231.560	512.740
462.400	— 19.040	+ 13.600	— 53.040	— 82.960	— 47.600	273.360
68 997.313	+ 795.712	+ 1.063.551	+ 1.637.740	+ 3.092.685	— 27.279.620	39.927.123
	— 822.817	— 2.065.575	— 1.140.401	— 4.351.465		
	— 27.105	— 1.002.024	+ 497.339	— 1.258.780		

Comprobación aritmética: + 69.494.652
— 29.567.529

39.927.123 = AS

BB	BC	BD	BE	BO	BS'
841	+ 0	+ 1,740	− 4,031	− 4,060	− 5,510
256	+ 1,072	+ 1,600	− 256	− 4,960	− 2,288
841	− 4,060	− 1,450	− 7,946	+ 22,040	+ 9,425
7,225	− 10,285	+ 8,925	− 35,615	+ 81,600	+ 51,850
11,881	+ 2,616	+ 24,852	− 9,592	+ 117,720	+ 147,477
3,721	+ 11,895	+ 915	+ 27,206	+ 73,200	+ 116,937
676	− 5,928	+ 2,028	− 14,560	− 36,660	− 54,444
10,201	− 9,999	+ 22,826	− 8,686	− 156,550	− 142,208
10,000	+ 9,900	+ 16,200	+ 36,100	− 132,000	− 59,800
2,304	+ 7,056	− 96	+ 17,472	− 44,160	− 17,424
16	− 392	+ 384	− 300	+ 1,120	+ 828
784	− 560	+ 2,184	+ 3,416	+ 1,960	+ 7,784
+ 48,746	+ 32,539	− 1,546	− 80,986	− 378,390	+ 334,301
	− 31,224	+ 81,654	+ 84,194	+ 297,640	− 281,674
	+ 1,315	+ 80,108	+ 3,208	− 80,750	+ 52,627

Comprobación aritmética: + 133,377
− 80,750
+ 52,627 BS'

CC	CD	CE	CO	CS''
0	0	0	0	0
4,489	6,700	− 1,072	− 20,770	− 10,653
19,600	7,000	38,360	− 106,400	− 41,440
14,641	− 12,705	50,699	− 116,160	− 63,525
576	5,472	− 2,112	25,920	29,856
38,025	2,925	86,970	234,000	361,920
51,984	− 17,784	127,680	321,480	483,360
9,801	− 22,374	8,514	153,450	149,391
9,801	16,038	35,739	− 130,680	− 69,102
21.609	− 294	53,508	− 135,240	− 60,417
9,604	− 9,408	7,350	− 27,440	− 19,894
400	− 1,560	− 2,440	− 1,400	− 5,000
180,430	38,135	408,820	734,850	1.024,527
	− 64,125	− 5,624	− 538,090	− 270,031
	− 25,990	403,196	196,760	754,496

Comprobación aritmética: 780,486
− 25,990
754,496 = CS''

DD	DE	DO	DS'''
3,600	— 8,340	— 8,400	— 13,140
10,000	— 1,600	— 31,000	— 22,600
2,500	+ 13,700	— 38,000	— 21,800
11,025	— 43,995	+ 100,800	+ 67,830
51,984	— 20,064	+ 246,240	+ 278,160
225	+ 6,690	+ 18,000	+ 24,915
6,084	— 43,680	— 109,980	— 147,576
51,076	— 19,436	— 350,300	— 318,660
26,244	+ 58,482	— 213,840	— 129,114
4	— 728	+ 1,840	+ 1,116
9,216	— 7,200	+ 26,880	+ 28,896
6,084	+ 9,516	+ 5,460	+ 21,060
+ 178,042	— 145,043	— 751,520	— 652,890
	+ 88,388	+ 399,220	+ 421,977
	— 56,655	— 352,300	— 230,913

Comprobación aritmética: 408,955
+ 178,042
— 230,913 = DS^{III}

EE	EO	ES^IV	OO
19,321	19,460	38,781	19,600
256	4,960	5,216	96,100
75,076	— 208,240	— 133,164	577,600
175,561	— 402,240	— 226,679	921,600
7,744	— 95,040	— 87,296	1.166,400
198,916	535,200	734,116	1.440,000
313,600	789.600	1.103,200	1.988,100
7,396	133,300	140,696	2.402,500
130,321	— 476,520	— 346,199	1.742,400
132,496	— 334,880	— 202,384	846,400
5,625	— 21,000	— 15,375	78,400
14,884	8,540	23,424	4,900
1.081,196	— 1.537,920	2.045,433	11.284,000
	1.491,060	— 1.011,097	
	— 46,860	1.034,336	

Comprobación aritmética: 1.081,196
— 46,860
1.034,336 = ES^{IV}

ECUACION NORMAL

(a)	68.997,313	— 27,105	— 1.002,024	+ 497,339	— 1.258,780	— 27.279,620
(1)		+ 11	+ 391	— 194	+ 491	+ 10,639
(b)		+ 48,746	+ 1,315	+ 80,108	+ 3,208	— 80,750
(2)			+ 14,549	— 7,216	+ 18,295	+ 396,100
(c)			+ 180,430	— 25,990	+ 403,196	+ 196,760
(3)				+ 3,580	— 9,063	— 196,413
(d)				+ 178,042	— 56,655	— 352,300
(4)					+ 22,964	+ 497,580
(e)					+ 1.081,196	— 46,680
(5)		— 0,00039	— 0,01452	+ 0,00720	— 0,01824	— 0,39537

