



Lej.
36

**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

FACULTAD DE CIENCIAS

**“FRANCIS GALTON Y KARL PEARSON: ANALISIS
HISTORICO DE ALGUNAS CONTRIBUCIONES
A LA ESTADISTICA”**

T E S I S

QUE PARA OBTENER EL TITULO DE:

A C T U A R I O

P R E S E N T A :

JORGE VICTOR ROSASLANDA SANDOVAL

México, D. F.

1986



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

I N D I C E

Introducción	2
Francis Galton (1822-1911)	10
Karl Pearson (1857-1936)	40
Conclusiones	67

I N T R O D U C C I O N

Con el presente trabajo pretendemos analizar el proceso histórico de la conformación del razonamiento estadístico, a fin de conocer las causas que lo originaron y sus implicaciones en el desarrollo del pensamiento científico ulterior.

El examen y discusión de los motivos que llevaron al planteamiento de problemas, cuya novedad consistió en el tratamiento que se les dió, nos permite ubicar mejor el producto de las investigaciones y trabajos que llegarían a constituir un cuerpo de conocimientos muy útil a la ciencia.

El análisis de las interrogantes que se formularon los creadores de las incipientes técnicas estadísticas, así como de las consecuencias de sus conclusiones, nos proporciona una comprensión más clara de lo que hoy puede ser y son sus aplicaciones en la investigación científica. Por lo tanto con la aproximación desde el punto de vista histórico a los conceptos de la Estadística se verá cómo y dónde se inició el desarrollo de esta disciplina y poder así entender mejor su naturaleza.

Se han escogido a dos científicos ingleses debido a que su labor la consideramos como el fundamento de la Estadística clásica, tanto teórica como aplicada. Los temas de investigación que desarrollaron fueron originales y los conceptos que establecieron son ahora esenciales en muchos campos del saber.

Francis Galton y, sucediéndole en tiempo, Karl Pearson aprovecharon los conocimientos de su tiempo al dirigir sus

inteligencias y sus vidas hacia horizontes no explorados. Cada uno con sus propias motivaciones trabajaron sistemáticamente en problemas nada nuevos pero con un enfoque distinto, a saber: La cuantificación y matemización de los fenómenos de la vida.

En la mente de estos dos hombres se llegaron a mezclar conocimientos que se desarrollaban de manera independiente: el azar (tratado por los matemáticos franceses y traducido al cálculo de probabilidades), la aritmética política dirigida principalmente por William Petty y Graunt, y las ciencias políticas de la escuela alemana. Así también teorías revolucionarias como la de Darwin e ideales políticos como el de desear la creación de un Estado sustentado sobre bases diferentes.

Por el interés de descubrir el proceso hereditario, por el afán de demostrar científicamente tesis sobre Evolución y por el rechazo a vaguedades existentes en su época en la medicina y biología, reestructuraron, adecuaron e inventaron herramientas matemáticas cuyas aplicaciones actuales han rebasado enormemente las originales. Aunado a ésto y como consecuencia natural, establecen conceptos tan amplios y tan válidos como el de correlación y regresión.

La inquietud de Pearson y Galton por aplicar sus herramientas matemáticas en campos que tradicionalmente no se utilizaban sirvió para mostrar las posibilidades del desarrollo de técnicas que constituirían el método estadístico. Al abordar temas o áreas que no habían sido tratadas cuantitativamente crearon un razonamiento, produjeron avances y profundizaron en disciplinas científicas como es el caso de la Biometría. Galton y Pearson son considerados los iniciadores de dicha ciencia,

cuya definición actual es: aplicación de los métodos estadísticos a la solución de problemas biológicos. Sin embargo, en un principio estuvo restringida al análisis de la variación o variabilidad biológica.

El ambiente intelectual de la Gran Bretaña a mediados del siglo XIX (el gran período victoriano medio de expansión industrial y comercial) resultó saturado de notables contribuciones en diversos campos del conocimiento. Junto con varias innovaciones tecnológicas aparecieron trabajos de economía política de Adam Smith; David Ricardo; Thomas Malthus, James y John Stuart Mill; en medicina Edward Jenner propuso métodos antisépticos en la cirugía; y en 1859 sale publicado "El Origen de las Especies" de Charles Darwin.

Esta última obra provocó grandes cambios en todos los aspectos del pensamiento prevaleciente y no solamente en los campos aludidos en la obra mencionada (biología, botánica y geología). Su amplio esquema explicatorio constaba de varios aspectos:

- 1) El mundo no es estático sino que evoluciona en forma gradual.

Durante el viaje del Beagle había quedado profundamente impresionado cuando descubrió en las formaciones de las Pampas grandes animales fósiles cubiertos de coraza, como las de los actuales armadillos; en segundo lugar, por la manera en que animales estrechamente emparentados se sustituyen unos a otros conforme se va hacia el sur del Continente; y en tercer lugar, por el carácter sudamericano de la mayor parte de los productos de las islas Galápagos, y más especialmente por la manera en que difieren ligeramente los de cada una de las islas del grupo sin que ninguna de ellas parezca muy vieja en sentido geológico. Era evidente que hechos como éstos, y también otros muchos sólo podían explicarse mediante

la suposición de que las especies se modifican gradualmente. (1)

2) La evolución explicada por medio de la Selección Natural:

Pronto me di cuenta que la selección era la clave del éxito del hombre cuando conseguía razas útiles de animales y plantas. Pero durante algún tiempo continuó siendo un misterio para mí la forma en que podía aplicarse la selección a organismos que viven en estado natural. En octubre de 1938, esto es, quince meses después de haber empezado mi estudio sistemático, se me ocurrió leer por entretenimiento el ensayo de Malthus sobre la población, y, como estaba bien preparado para apreciar la lucha por la existencia que por doquier se deduce de una observación larga y constante de los hábitos de animales y plantas, descubrí enseguida que bajo estas condiciones las variaciones favorables tenderían a preservarse y las desfavorables a ser destruidas. El resultado de ello sería la formación de especies nuevas. (2)

3) Finalmente, la comunidad de la descendencia:

Si optamos por dejar las conjeturas correr a su antojo, los animales hermanos en el dolor, la enfermedad, la muerte, el sufrimiento y el hambre -nuestros esclavos en los momentos más penosos, nuestros compañeros en las diversiones- acaso participen en nuestro origen en un antepasado común: quizá procedamos todos de un mismo crisol. (3)

Galton, después de la lectura de este libro revolucionario, cambió de geógrafo a eugenista y antropólogo, pues así lo

(1) Charles Darwin: Autobiografía y cartas escogidas. Madrid; Alianza Editorial, 1984. (2a. Edición) Vol. 1 p. 85.

(2) Ibidem., 86.

(3) Ibidem., 88.

muestran sus publicaciones. Aquí el paso trascendental consistió en utilizar los métodos estadísticos como el instrumento indispensable para probar las tesis mencionadas. En principio, respecto a la obra de Darwin, Galton empezó por tratar de darle soporte experimental al concepto, por medio del cual, Darwin explicaba el proceso hereditario: la teoría de la Pángénesis.(1) Calificada por su autor como 'hipótesis provisional', ya que reconocía la falta de sustento de pruebas científicas de esta parte de su obra; Galton se dispuso a demostrarla, llegando a negar, por lo menos lo que se refiere al hecho de que la sangre fuera el vehículo de las gémulas.