1.ª ECUACION REDUCIDA

(a)	+ 48,735	+ 924	+ 80,302	+ 2,717	— 91,389
(1)		+ 17	+ 1,522	+ 51	— 1,732
(b)		+ 165,881	— 18,774	+ 384,901	— 199,340
(2)			+ 132,315	+ 4,476	— 150,583
(c)			+ 174,462	— 47,592	— 135,887
(3)				+ 151	— 5,094
(d)				+ 58,232	— 544,440
(4)		+ 0,01896	+ 1,64772	+ 0,05574	— 1,87522

2.ª ECUACION REDUCIDA

(a)	+ 165,864	— 20,296	+ 384,850	— 197,608
(1)		+ 2,483	— 47,090	+ 24,179
(b)		+ 42,147	— 52,068	— 5,304
(2)			+ 892,955	— 458,503
(c)			+ 1.058,081	— 539,346
(3)		— 0,12236	+ 2,32027	— 1.19138

3.ª ECUACION REDUCIDA

(a)	+	39,664	—	4,978	—	29,485
(1)			+	624	+	3,700
(b)			+	165,126	—	80,843
(2)			—	0,12550	—	0,74331

4.ª ECUACION REDUCIDA

+	164,502	—	84,543

CALCULO DE LOS FACTORES AUXILIARES

$$V_6 = \frac{84543}{164502} = 0{,}51393\,;$$

$$V_5 = +\,0{,}74331 - (0{,}51393 \times (-\ 0{,}1255) = 0{,}83780$$

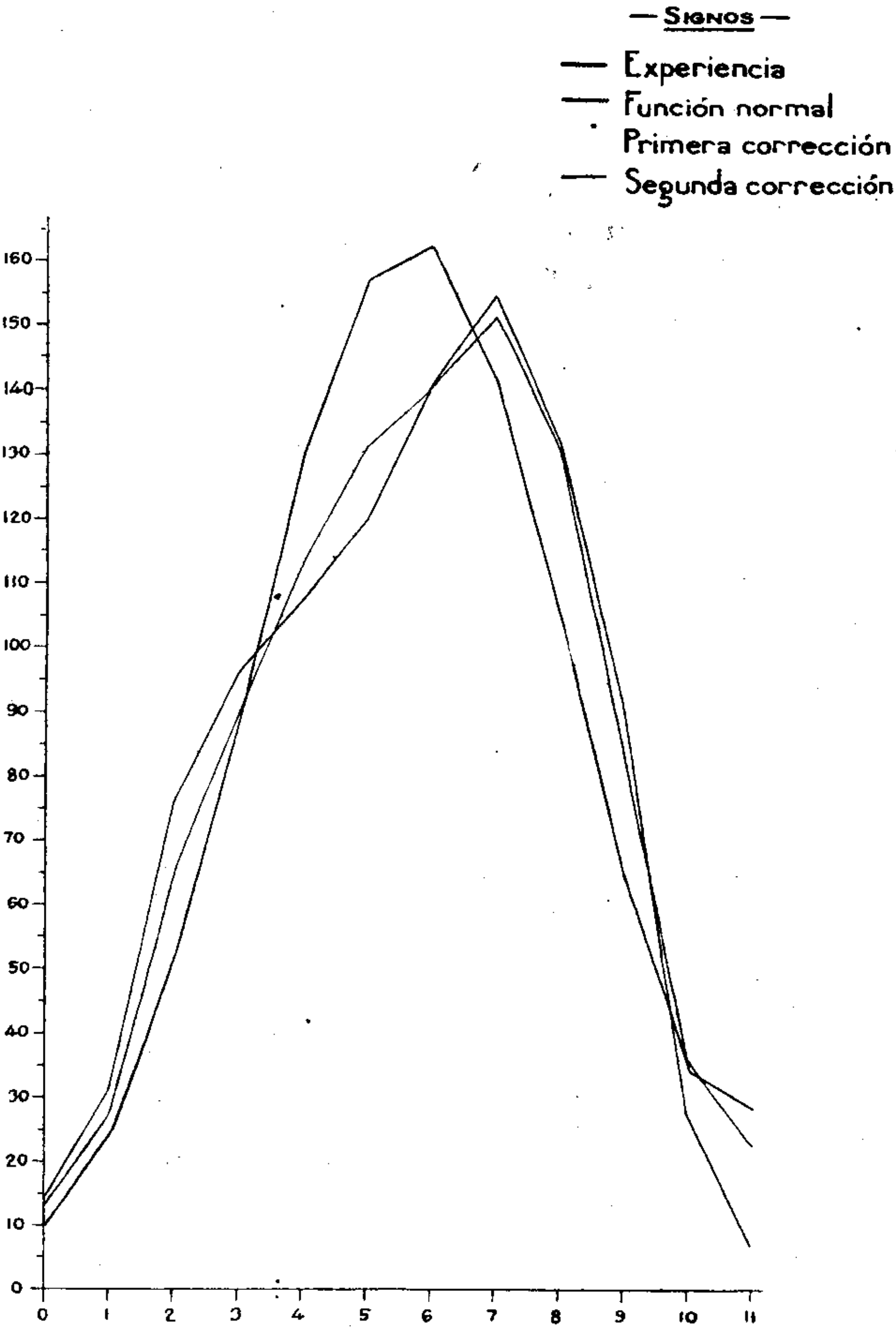
$$V_4 = +\,1{,}19138 - (0{,}8378 \times (-\,0{,}12236) - (0{,}51393)\ (2{,}32027) = 0{,}101442$$

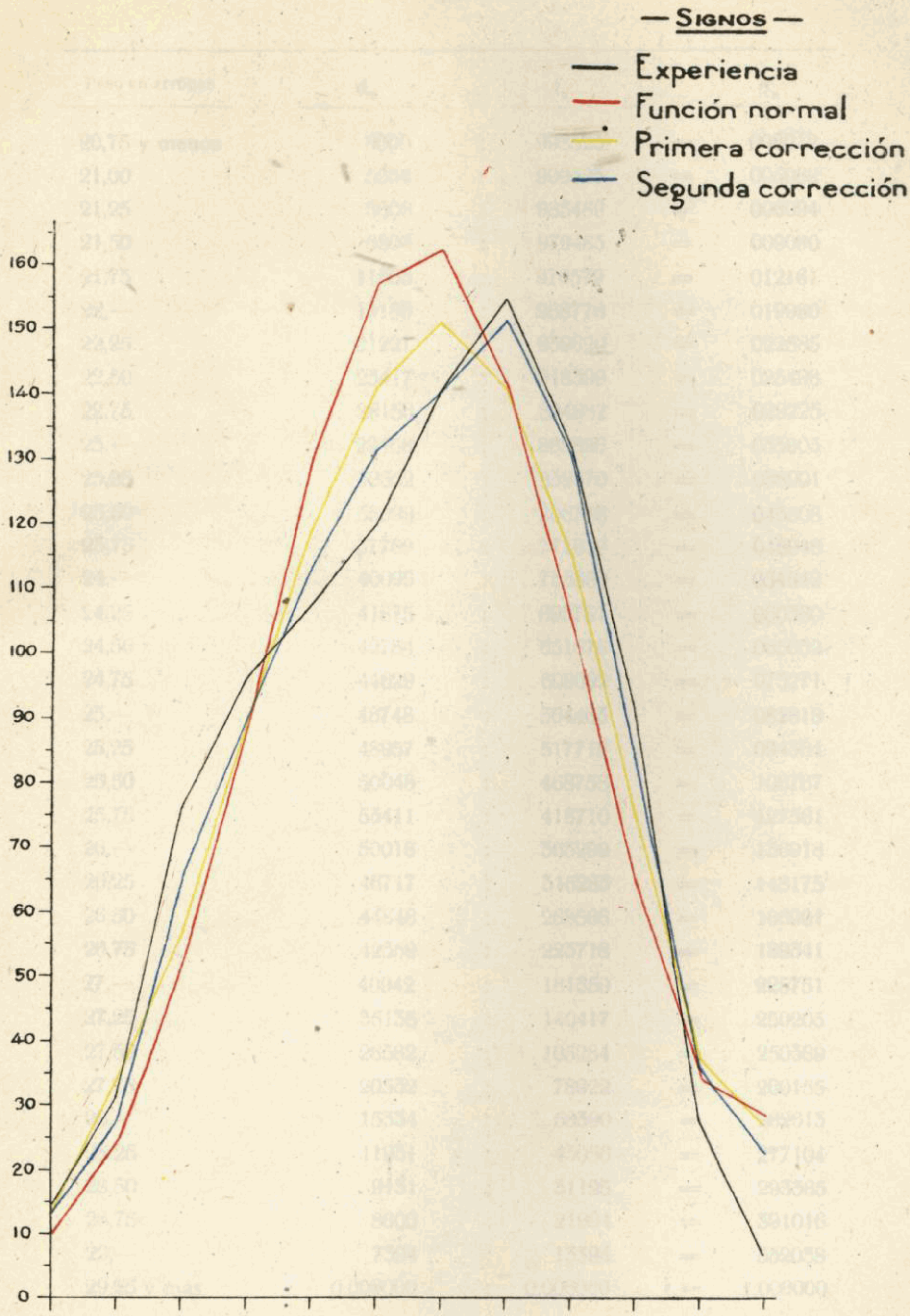
$$V_3 = \quad 1{,}87522 - (10144)\ (0{,}01896) - (0{,}83780)\ (1{,}64772) - (0{,}51393)\ (0{,}05574) = 0{,}46421$$

$$V_0 = +\,0{,}39539 - (0{,}46421)(-\,0{,}00039) - (0{,}10144)(0{,}01452) - (0{,}83780)(0{,}00720) - (0{,}51393)(-\,0{,}01824) = +\,0{,}40373$$

Cuadro núm. 5.