Por otro lado, en el número 1 de la revista *Biometrika*, Karl Pearson escribe en su editorial:

La primera condición necesaria para que cualquier proceso de selección natural pueda iniciarse en una raza es la existencia de diferencias entre sus miembros y el primer paso para un examen del posible efecto de un proceso selectivo sobre cualquier característica de una raza debe ser una estimación de la frecuencia con que se presentan individuos, exhibiendo un cierto grado de deformidad con respecto a ese carácter. La unidad de investigación no debe ser un individuo, sino una raza y el resultado debe tomar la forma de una relación numérica, mostrando la frecuencia relativa con que se presentan las varias especies de individuos que componen la raza.(2)

(1) Pángénesis: Consistía, según Darwin, en suponer que cada célula del organismo expide partículas invisibles, llamadas gémulas que llevan la copia de la célula emisora, encargándose la sangre de transportarla hasta donde se encuentran las células sexuales, donde se reúnen. Estos supuestos portadores de caracteres hereditarios pueden reproducir los elementos anatómicos de que provienen y además se consideran corpúsculos blandos, modificables gracias a estos, es posible que los caracteres adquiridos se transmitan a la descendencia.

(2) E.S. Pearson: Pearson, Creador de la Estadística Aplicada. Argentina Espasa Calpe. 1948. p.14.

Así, Pearson nos deja testimonio tanto del razonamiento estadístico a seguir en la demostración de las afirmaciones darwinianas, como del estímulo de la obra mencionada en la revista que se publicó por primera vez en el año 1901 con el apoyo financiero y moral de Galton y que actualmente continúa emitiéndose.

También es conveniente dar lugar a la referencia explícita de otra influencia significativa en las primeras aplicaciones del método estadístico y que posteriormente analizaremos con más detalle.

En *Los Elementos de la Ciencia* (1892), obra esencialmente de contenido filosófico escrita por Pearson, encontramos lo siguiente:

Nuestro objetivo en biología es idéntico al de la física, a saber, describir las más amplias agrupaciones de fenómenos en las fórmulas más breves posibles.

Es claro el deseo de aplicar la matemática fuera de sus campos ortodoxos de aplicación pretendiendo imitar el éxito obtenido en la Física debido a las leyes establecidas por Newton. E.G. Pearson supone que su padre, al hacer tal afirmación, pensaba en la Ley de Herencia Ancestral propuesta por Galton (en su obra más connotada [Herencia Natural], como una ley tan general como la Ley de la Gravitación Universal. (1)

A lo largo del presente trabajo se discutirán los logros de dos autores: Karl Pearson y Francis Galton, pero conviene

(1) E.S. Pearson. Pearson, Creador de la Estadística Aplicada. Argentina. Espasa Calpe. 1948. p. 14.

advertir que seguramente se tuvo la intervención de otros individuos quienes de igual manera participaron en la formación del razonamiento estadístico. Sin embargo, al retroceder en la búsqueda de los orígenes de la 'aproximación estadística' hemos restringido nuestro modesto trabajo a ellos dos. El no hacerlo implicaría extenderlo más allá de nuestras posibilidades. Esto implica, de algún modo, soslayar contribuciones muy valiosas que por lo mismo, justificarían una investigación para ellas solas. Por ejemplo la obra de Ronald Fisher.

Procederemos entonces al desarrollo del trabajo, no sin temor a las fallas que son inevitables en una primera investigación de esta naturaleza.

Se tienen pruebas materiales de que en un tiempo entre el hombre prehistórico de hace 400 siglos y el tercer milenio antes de Cristo el "Homo Sapiens" inventó los juegos y entre éstos juegos los llamados de azar*. Para posteriormente descubrir ciertas analogías con los sucesos de la Naturaleza.

* F.N. David, Dicing and Gaming (a note on the history of probability) Biometrika 42, (1955). 2.

FRANCIS GALTON
(1822 - 1911)

Sir Francis Galton nació en Inglaterra a principios del siglo pasado. Después de pasar por varias escuelas, en 1840 ingresa al Trinity College a la edad de 18 años como estudiante de matemáticas; tres años más tarde obtiene su grado por votación.

Ya antes había cursado dos años de medicina sin realmente atraerle. Cuando muere su padre resulta dueño de una gran fortuna, hecho que le permite llevar una vida más cómoda. Realiza viajes y expediciones a lugares recónditos; a causa de uno de ellos enferma y permanece cuatro años en Inglaterra para posteriormente interesarse en el Cabo, Sudáfrica, en donde escribe un diccionario de Inglés y la lengua aborígen. También escribe leyes que propone a los jefes locales; al regreso a su patria recibe una medalla otorgada por la Sociedad Geográfica Real y otra por la Sociedad Geográfica Francesa.

Su primera publicación fué *El Arte de Viajar y le siguió Turistas Vacacionales y Notas de Viaje*. En ellos y otros subsecuentes da muestra de curiosidad científica, específicamente sobre meteorología. En 1858 ya pertenecía al Comité Directivo del Observatorio Kew, donde se probaban y calibraban instrumentos geográficos, haciendo uso de la "Ley de error" de Gauss para determinar los límites permisibles de error, de un instrumento dado. Un problema para los meteorólogos de esa época (1850-1860) era el de la predicción del tiempo a largo plazo. Galton primero escribió artículos sobre construcción y prueba de instrumentos y después, en base a gráficas y mapas con

direcciones de vientos y presiones atmosféricas, en 1870 intentó predecir la velocidad del viento en función de la altura barométrica, temperatura y humedad, pero falló. Sin embargo este fracaso le llevó a familiarizarse con el uso de la curva normal y con funciones lineales de predicción. Hecho que se puede considerar como un primer acercamiento al concepto de Regresión. Cabe recordar que los meteorólogos de su tiempo consideraban al modelo normal solamente como modelo de los errores de medición.

Galton se inicia propiamente con la Estadística en 1865 con la publicación de *Cardcter y Talento Hereditario*, en el cual aborda el tema de la herencia. Galton se plantea las preguntas ¿Puede ser heredada la inteligencia?, ¿Los hijos de padres brillantes también serán brillantes?. Su respuesta fué afirmativa y la fundamentó con una investigación que consistió en contar el número de personajes importantes aparecidos en diccionarios biográficos y mostrar su parentesco entre ellos. Cuando compara las frecuencias relativas con aquellas provenientes de la población en general, concluye que debe ser cierta su afirmación debido a que resultan mayor las primeras. Es decir, con mucha frecuencia, los individuos 'más dotados' son parientes próximos de individuos también 'muy dotados'.

Para él, la comprensión de un problema consiste en matematizarlo. Esta tendencia lo lleva a introducir las matemáticas de la probabilidad. En la siguiente publicación llamada *Genio Hereditario*, muestra la concordancia entre las distribuciones de las mediciones de características físicas, como estatura, medidas de tórax, etc., pertenecientes a soldados, con lo que teóricamente se puede obtener al hacer uso de la ley de la 'desviación del promedio' (así llamada por él al modelo normal).

De manera análoga, añade, se puede observar o se debe observar la distribución de la inteligencia de acuerdo a la misma ley. Y para no caer en una simple analogía, muestra la distribución de los resultados de exámenes aplicados a aspirantes a puestos públicos.