(F)	$\varphi_0 \times C_0$	$C_3 \times \varphi_3$	$C_6 \times \varphi_6$	$C_4 \times \varphi_4$	$C_5 \times \varphi_5$	Σ
10	10,681	1,386	— 7,153	6	5,064	9,984
34	25,474	769	— 857	688	8,360	34,434
70	52,197	— 1,353	14,075	1,429	4,188	70,536
97	87,828	— 3,985	21,540	1,236	— 8,831	97,788
107	127,007	— 5,077	4,551	— 0,253	— 19,106	107,122
125	154,778	— 2,870	— 22,940	— 1,987	— 1,273	125,708
135	159,333	1,219	— 28,823	— 2,321	6,533	135,941
156	138,373	4,704	— 4,459	— 1,010	19,019	156,627
139	101,936	4,682	18,596	1,011	13,575	139,800
85	63,402	2,263	18,720	1,499	— 167	85,717
29	33,333	— 198	3,871	1,001	— 8,044	29,963
13	27,408	— 1,322	— 6,290	0,207	— 6,541	13,462
1000						1.007,082


— Signos —
Experiencia
Función normal
Primera corrección
Segunda corrección
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
10
11



— Signos —
Experiencia
Función normal
Primera corrección
Segunda corrección
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
0 1 2 3 4 5 6 7 8 9 10 11

Peso en arrobas	d_x		l_x		q_x
20,75 y menos	8000	:	998523	=	008012
21,00	5034	:	990523	=	005082
21,25	6006	:	985489	=	006094
21,50	8904	:	979483	=	009090
21,75	11803	:	970579	=	012161
22,—	19156	:	958776	=	019980
22,25	21221	:	939620	=	022585
22,50	23417	:	918399	=	025498
22,75	26156	:	894982	=	029225
23,—	29456	:	868826	=	033903
23,25	32652	:	839370	=	038901
23,50	35099	:	806718	=	043508
23,75	37769	:	771619	=	048948
24,—	40099	:	733850	=	054642
24,25	41875	:	693751	=	060360
24,50	42784	:	651876	=	065632
24,75	44629	:	609092	=	073271
25,—	46748	:	564463	=	082818
25,25	48957	:	517715	=	094564
25,50	50048	:	468758	=	106767
25,75	53411	:	418710	=	127561
26,—	50016	:	365299	=	136918
26,25	46717	:	315283	=	148175
26,50	44848	:	268566	=	166991
26,75	42359	:	223718	=	189341
27,—	40942	:	181359	=	225751
27,25	35133	:	140417	=	250205
27,50	26362	:	105284	=	250389
27,75	20532	:	78922	=	260155
28,—	15334	:	58390	=	262613
28,25	11931	:	43056	=	277104
28,50	9131	:	31125	=	293365
28,75	8600	:	21994	=	391016
29,—	7394	:	13394	=	552038
29,25 y más	0,006000	:	0,006000	=	1,000000

Los valores de las frecuencias teóricas obtenidos por ambos procedimientos comprenden en la mayoría de los casos los valores de la frecuencia obtenida en la experiencia.

Resultado final.

En el cuadro número 6 figura el resultado final, no ya por grupos de tres valores de la variable, como habíamos empezado el tratamiento. La columna (1), encabezada d_x , es la probabilidad correspondiente al caso considerado. Como el elemento diferencial de medición es el $^1/_4$ de arroba, cuando se dice la probabilidad de obtener en peso 23 arrobas, es la probabilidad de que el peso del toro esté comprendido entre 22 $^7/_8$ de arroba y 23 $^1/_8$ de arroba. La columna l_x es la suma de las d_x anteriores. La columna q_x es el cociente de las dos anteriores.

Se ha encabezado este cuadro con los símbolos, tan familiares a los Actuarios, del Seguro de Vida, por ser éste el básico y primeramente estudiado de todos los Seguros; con ello queda demostrado que se podría obtener la tabla de mortalidad por el tratamiento de un complejo aleatorio con muchísima más facilidad y sin pérdida de exactitud que por las habituales fórmulas de Makeham.

DIAGNOSIS

II.—La bravura de las vacas bravas.

EL COMPLEJO ALEATORIO.

Es el imponderable que pudiéramos llamar bravura de las vacas de casta. La estadística se compiló hace aproximadamente cincuenta años, en las tientas anuales de una ganadería brava castellana para seleccionar las madres. La selección consistía en probar la vaca mediante la suerte de varas, aceptando solamente las que tomaban el sexto puyazo. Las varas habían de tomarse arrancándose de lejos, sin más excitación que la del jinete y recargando y sin volver la cara en la suerte. El resultado es el que se expresa en el cuadro para el número total de 512 vacas, que fueron las tentadas en varios años.

La estadística, como puede apreciarse, es extraordinariamente disimétrica y aguda. Cualquier sistema de computación dará resultados muy alejados de la realidad. Por esta razón se ha empleado la transformación logarítmica de la variable.

TRATAMIENTO MATEMÁTICO.

Hay muchos complejos aleatorios que están ligados por una ley logarítmica que ya expresó Fechner existía para causas y efectos. Esta ley va teniendo cada día en todas las ciencias mayor universalidad, sobre todo cuando se dispone de la elección de la base de logaritmos, aunque la serie de logaritmos naturales es la que tiene mayor aplicación. Ello sugiere un cambio de la variable x por la variable log x, conservando la misma frecuencia para cada desviación.