Así mismo afirma no estar de acuerdo con una supuesta igualdad innata de los recién nacidos, quienes después, en edad escolar o como adultos mostrarán gran diferencia de potencia mental, producto esto de su entrenamiento escolar. El pensaba que si alguien es realmente excepcional, es por motivos de linaje y no por influencias educacionales y sociales. El título del siguiente artículo es sugerente: *Hombres de Ciencia Ingleses: su Naturaleza y Crianza* (1874).

En la segunda de sus obras hasta ahora mencionadas, Galton dice:

Yo soy el primero en tratar la materia de manera estadística, llegar a resultados numéricos e introducir la ley teórica de 'desviación del promedio' en una discusión de Herencia.

Respecto a esto cabe considerar la influencia intelectual de Adolphe Quetelet (1796-1874) Astrónomo Real de Bélgica, sobre Galton. Por comentarios del mismo Galton es posible afirmar que es innegable dicha influencia, aunque Karl Pearson su biógrafo más importante no lo acepte del todo. (1)

Pues afirma que, ninguno de los libros de Quetelet estaban en su biblioteca personal; y además, los apuntes que tomé

(1) Ruth Schwartz Cowan. Francis Galton's Statistical Ideas: The Influence of Eugenics. Isis 63 (1972) 512.

sobre Quetelet y su Física Social, llenan solamente una página de su libreta.

Sin embargo, veamos algunos comentarios respecto al uso del modelo normal, primero de Quetelet:

Yo muestro que la ley de 'las causas accidentales' es una ley general, la cual se aplica tanto a individuos como a gente; así como gobierna nuestras cualidades morales e intelectuales, de la misma forma que lo hace con nuestras cualidades físicas. Así, lo que se considera como accidental deja de serlo cuando las observaciones se extienden a un número considerable de casos. (1)

Cuando se refiere a cualidades intelectuales se podría pensar en la inteligencia como una de ellas, pero nos hacen saber D. Landau y P. Lazarsfeld (2) que ocasionalmente en las partes más especulativas de sus escritos, propuso solamente una forma de medir la productividad escolar (contando el número de publicaciones producidas por el hombre), sin realmente demostrarlo experimentalmente, puesto que nadie había coleccionado tales datos. La información que sí manejó era sobre tasas de criminalidad, suicidios y casamientos de varias poblaciones y por generaciones sucesivas.

Enseguida vemos como Galton evoca a Quetelet en líneas previas al uso de la ley en *Genio Hereditario*:

La ley es muy general. El señor Quetelet, Astrónomo Real de Bélgica y autoridad máxima en estadística social y de población, la ha usado ampliamente en sus investigaciones.

(1) Adolphe, Quetelet. Letters and Probabilities. Londres Layton and Co. 1849. p. IX. (Citado por D. Landau y Paul F. Lazarsfeld: Quetelet, Adolphe, Incluido en International Encyclopedia of Statistics, New York. Free Press (A) 1978 p. 828).

(2) D. Landau, Paul F. Lazarsfeld. Ibidem., p. 831.

Ha construido también tablas numéricas, a partir de las cuales pueden hacerse cálculos, siempre que sea menester utilizar la ley.

Quienes quieran saber más de lo que yo puedo relatar aquí, deben consultar su trabajo que está expuesto en un libro muy agradable, en octavo, que debería ser mejor conocido por los estadísticos. Su título es *Letters and Probabilities*, traducido por Downes, al Inglés. Editado por Layton y Co., Londres, 1849. (1)

Por lo expuesto, Galton tiene razón, en considerase pionero en el tratamiento estadístico de la herencia, por lo menos en cuanto a Quetelet, pero su influencia sobre Galton la consideramos admisible.

Regresando a la distribución de la inteligencia, Galton deseaba determinar mediciones o rangos de 'genio' para que su estudio fuera preciso. Con el uso de la ley de 'desviaciones' y alguna medida disponible de genio, podría evaluar sus proposiciones. La curva normal de inteligencia sería dividida en rangos y así poder obtener el número de individuos que cayeran en cada uno.

Así lo intentó en *Genio Hereditario* tomando como base 15 millones de gente, obteniendo una distribución del número de hombres brillantes y tontos; partiendo de una estimación previa que realizó con las calificaciones obtenidas al aplicar una prueba de inteligencia a un gran número de individuos.

Ejemplificamos en la tabla que aparece en la siguiente página.

(1) James R. Newman. El Mundo de las Matemáticas. México, Grijalbo, 1976 (3a. edición) Vol. III, p. 248.

CLASIFICACIÓN DE LOS HOMBRES SEGÚN SUS DOTES NATURALES

Grados de habilidad natural, separados por intervalos iguales		Número de personas comprendidas en los diferentes grados de habilidad natural, con respecto de su capacidad general o de sus aptitudes especiales							
Por debajo del promedio	Por encima del promedio	Proporcional, o sea, uno de cada	En cada millón de la misma edad	En el total de la población del Reino Unido, es decir, 15 millones, de las siguientes edades					
				20-30	30-40	40-50	50-60	60-70	70-80
a	A	4	256.791	651.000	495.000	391.000	268.000	171.000	77.000
b	B	6	162.279	409.000	312.000	246.000	168.000	107.000	48.000
c	C	16	63.563	161.000	123.000	97.000	66.000	42.000	19.000
d	D	64	15.696	39.800	30.300	23.900	16.400	10.400	4.700
e	E	413	2.423	6.100	4.700	3.700	2.520	1.600	729
f	F	4.300	233	590	450	355	243	155	70
g	G	79.000	14	35	27	21	15	9	4
x	X								
Todos los grados por debajo de g	Todos los grados por encima de G	1.000.000	1	3	2	2	2		
A ambos lados del promedio			500.000	1.268.000	964.000	761.000	521.000	332.000	149.000
Total, de ambos lados . . .			1.000.000	2.536.000	1.928.000	1.522.000	1.042.000	664.000	298.000

La proporción de hombres de diferentes edades ha sido calculada a partir de las proporciones que son válidas para Inglaterra y Gales. (Censo 1861, apéndice, p. 107.)

Por ejemplo, la clase F contiene uno de cada 4.300 hombres. O sea, hay 233 de esta clase en cada millón de hombres. Lo mismo sucede dentro de la clase J. En todo el Reino Unido hay 590 hombres de la clase F (y el mismo número de la J) entre las edades de 20 y 30 años; 450 entre los 30 y 40; etc.

La labor más fructífera de Galton, para la Estadística resultó cuando se interesó por la herencia. ¿Qué fué lo que le llevó a involucrarse en semejante empresa?

Sabemos que como matemático no tuvo éxito extraordinario, pero fué un medidor y contador compulsivo. (1) Es posible que siguiera fielmente la exhortación de Galileo: Hay que medir lo medible y hacer medible lo que no es. Y entonces, una actitud meramente científica hizo que considerara inadecuadas las teorías genéticas de Darwin, estimulándolo

(1) Ruth Schwartz Cowan. Ibidem, p. 510.

para tratar de resolver el problema de la herencia. Como veremos más adelante, propone inclusive su propia teoría del mecanismo hereditario.

Bien, pero se afirma(1) que atrás de esa actitud científica existió una motivación dominante de tipo político ideológico: mejorar a la humanidad. O como Pearson dijo: el mejoramiento en la raza del hombre. Para esto, la clave la encontró en Darwin.

Galton pensaba que las 'cualidades mentales' se podían controlar. Los hijos de padres 'bien dotados', también serían 'bien dotados', y así como el hombre ha podido, artificialmente mejorar las variedades o razas de animales, así también es posible hacerlo con el hombre, en términos de carácter e inteligencia ¿Cómo?, favoreciendo la reproducción en las parejas 'bien dotadas' y desalentando la procreación en los 'menos dotados'.

Gracias al deseo de demostrar todo lo anterior, recabando información, obteniendo gran cantidad de datos y manejando la información con mentalidad estadística, sentó las bases para el desarrollo de los métodos estadísticos.

Galton le escribió en 1869 a Charles Darwin, su primo por parte de su madre, haciéndole saber la verdadera crisis que le había causado la lectura de su obra *El origen de las especies*.

Podemos decir que Galton analiza la obra de Charles Darwin abordando dos de sus supuestos:

(1) Ruth Schwartz Cowan. Ibidem., P. 509.

En primer lugar, el mecanismo hereditario explicado por medio de su hipótesis de la Pangénesis; y en segundo lugar, la transmisión de caracteres adquiridos.

Respecto a la Pangénesis, Darwin cree que cada célula, cada órgano o cada tejido unitario produce y emite átomos diminutos a los que llama gémulas, y que éstas, más tarde esparcidas en todo el cuerpo por la sangre y otros jugos se reúnen cuando se requiere y después recrean aquellas partes de las cuales provienen. Los órganos sexuales contienen gémulas de todas partes del cuerpo las cuales se combinarán posteriormente en el embrión.

Para Galton, esta hipótesis es considerada potencialmente cuantificable.

La Teoría de la Pangénesis recoge todas las influencias que intervienen en la herencia y que le dan forma apropiada para comprenderla bajo análisis matemático. (1)

Realizó experimentos por dos años, en conejos de raza pura. Transfirió sangre de miembros de una camada en miembros de otra. Después de cruzarlos entre camadas esperaba que los hijos de una camada tuvieran características de la otra. Pero no fué así. De todos modos ensayó modelos matemáticos que pudieran explicar el mecanismo.

Bajo el supuesto de que el número de factores que intervienen en el proceso, se reduce al mínimo, considera el caso del color de pelo en animales. Si los padres tienen, digamos, blanco y negro, sus hijos resultarán de tres colores: negro, blanco y gris. En una población grande se deben

(1) Ruth Schwartz Cowan. Ibidem., p. 515.

observar las siguientes fracciones: 1/4 de blancos, 1/4 de negros y 1/2 de grises. En un siguiente cruce se tendrán cuatro colores: dos grados diferentes de gris, blanco y negro. Y así sucesivamente. Con la ayuda del triángulo de Pascal se puede seguir el número de partículas hereditarias y la verificación de sus fracciones o proporciones se podrían llevar a cabo al registrar las frecuencias relativas de los casos. Usá la expansión binomial:

$$(a + b)^x = a^x + xa^{x-1}b + \frac{x(x-1)}{2!}a^{x-2}b^2 + \dots + xab^{x-1} + b^x.$$

Donde el lado derecho nos da el número de combinaciones que se pueden dar y además su frecuencia, cuando a y b se combina de x formas.

Galton se convierte en uno de los críticos de Darwin más capaces(1) al negar experimentalmente la hipótesis de la Pangénesis. Y en su lugar propone su propia hipótesis, manteniendo la creencia en la existencia de las gémulas (aunque les llama de otro modo).

A todas las unidades orgánicas que existen, según Galton, en un huevo fertilizado les llama 'stirps'. Las cuales en su mayoría darán lugar a la estructura orgánica del ser, y las sobrantes a las células sexuales, permaneciendo estas sin ser influenciadas por las condiciones de vida del individuo.

Galton se opone a la suposición de que los caracteres adquiridos son hereditarios; nunca dejó de creer en la

(1) Peter J. Vorzimmer, Charles Darwin: The years of controversy The Origin of Species and its critics, 1859-1882. Temple University Press, 1970. p.255.

supremacía de la herencia sobre los factores ambientales en la constitución de los individuos.

El hecho de que haya utilizado la expansión binomial con el cual, se supone que los factores hereditarios, respecto a una sola característica, son pocos y están bien definidos, confirma tal creencia. Si en cambio, hubiera recurrido al modelo normal, tendría que haber aceptado a un número infinito de causas fundamentales que serían consecuentemente las ambientales. Sin embargo, es probable que por el año de 1876, Galton haya abandonado su primer modelo del proceso hereditario por el modelo de 'desviaciones del promedio', comprendiendo que solamente con desviaciones expresadas en 'unidades estadísticas' era posible representar el fenómeno en cuestión; esto lo hizo sin rechazar su concepción de la herencia.

Cuando se dice que se interesó en la herencia debido a que, una vez descubiertas sus leyes, éstas se utilizarían para mejorar la 'raza del hombre' (causando impacto social y político), tal parece que no se refiere al género humano en general. Veamos cuál es su deseo, en el momento de definir un concepto que será la base de una doctrina fundada por Galton: nos referiremos a la Eugenesia.

Eugenesia: Es la ciencia de la mejora del linaje, en modo alguno limitada a cuestiones de uniones juiciosas, sino que, particularmente en el caso del hombre, se ocupa de todas aquellas influencias susceptibles de dar a las razas mejor dotadas mayores oportunidades de prevalecer sobre las peores. (1)

(1) Pierre Thuillier. La Tentación de la Eugenesia, Mundo Científico. No. 38. (1985) 775.

Cabe preguntarse ¿cuáles son las razas mejor dotadas?. Sin embargo, lo conveniente es analizar cómo se desarrollaron las ideas que nos ocupan.

Galton proponía la hipótesis de que la salud y la inteligencia de la población se deterioraban, debido a la emigración de la gente más capaz, hacia las ciudades en donde su reproducción disminuía notablemente. Entendía que para demostrar lo anterior necesitaba información. Tomando características como: peso y estatura, se dedicó a obtener registros en las escuelas de Inglaterra; pero dicha información no era suficiente puesto que faltaban datos de sus padres.

Poco antes de 1876 Galton se encontraba cultivando chícharos* para estudiarlos en generaciones sucesivas con el fin de conocer, no precisamente el mecanismo hereditario, sino los efectos de tal mecanismo. Pero, ¿Cómo se le ocurrió que el seguimiento de características como peso y diámetro de chícharos por varias generaciones podrían sustituir temporalmente los registros que necesitaba de poblaciones humanas?; una respuesta inmediata podría ser la de tipo práctico, los datos de niños de las escuelas inglesas y los de sus padres implicaban mucho esfuerzo y además Darwin tiempo atrás le había sugerido llevar a cabo experimentos con plantas para la demostración de la Pangénesis. ¿Porqué no hacerlo ahora para investigar la posible degeneración de la población en las ciudades?.

* No podemos dejar de preguntarnos el porqué chícharos. ¿Hubo alguna relación con los experimentos de Gregor Mendel? Dejamos abiertas tales preguntas.

Quetelet había demostrado que la distribución de algunas características físicas humanas era de acuerdo al modelo normal y había sugerido el hecho como una ley invariable en el tiempo:

Si supieramos cuál ha sido la altura (media del hombre) de un siglo a otro, tendríamos una serie de medidas, las cuales expresarían la ley del desarrollo de la humanidad en lo que se refiere a la altura.(1)

Un paso decisivo por parte de Galton fué el empleo de la ley 'de las desviaciones'(2) para explicar los resultados de los chícharos de diferentes diámetros y pesos y sus productos, de esta investigación se publicó: *Experimentos*

(1) D. Landau, P.F. Lazarsfeld. *Ibidem.*, p.830.