El desarrollo en serie de una curva de frecuencia es

$$F(\varphi) = C_0\,\varphi_0(\varphi) - C_1 : \underline{|1}\ \varphi_1(\varphi) + C_2 : \underline{|2}\ \varphi_2(\varphi) - C_3 : \underline{|3}\ \varphi_3(\varphi) + \dots$$

En la cual la función generadora es la de Laplace, para $p = q = 0{,}50$ $\varphi_0(\varphi) = 1 : \sigma\sqrt{2\pi} \times e^{-(\varphi - m)^2 : 2\sigma^2}$ y las φ_i sus derivadas de orden$_i$

En la cual *m* y σ son los parámetros llamados media y dispersión. En la nueva serie motivada por el cambio de variable, la función generadora es la siguiente:

$$\psi_0(\varphi) = 1 : n\sqrt{2\pi} \times e^{-\frac{1}{2} x \left(\frac{\log x - m}{n}\right)^2}$$

En la cual *m* y *n* son dos parámetros ligados por las fórmulas que en cuadro se ponen. Los valores de *x* se cuentan a partir del 0 matemático, que se encuentra por las fórmulas de Jorgensen, pero que en el caso de nuestro trabajo corresponde al de siete varas, que no se le obligaba a tomar a ninguna vaca.

Cuadro núm. 1.

(1) Varas	(2) Vacas	(3) φ	(4) (3) $\times$ (2)	(5) $(3)^2 \times (2)$	(6) log φ	(7) (6) — m	(8) (7) : n	(9) $\varphi_0(z)$	(10) Frecuencia teórica (9) $\times K_0$: n
7	—	0	—	—	—	—	—	—	—
6	348	1	348	348	0,0000	+ 0,1450	+ 0,21	0,3900	269
5	63	2	126	252	0,6932	0,8382	1,24	0,1800	124
4	42	3	126	378	1,0986	1,2436	1,82	0,0750	52
3	28	4	112	448	1,3863	1,5313	2,25	0,0317	22
2	15	5	75	375	1,6094	1,7544	2,58	0,0146	10
1	10	6	60	360	1,7918	1,9368	2,84	0,0071	5
0	6	7	42	294	1,9456	2,0906	3,07	0,0035	3
	512		889	2455					485
	S_0		S_1	S_2					

Cálculo de parámetros $K_0 = S_0^3 \; S_2 : S_1^3$ $\quad e^{2m} = S_1^8 : S_0^5 \; S_2^3$ $\quad e^{n^2} = S_0 \; S_2 : S_1^2$

log S_2	3,39005		
3 log S_0	8,12781		
3 log S_1	8,84670		
log K_0	2,67116	$K_0 = 468{,}99$	$K_0 : n = 689$

Log S_0 S_2	6,09932	
2 log S_1	5,89782	
log e^{n^2}	0,20150	$n^2 = 0{,}4640$ $n = 0{,}68$

8 log S_1	23,59120	
5 log S_0	13,54635	
3 log S_2	10,17015	
log e^{2m}	1,87470	$2m = -0{,}289$
		$m = -0{,}145$

Cuadro número 1.

La primera columna expresa los distintos valores que puede tomar la variable, habiendo escogido como 0 matemático el valor 7, que no puede dar nunca valor alguno de frecuencia. Los desvíos sobre ese 0 están expresados en la columna (3). La columna (2) es la de las frecuencias observadas y da, a su vez, el momento de orden 0. Las (4) y (5) dan los momentos de orden 1 y 2. La (6) expresa los logaritmos naturales de los desvíos. Las (7) y (8), las operaciones indicadas, y la (9) expresa la función normal de la probabilidad para el desvío $z = \frac{\log \varphi - m}{n}$, o sea para el desvío expresado en la columna (8).

Debajo del cuadro se expresan las fórmulas de Jorgensen que ligan los parámetros con los momentos y se pone su cálculo, que ha servido para formar, luego de conocidos, las columnas (7) y (8).

La columna (10) da la frecuencia teórica si tomamos solamente un término de la serie: es decir, $F\ (x) = K_0\, \varphi_0\ (z)$ Comparada con la frecuencia práctica, vemos que, aun a pesar de la transformación logarítmica, el resultado deja mucho que desear, si bien es más aproximado que cualquier sistema de cálculo sin esa transformación, ya que la frecuencia teórica muestra una curva del género, en agudeza, del fenómeno observado en la práctica. En busca de mayor aproximación vamos a calcular los segundo y tercer términos de la serie y formamos el

Cuadro núm. 2.

(1) z	(2 φ_0 (z)	(3) φ_3 (z)	(4) φ_4 (z)	(5) $K_0\, \varphi_0$	(6) $K_3\, \varphi_3$	(7) $K_4\, \varphi_4$	(8) Frecuencia teórica
0,21	0,3900	+ 0,240	+ 1,07	287,57	— 10,01	+ 67,55	345,12
1,24	0,1800	+ 0,333	— 0,71	132,72	— 13,89	— 44,82	74,00
1,82	0,0750	— 0,045	— 0,45	55,30	+ 1,88	— 28,41	28,77
2,25	0,0317	— 0,147	— 0,05	23,37	+ 6,13	— 3,15	26,35
2,58	0,0146	— 0,135	+ 0,10	10,76	+ 8,50	+ 6,31	25,57
2,84	0,0071	— 0,102	+ 0,13	5,23	+ 4,17	+ 8,27	17,67
3,07	0,0033	— 0,070	+ 0,12	2,58	+ 2,87	+ 7,00	12,45
							529,94

A	B	C	O	S	S'	S''
1950	+ 480	+ 1070	— 3480	+ 20	— 1930	— 2410
900	+ 666	— 710	— 630	+ 226	— 674	— 1340
375	— 90	— 450	— 420	— 585	— 960	— 870
158	— 294	— 50	— 280	— 466	— 624	— 330
73	— 270	+ 100	— 150	— 247	— 320	— 50
35	— 204	+ 130	— 100	— 139	— 174	+ 30
18	— 140	+ 120	— 60	— 62	— 80	+ 60
+ 3509	+ 1146	+ 1420	— 5120	— 1499		
	— 998	— 1210	+ 3967	+ 246		
	+ 248	+ 210		— 1253		

AA	AB	AC	AO	AS
+ 3.802,500	+ 936,000	+ 2.086,500	— 6.786,000	+ 39,000
810,000	+ 599,400	— 639,000	— 567,000	+ 203,400
140,625	— 33,750	— 168,750	— 157,500	— 219,375
24,964	— 46,452	— 7,900	— 44,240	— 73,628
5,329	— 19,710	+ 7,300	— 10,950	— 18,031
1,225	— 7,140	+ 4,550	— 3,500	— 4,865
324	— 2,520	+ 2,160	— 1,080	— 1,116
4 784,967	+ 1.535,400	+ 2.100,510	— 7.570,270	+ 242,400
	— 109,572	— 815,650		— 317,015
	+ 1.425,828	+ 1.284,860		— 74,615

BB	BC	BO	BS'	CC	CO	CS''
+ 230,400	+ 513,600	— 1.670,400	— 926,400	+ 1.144,900	— 3.723,600	— 2.578,700
+ 443,556	— 472,860	— 419'580	— 418,884	504,100	+ 447,300	+ 951,400
+ 8,100	+ 40,500	+ 37,800	+ 86,400	202,500	+ 189,000	+ 391,500
+ 86,436	+ 14,700	+ 82,320	+ 183,456	2,500	+ 14,000	+ 16,500
+ 72,900	— 27,000	+ 40,500	+ 86,400	10,000	— 15,000	— 5,000
+ 41,616	— 26,520	+ 20,400	+ 35,496	16,900	— 13,000	+ 3,900
+ 19,600	— 16,800	+ 8,400	+ 11,200	14,400	— 7,200	+ 7,200
902,608	+ 568,800	— 2.089,980	— 1.375,284	1.895,300	— 3.758,800	— 2.583,700
	— 543,180	+ 189,420	+ 402,952		+ 650,300	+ 1.370,500
	+ 25,620	— 1.900.560	— 972,332		— 3.108,500	— 1.213,200

Cuadro número 2.

Las columnas (1) y (2) proceden del cuadro anterior: columnas (8) y (9). Las columnas (3) y (4) son las tercera y cuarta derivadas de la función normal y están copiadas de las tablas para Biometría. Para calcular los coeficientes de las (2), (3) y (4) empleamos el método de mínimos cuadrados, que parece más rápido en este caso que el cálculo directo de ellos. Tenemos siete ecuaciones con tres incógnitas. Asignamos provisionalmente a los coeficientes los siguientes valores: $K_0 = 500$, $K_3 = 200$, $K_4 = 100$, a fin de dar sencillez a los cálculos, y con el mismo fin multiplicamos todos por 10, así como la frecuencia de la práctica.

Así se forma el cuadro general *A - B - C - O* de ecuaciones de observación, que tiene los productos de los coeficientes provisionales por las funciones normal o derivadas y el producto de 10 por la frecuencia. Formamos las columnas S - S′ - S″, que han de servir de contraste en las operaciones sucesivas mediante el mecanismo siguiente:

$$S = A + B + C + O. \quad S' = B + C + O. \quad S'' = C + O.$$

Formamos luego las sumas productos AA - AB - AO, BB - BO, etc., comprobando las operaciones mediante las columnas AS, BS′, CS″, y formamos las ecuaciones normal y reducidas, que nos darán los valores definitivos de los coeficientes. El método seguido es el de Gauss. Las ecuaciones normales (a), (b) y (c) expresadas en el cuadro "Ecuación normal" son los resultados obtenidos de las sumas-productos AA, AB, AC, AO, para la (a); BB, BC, BO, para la (b), y así sucesivamente. Las cifras de la línea titulada (3) en este cuadro están formadas dividiendo cada una de las cifras de la línea (a) por su primer término, o sea por 4.784.967. La línea titulada (1) contiene los productos de los términos de la línea titulada (a) por el factor 0,29798, primer término de la línea (3). La línea (2) es la de los productos por el factor 0,26852 de la línea (a).