(2) Es interesante conocer la importancia que le concede Galton a la curva normal: 'Necesito difícilmente recordarle al lector que la ley del error sobre la cual se basan los Valores Normales fué pensada para el uso de los astrónomos y otros quienes estén interesados con precisión extrema de mediciones y sin la menor idea, hasta el tiempo de Quetelet, que ellos podrían aplicarla a mediciones humanas. Pero, los Errores, Diferencias, Desviaciones, Dispersiones y Variaciones individuales, todas emanan de la misma clase de causas. Objetos que llevan el mismo nombre o que pueden ser descritos por la misma frase, son de este modo reconocidas como miembros de la misma clase, especie o como queramos llamarle al grupo. Por otro lado, cada objeto tiene diferencias peculiares en sí mismos, por lo que se distinguen de otros.

Esta afirmación general es aplicable a miles de ejemplos. La ley del error encuentra su fundamento donde las peculiaridades son completamente debidas a influencias combinada de una multitud de 'accidentes', en el sentido en que la palabra ha sido definida. Todas las personas versadas en Estadística están conscientes de que esta suposición lleva a la Variabilidad dentro del campo de las leyes del azar, con el resultado de que las frecuencias relativas a las Desviaciones de diferentes cantidades admiten ser calculadas, cuando estas cantidades sean medidas en términos de cualquier unidad, por sí misma contenida, de variabilidad, como muestra Q'. (E.S.Pearson. Some Incidents in the Early History of Biometry and Statistics, Includido en: E.S.Pearson, M.G. Kendall, Studies in the History of Statistics and Probability. London Griffin. 1970. p.327.

con plantas sobre las causas de uniformidad estadística en generaciones sucesivas en 1876, antecedente de *Las Leyes Típicas de la Herencia* aparecido un año más tarde. En esta última obra, tenemos las conclusiones de sus experimentos.

1. Los descendientes de plantas extremosas (chicharos con diámetros muy pequeños o muy grandes) estaban más cerca de la media poblacional que sus 'padres' lo estaban, y
2. un grado de variación, una unidad estadística*, es la misma medición absoluta, ya sea que se considere a los productos de plantas extremosas, o bien a los productos de plantas normales.**

Para su explicación estadística hace uso de un dispositivo mecánico llamado *quincunx*: el modelo normal se puede formar por la caída de balas de plomo desde un embudo sobre el enrejado de rectas verticales o alfileres formando compartimientos; el modelo nos representa la distribución de la característica, en este caso, el diámetro de chicharos. En una siguiente generación (ver la siguiente gráfica) la distribución de la característica estudiada sería explicada, según la misma ley con mayor o menor variabilidad. Si se ha concluido que los descendientes en una primera generación 'están más cerca de la media poblacional', la curva sería con menos dispersión, pero sin embargo si pensamos en que podemos retirar la base del modelo, al caer libremente las balas, se reproduciría la

* Pearson señala que la 'unidad estadística' a la que se refiere es a Q (rango intercuartílico) al que llama 'error probable'. Ruth Schwartz Cowan, *Ibidem.*, p.517.

** Es decir, la variabilidad de los productos de semillas pequeñas era igual a la variabilidad del producto de semillas grandes.

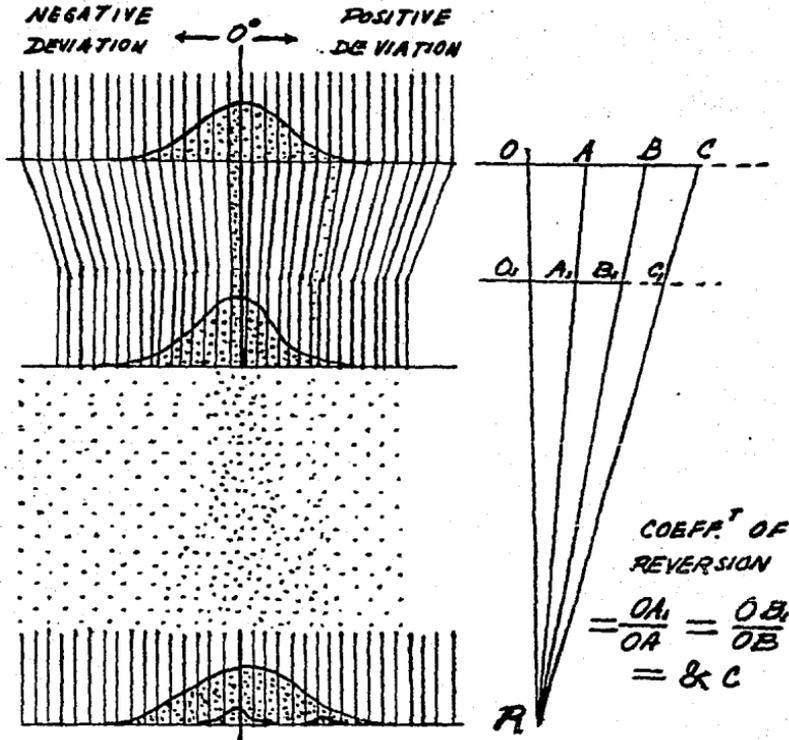


Ilustración de Galton sobre la naturaleza de la reversion tomada de " Las Leyes Típicas de la Herencia ", Proceedings of the Royal Society Institution. February 9, 1877.

misma curva que la de sus padres o antecesores. Es decir, la variabilidad se mantendría constante de una generación a otra, de acuerdo con la segunda conclusión descrita arriba. Pero entonces ¿Qué significa que la población empeore?

De las dos revelaciones a partir de los experimentos con las semillas de chícharos, la primera constituye el principio de lo que será una de las principales contribuciones de Galton a la estadística. El hecho de que las generaciones tienden hacia la media o promedio poblacional, lo considera como un rasgo hereditario y cuantifica esta 'reversión' en un valor numérico:

Reversión es la tendencia del tipo filial a separarse del tipo paternal, revertiéndose a lo que ásperamente y quizá descrito moderadamente como el tipo ancestral medio. Si la variabilidad familiar fuera el único proceso en descendencia simple, que afecte las características de una muestra, la dispersión de la raza con respecto a su tipo ideal medio aumentaría indefinidamente con el número de generaciones, pero la reversión registra este aumento y lo detiene. (1)

Graficando el tamaño de las semillas padres con el tamaño de sus productos, obtiene una línea recta cuya pendiente la considera como el coeficiente de reversión.