El cuadro de las "primeras ecuaciones reducidas" (a) y (b) se forma restando en el cuadro de las ecuaciones normales la línea (1) de la línea (b), la línea (2) de la línea (c). Estas ecuaciones reducidas se tratan de la misma manera que las normales y conducen a la "segunda ecuación reducida".

Así llegamos a la obtención de los factores auxiliares V_0 V_3 V_4, que figuran en el cuadro de cálculo, y que multiplicados por los coeficientes provisionales que habíamos elegido, nos darán los definitivos. Así, pues,

$$
\begin{aligned}
K_0 &= 500 \times 1{,}47475 = 737{,}375 \\
K_3 &= 200 \times (-\,0{,}20868) = -\,41{,}736 \\
K_4 &= 100 \times 0{,}63133 = 0{,}63133
\end{aligned}
$$

Obtenidos los coeficientes, formamos las columnas (5), (6) y (7) del cuadro número 2. La columna (8) es la suma de las tres columnas anteriores y nos da la frecuencia teórica. Salta a la vista la mayor aproximación obtenida con los tres términos de la serie sobre la obtenida con uno solo. Si tomásemos 5 términos de la serie, la aproximación sería mayor aún, sin que haga falta comprobarlo, pues es sabido que la frecuencia total disminuye y quedaría próxima a la de la experiencia, al mismo tiempo que bajaría la frecuencia del desvio próximo a 1,25 unidades dispersión y la de los desvíos próximos a 3 unidades dispersión.

Con ello se comprueba la excelencia del método de transformación logarítmica y su notable sencillez en casos como el que hemos estudiado, de profunda agudeza.

ECUACION NORMAL

+ 4.784,967	+ 1.425,828	+ 1.284,860	− 7.570,270	(a)
	+ *424,868*	+ *382,862*	− *2.255,789*	*(1)*
	+ 902,608	+ 25,620	− 1.900,560	(b)
		+ *345,010*	− *2.032,768*	*(2)*
		+ 1.895,300	− 3.108,500	(c)
	+ *0,29798*	+ *0,26852*	− *1,58209*	*(3)*

1.ª ECUACIÓN REDUCIDA

+ 477,740	− 357,242	+ 355,229	(a)
	+ *267,134*	− *265,629*	*(1)*
	+ 1.550,290	− 1.075,732	(b)
	− *0,74777*	+ *0,68076*	*(2)*

2.ª ECUACIÓN REDUCIDA

+ 1.283,156 — 810,103

$$V_4 = \frac{810{,}103}{1.283{,}156} = 0{,}63133$$

$$V_3 = -\,0{,}68076 - (0{,}63133 \times -\,0{,}74777) = -\,0{,}20868$$

$$V_0 = +\,1.58209 - (-0{,}20868 \times +\,0{,}29798) - (0{,}63133 \times 0{,}26852) = 1{,}47475$$

APÉNDICE

En los dos ejemplos de diagnosis que he propuesto se ha pretendido encontrar una ley generadora del complejo o, lo que es lo mismo, una expresión analítica que nos dé la probabilidad de presentación de los distintos valores que puede alcanzar la aleatoria. Esta ley, aplicable en el "momento actual", podrá no serlo en el futuro, o no haber sido la misma en el pasado, ya que las aleatorias tienen vitalidad y se desarrollan y transforman aun en aquellas que parecen más inconmovibles. Es bien sabido, por ejemplo, que la estatura de una raza no es inalterable a través de los siglos. También hay certeza de que las tablas de mortalidad humana evolucionan profundamente. Los complejos podemos compararlos al sistema solar, que en cada momento físico tiene posiciones diferentes entre sus planetas, a la vez que el sol (la media que pudiéramos decir) sigue una marcha fija al parecer, pero quizá aleatoria también. Y aun hay complejos en que, además de la evolución del valor medio, varía también la fluctuación de los valores alrededor de la media. He ahí una labor que los Actuarios legamos a las generaciones venideras. ¿Quién puede medir la importancia que tendrá en el futuro una estadística que hoy compilamos y a la que por falta de antecedentes no podemos sacar toda la enseñanza que en sí encierra? Y claro está que no me refiero a los ejemplos que he propuesto, que, como he dicho al principio, no pretenden nada utilitario, sino que han salido a luz por españolísimos y para que en español practiquen los que tengan a ello vocación.

Pero prescindiendo de la labor potencial de la diagnosis y de la actual de investigación de causas (que para que sea fructífera ha de hermanarse con una aptitud filosófica), se pueden obtener para la vida práctica resultados muy eficaces al expresar analíticamente los complejos aleatorios. De la misma manera que la tabla de mortalidad sirve para resolver todos los casos del Seguro de Vida, también la expresión analítica de la probabilidad de presentación de diversos valores de una aleatoria resuelve casos de Seguro con ella relacionados. Esta actividad es precisamente la actividad actuarial, que no se limita al Seguro de Vida; es la labor de los llamados "matemáticos asesores", que no faltan en las Empresas o actividades modernizadas, y de la que doy en visión panorámica rapidísima, algunos ejemplos en una sistematización, quizá arbitraria, que me atrevo a presentar, y con la que doy fin al trabajo.

a) *Riesgo cierto.*

1. La Red de Compañías de electricidad de una nación conoce el complejo aleatorio de "potencia disponible procedente de saltos de agua cada año", según el régimen de lluvias. Se conoce la energía contratada por la red con sus clientes y se quiere concertar el Seguro de la oferta, es decir, que se ha de poder servir toda la fuerza contratada.

Claro es que el Seguro puede ser contratando una indemnización o previendo la red, que en este caso sería auto-aseguradora, la cantidad de energía que ha de suministrar con hulla negra para satisfacer la demanda.

2. Una Compañía de navegación aérea vuela 30 millones de kilómetros al año y acepta como riesgo cierto (tomando una dispersión del riesgo de accidente igual a dos veces el desvío medio cuadrático), la rotura de un total de 12 aviones. Se conoce por estadística el complejo aleatorio "rotura del avión en accidente", que tiene como valores de la variable desde el 1 por 100 del valor del avión hasta el 100 por 100 o pérdida total. La Compañía contrata el Seguro de reparación del material sobre las siguientes bases: Las roturas de más del 95 por 100 no admiten reparación; las de menos del 5 por 100 son reparables "in situ", y se concreta, por tanto, solamente las comprendidas entre el 5 y el 95. El problema aquí, como en todos los casos, puede ser: exterior a la Compañía si es solamente de indemnización de la Compañía aseguradora, en cuyo caso se obtiene la prima en función de la probabilidad y el valor del avión; o puede ser interior a la Compañía si ésta es auto-aseguradora, en cuyo caso, lo que más importa conocer es el número de horas de trabajo o de operarios-día que necesita la Compañía para la reparación de este material, lo cual determina la plantilla óptima de personal especialista.

b) *Riesgo incierto.*

Es el del Seguro de incendio, de accidente de vuelo, etc., determinados también en los complejos aleatorios correspondientes.

Ejemplos: La rotura por accidente en la red actual de las líneas aéreas mundiales es un avión por cada 2 millones de kilómetros, con una dispersión de 400.000 kilómetros. Se contrata un vuelo de 2.000 kilómetros para un avión que vale 2 millones de pesetas.

c) *Riesgo definido.*

1. Una fábrica de óptica de precisión tiene la estadística de su fabricación en serie, que contiene siete grupos, que pueden ser "defectos de centraje", "defectos de colorido", "perfectos", "defectos de calibre", etc., etc., y la probabilidad de cada uno; de los cuales solamente uno es el que puede ofrecer a la clientela, desechando los otros seis por defectos de construcción irremediables; cuyas primeras materias se recuperan, perdiéndose la mano de obra.

En este caso, el Seguro es el de fabricación y el auto-seguro determina cuál es el sobreprecio con que debe cargarse cada ejemplar terminado.

2. En una Empresa de productos químicos procedentes de la naturaleza cultivada, se conoce el complejo "acidez del arbusto", que está escalonado en quince grupos, que van del valor 20 al 35. Para la fabricación sólo se emplea los ejemplares del 24 al 31 de acidez, desechándose totalmente los demás.

El auto-seguro consiste en este caso en saber a cómo ha de pagarse cada ejemplar de arbusto, ya que su recolección es por millares y se hace en el campo en su totalidad y sin análisis previo.

Un caso muy conocido de este Seguro es la determinación de los reclutas que pueden ser útiles para el servicio X en cada reemplazo, si se les exige una estatura y un perímetro torácico definidos.

d) *Riesgos elegidos.*

1. El Sindicato Hotelero de un archipiélago de turismo conoce el completo aleatorio "viajeros-día que le visitan en cada mes del año" y quiere decidir el número de habitaciones (oferta) que opone a la demanda, a fin de que el rendimiento económico sea el óptimo. El auto-seguro de la Empresa es (en este caso) Seguro de máximos beneficios. Ha de pensarse que si todos los hoteles del Sindicato estuvieran llenos siempre, su rendimiento individual sería máximo (100 por 100), pero habría una

posible corriente turística que acabaría no visitando el archipiélago por falta de alojamiento, y de la que no se beneficiaría el Sindicato. La superproducción temporal es más beneficiosa en el conjunto del año, porque todos los fenómenos económicos de oferta y demanda tienen *rendimiento* en el sentido técnico de la palabra de cociente de demanda por oferta, que no puede ser constantemente igual a la unidad por la naturaleza aleatoria del complejo. Cuando la demanda eficiente no puede ser satisfecha en períodos largos, se disminuye y el rendimiento deja de ser total, con riesgo inclusive para la Empresa que pudo ser próspera con rendimientos menores que la unidad.

2. El avión o aviones que sirven una línea internacional tiene diariamente una capacidad de carga de V viajeros y M mercancías. El Estado se ha reservado para el correo y para sus fines la carga M, a cambio de la subvención que da a la Empresa.

Se conoce el complejo "utilización de la línea por viajeros", en que se expresa la probabilidad de llevar $\dot{X}$ viajeros en cada viaje. El Estado quiere reservarse la carga $M' > M$, de tal modo que tenga el 80 por 100 de probabilidad de que los viajeros nacionales no tengan que tomar el billete en otra Compañía extranjera, al encontrarse que no tenían billete en la nacional por disminución de los sitios disponibles para viajeros, al haber aumentado la carga reservada a la mercancía.