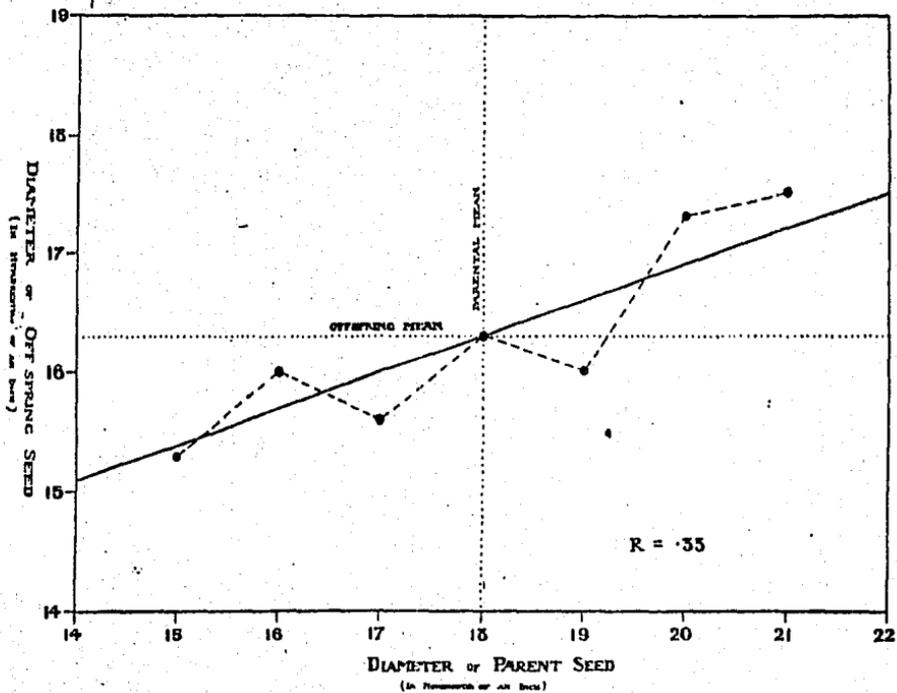
Pasarán casi 10 años para que Galton llegue finalmente a tener claro el concepto de regresión y correlación, pero el paso primordial estaba dado con lo que sería la primera recta de regresión en un análisis estadístico. (Ver la siguiente gráfica).

En el lapso comprendido entre 1877-1888 Galton se dedica a

(1) Francis Galton. Leyes Típicas de la Herencia. p. 10. Citado en: K. Pearson: Notes on the History of Correlation, en E.S. Pearson, M.G. Kendall. Ibidem., p. 193.

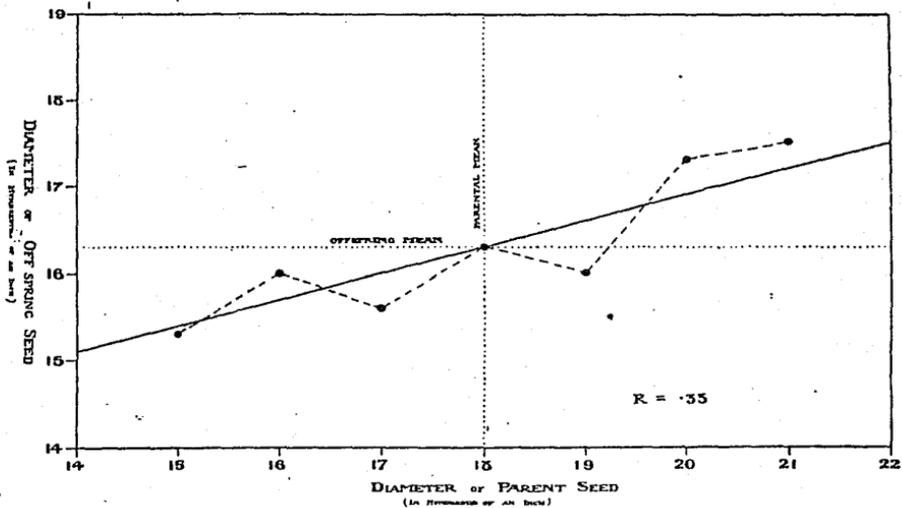
INHERITANCE IN SIZE OF SWEET PEA SEEDS.

CALTON - ROYAL INSTITUTION LECTURE 1877



INHERITANCE IN SIZE OF SWEET PEA SEEDS.

CALTON - ROYAL INSTITUTION LECTURE 1877



obtener información tanto de centros escolares, como de un sitio creado exprofeso, llamado Laboratorio Antropométrico. Este laboratorio ideado y dirigido por el mismo, consistía en oficinas para medir y pesar a seres humanos, se invitaba a que los padres mandaran periódicamente a sus hijos; inclusive se llevaba un registro personal de habilidades y facultades como: altura, peso, palma de la mano, potencia al respirar, fuerza de tiro y apretujamiento manual, potencia al soplar, oído, vista, sentido de color y otros datos personales. Galton proponía que en el futuro este laboratorio se convirtiera en una dependencia de Estadísticas Nacionales para apoyar las funciones de gobierno.

Ahí, en el laboratorio se obtuvieron datos de 9,337 personas que servirían para investigar la selección sexual, el coeficiente de reversión, la constancia de la variabilidad y en caso contrario la influencia de la selección natural y de la fertilidad diferencial.

Antes de continuar, es oportuno analizar la metodología que Galton va introduciendo en el manejo de la información que obtenía. En un principio, señala que no se debe limitar un análisis estadístico, solamente a calcular un promedio por ser un hecho aislado, sino que se debe proceder a 'calcular una distribución', y asignar un esquema gráfico para visualizarla. El primer trabajo (1875)(1) de Galton consistió en inventar un esquema de distribución de frecuencias acumuladas y su gráfica la ojiva. Los datos que

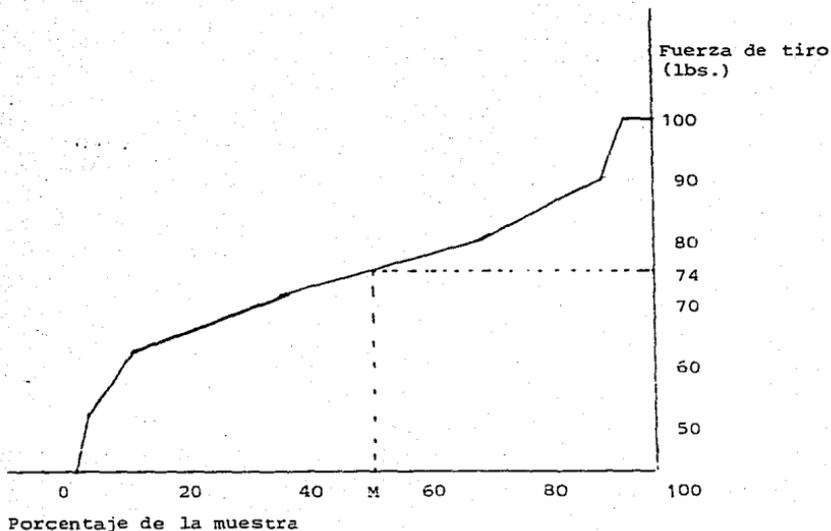
(1) John Dewey. Galton's Statistical Methods. Journal of American Statistical Association. Vol. I. (1901) 331.

maneja son los correspondientes a la fuerza de tiro (pull) de 519 hombres:

Fuerza de tiro	Porcentaje
abajo de 50 Lbs.	2
" " 60 "	10
" " 70 "	37
" " 80 "	70
" " 90 "	91
" " 100 "	95
arriba de 100 "	100

Para la tabla anterior propone una gráfica (ver la siguiente hoja) de la forma siguiente:

En una línea base se localizan los porcentajes de fuerza y el número de libras e nuna línea perpendicular a la derecha. Después, a partir de cada porcentaje se erige una perpendicular a una altura igual al número de libras correspondientes. Al unir los puntos de intersección se obtiene una línea quebrada, la cual se suavizaría pareciéndose a una curva al tomar suficientes datos y los intervalos de fuerza fueran, digamos, cada libra. La figura resultante es el esquema de distribución. Su uso es evidente si se toma la fuerza medida de cualquier individuo, por ejemplo, 74 libras y se traza una línea horizontal hasta dar con la curva y de ahí una perpendicular hasta encontrar la línea base, así el porcentaje de la muestra correspondiente en este caso cae en 50, es decir, ocupa el lugar medio y la fuerza es la fuerza media; a esta posición le asigna la letra M, parámetro muy importante para Galton del cual observa las siguientes propiedades:



a) la probabilidad de que cualquier observación (medición de fuerza de un individuo) desconocida previamente caiga antes o después de M es la misma; b) el valor más probable de cualquier medición previamente desconocido es M ; c) si la curva del esquema es simétrica bilateralmente con respecto a M , entonces M es idéntico al promedio ordinario o media aritmética.

Del análisis anterior, podemos constatar la inclusión de conocimientos teóricos de la probabilidad en el manejo de datos empíricos, así como una explicación entre lo observado y lo esperado.

Aprovechemos también para describir otra constante muy útil a Galton simbolizada por Q (el rango intercuartílico).

Se toman perpendiculares a los porcentajes 75 y 25, se restan sus valores correspondientes y se divide entre dos. El resultado lo considera como una medida de la pendiente de la curva de distribución; siendo realmente una medida de dispersión. En el ejemplo descrito, el valor de Q es de 9 libras. Quiere decir que la mitad de la población difiere menos de 9 libras valor medio (la mediana) de la población entera, en ambos sentidos. El valor de Q lo calcula para cada una de las características medidas en su laboratorio, estimado inclusive una Q promedio con todas ellas.*

En 1886 y 1888 publica *Parecido Familiar en la Estatura y Correlaciones y sus Mediciones, principalmente a partir de Datos Antropométricos*, respectivamente, ambos de la Royal Society of London (proceedings 40 y 45). En ellos realiza un análisis exhaustivo de toda su información. Para darnos una idea del procedimiento que siguió, en lo que se refiere a la estatura, Galton decía que para pronosticar la estatura de un hermano en función de la estatura de otro hermano era necesario formar un grupo con aquellos que tuvieran una misma estatura conocida y construir un conjunto de estaturas de todos los hermanos del primer grupo llamado *cofraternidad* y de ahí discutir; la variación en altura de todos los individuos alrededor de la gran media, la variación de los individuos de las *cofraternidades* alrededor de sus medias *cofraternales* respectivas.

Cuando calculó el coeficiente de reversión para las estaturas entre padres e hijos le resultó $2/3$, el mismo que obtuvo de los chicharos. Al invertir los valores en los ejes o sea cuando los valores individuales de los hijos

* Si la curva, a partir de la cual se calcula Q , es simétrica entonces Q es lo que se llama error probable (John Dewey, *Ibidem.*, p. 333).

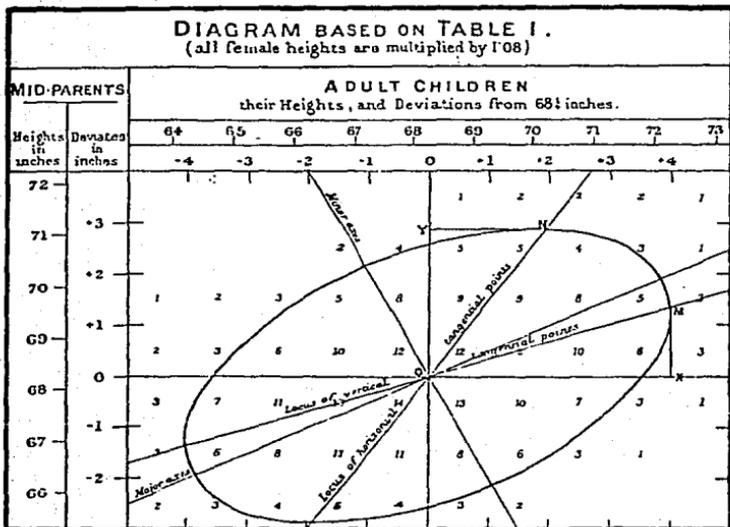
fueron graficados contra la altura media de los padres, el coeficiente fué igual a $1/3$. Esto lo explica de la siguiente manera:

El número de individuos que están cercanamente a la media, es tan preponderante, que a un hombre excepcional se le encuentra más frecuentemente como el hijo excepcional de padres mediocres, que al hijo promedio de padres muy excepcionales. (1)

Junto con la línea recta (de regresión) que trazó, elaboró lo que llamó 'tabla de correlación', cuyas entradas eran por un lado la estatura de todos y cada uno de los hijos y por otro la estatura media de los padres (la estatura media paternal era el promedio de la estatura de los progenitores, después de que la altura de la madre había sido transformada a unidades que eran probabilísticamente similares al del padre, multiplicando por 1.08). Si por ejemplo la altura media de los padres es de 68 pulgadas y si las alturas de sus cuatro hijos habían sido 64, 65, 66 y 67 entonces se marcaba en cada una de las cuatro columnas verticales correspondientes a aquellas estaturas de los hijos en la línea horizontal de 68 pulgadas. (Ver la gráfica de la siguiente página). Los resultados los graficó haciendo un conteo de las intersecciones.

Galton realizó un gran avance cuando decidió cambiar el coeficiente de reversión por coeficiente de regresión y esto, no por el cambio de nombre, sino porque, sucedieron dos hechos significativos. El primero es notorio al observar las gráficas 1 y 2. En la primera, el coeficiente de

(1) Ruth Schwartz Cowan, Ibidem., p. 522.



Encontré difícil, al principio, captar el significado completo de los datos aunque pronto descubrí relaciones curiosas y aparentemente muy interesantes entre ellos. Líneas trazadas a través de las entradas del mismo valor formaban una serie de elipses similares y concéntricas. El centro común quedaba en la intersección de las líneas vertical y horizontal que corresponden al valor de 68 1/4 pulgadas (el valor, tanto de la media poblacional de los padres como la de los hijos). Sus ejes estaban similarmente inclinados y los puntos donde cada elipse era tocada por una tangente horizontal quedaban en una línea recta inclinada con respecto a una horizontal en la razón de 1/3. (1)

(1) Francis Galton. Natural Inheritance. p. 191. Citado por Ruth Schwartz Cowan. Ibidem., p. 322.

reversión es igual a la pendiente de la recta, cuyo valor depende de la unidad de medida (dado que los valores son absolutos) de las variables involucradas; sin embargo, en la gráfica 2, las estaturas han sido convertidas a valores que actualmente les llamamos tipificados o estandarizados; es decir, se han tomado las diferencias de los valores con respecto a la mediana M , divididas por su variabilidad medida en términos de rango intercuartílico. Con esto tenemos el camino preparado para medir el grado de relación entre cualquier par de variables.

El segundo suceso significativo que se podría considerar de tipo anecdótico, pero que finalmente fué determinante para la generalización de la medida (que en principio fue) de la tendencia a la medianía o mediocridad.

Galton después de observar las elipses concéntricas que le resultaron al graficar la tabla de estaturas entre padres e hijos, decide recurrir a J.D. Hamilton Dickson matemático de Cambridge para pedirle que analizara la superficie de frecuencias de error proporcionándole solamente el error probable de la distribución de la característica paterna, el error probable de la generación de los hijos, así, como el coeficiente de reversión de este último sobre el primero. Y para sorpresa y admiración de Galton (1), Dickson dedujo, usando las matemáticas de la probabilidad, a partir de la escueta información y lo que es más importante, sin conocer el hecho de que las variables tenían que ver con caracteres hereditarios, el diagrama casi idéntico al que Galton había obtenido de forma empírica. El nombre de regresión liberaba a la medida del contexto que le dió origen, el de la herencia.

(1) Ruth Schwartz Cowan. Ibidem., p. 523.

En la introducción de nuestro trabajo hemos dicho que la obra más notable de Francis Galton fué *Herencia Natural* (1889). Ahora estamos listos para justificar tal afirmación. En esta obra resume el desarrollo de sus investigaciones por lo menos de diez años atrás y propone la llamada Ley de Herencia Ancestral, involucrando en ella sus revelaciones anteriores. Esta ley, llamada Ley de Galton por sus seguidores, se origina por ciertas interrogantes en el sentido de que los descendientes se parecen a sus padres aunque no completamente y Galton se cuestiona: ¿qué tanto se parecen? ¿influyen sus antecesores más remotos?, ¿en qué medida?.

Al analizar sus datos, en particular lo referente a las estaturas, obtuvo su coeficiente de reversión o regresión. Si la altura promedio de padres es $M + D$ (M : media de la población), la altura promedio de los hijos es de $M + (2/3)D$, siendo $2/3$ el coeficiente de regresión, que cuando se calcula de la media paterna sobre los hijos resulta $1/3$. Considera estos y otros parentescos calculando sus coeficientes y a veces deduciéndolos. Por ejemplo, la reversión de los nietos sobre los abuelos es $1/9$ ya que $1/3 \cdot 1/3 = 1/9$. Con los valores obtenidos, Galton se plantea un siguiente problema ¿Cuánto, de una característica se debe a cada ancestro? o mejor aún ¿Cuánto de la desviación del promedio se debe a cada ancestro?(1)

Al hacer la afirmación de que los padres medios* cuya estatura es $P + D$ tienen hijos con estatura promedio $P + (2/3)D$, se supone que no se ha tomado en cuenta lo ancestral previamente. Todavía, aunque no se sepa nada

(1) R.G. Swinburne. Galton's Law-Formulation and Development. Annals of Science. 21 (1965).

* Con padres medios se refiere al promedio de las estaturas de ambos padres, después de que se multiplica la estatura de la madre por 1.08.

de ellos, algo está implícito y ha sido tácitamente aceptado y se requiere que sea eliminado antes de que podamos captar la cantidad del legado paterno puro y simple. Lo que sea, lo debemos ahora tratar de descubrir. Cuando se habla de regresión converso, se demostró que una peculiaridad en un hombre implicaba una peculiaridad de $1/3$ de esa cantidad en su padre medio. Llamémosle a esa peculiaridad del padre medio D , entonces la peculiaridad implicada del padre medio del padre medio, es decir, del abuelo medio sería, según la suposición anterior $(1/3)D$ y del bisabuelo medio $(1/9)D$ y así sucesivamente. Por tanto la propiedad heredable total se acumularía en $D(1 + 1/3 + 1/9 + \dots) = (3/2)D$.

Ahora bien, ¿Cómo se hereda esa cantidad?, la contribución de cada ancestro es equitativa? ¿Los ancestros contribuyen más o menos?

Primero, supongamos que los legados de varias generaciones se recaban equitativamente, entonces como una acumulación de contribuciones ancestrales cuya suma es $(3/2)D$, proporciona una herencia efectiva solamente de $(2/3)D$; de aquí que cada pieza de propiedad heredable deber haber sido reducida a $4/9$ de su cantidad original, debido a que $(3/2)(4/9) = 2/3$.

Segundo, supongamos que la contribución no es uniforme pero se repite en cada transmisión sucesiva y es igual a $1/r$ de la cantidad de la propiedad en cada estadio. En este caso, la herencia efectiva sería

$$D(1/r + 1/3r^2 + 1/3^2r^3 + \dots) = D 3/(3r - 1)$$

lo cual debe ser como antes; igual a $D(2/3)$, por lo tanto $1/r = 6/11$.

Tercero, podría suponerse que el ancestro medio en una generación remota, en promedio, debe contribuir más con el niño que con el padre medio, pero esto es contrario a lo observado.

Los resultados de cada una de las suposiciones son, por lo tanto, que las peculiaridades paternas medias puras y simples influyen en el hijo, en $4/9$ de su cantidad. Estos valores difieren ligeramente de $1/2$ y su media es cercana a $1/2$. Así que, bien podemos aceptar este resultado. Por lo tanto la influencia pura y simple paternal media puede ser tomada como $1/2$ y la de los abuelos promedio $1/4$, etc. Consecuentemente, la influencia del padre individual sería $1/4$, del abuelo individual $1/16$, etc.*

* R.G. Swinburne, Ibidem., p. 19.

Al seguir los razonamientos anteriores se pueden sacar dos errores fundamentales: 1) los valores que Galton considera que difieren ligeramente y aunque sean aproximadamente iguales ($4/9$ y $6/11$) provienen de puntos diferentes, $4/9$ representa la contribución igual de cada ancestro, mientras que $6/11$ representa la contribución paternal media que se eleva a potencias más altas para generaciones más lejanas. Galton escoge la segunda ecuación sin argumentar su decisión y toma a $1/2$ como la constante; 2) el segundo error consiste en que la forma de la ecuación que selecciona, no es precisamente la requerida para exponer lo que quiere explicar, a saber la fuente de desviaciones de la media. Galton toma a lo paternal medio con una cierta desviación, (cuyos ancestros tienen sus desviaciones en términos de los coeficientes de correlación) y después argumenta que ya que el hijo tendrá una desviación en promedio de $(2/3)D$, el problema consistirá en contabilizar esta desviación en términos de las contribuciones de varios ancestros. Pero su problema no es contabilizar las desviaciones de hijos. Para el hijo con una desviación D , lo paternal medio tendrá en promedio una desviación de $(1/3)D$, lo abuelo medio de $(1/9)D$ y así sucesivamente. Por lo tanto la ecuación del segundo tipo debe ser igual a D . Esto da una participación negativa de las contribuciones de cada ancestro; el padre medio contribuye con $1\ 1/2$ veces su peculiaridad, el abuelo medio con $1\ 1/4$ veces, etc.

Las falacias señaladas hacen pensar, según R.G. Swinburne,(1) que Galton tenía el resultado preconcebido y que solamente intentó sustentarlo experimentalmente.

(1) R.G. Swinburne, Ibidem., p.20.

En un artículo de 1897, *La Contribución Promedio de cada uno de varios Ancestros a la Herencia Total del Hijo*, aplica su ley de Herencia Ancestral con datos coleccionados por Sir Everett Millais correspondientes al color de pelo en perros. Aquí, establece su ley sin calcular coeficientes de regresión y propone proporciones de perros tricolores y no tricolores en cualquier generación, calculadas en base a datos de colores de varios ancestros.

Aunque los intentos de Galton de revelar las leyes de la herencia por los medios descritos arriba hayan sido en vano, su metodología se convertiría en el fundamento para la aplicación de la Estadística a la Biología, atrayendo a una generación de investigadores, como veremos adelante el caso de Karl Pearson. Finalmente, veamos lo que se refiere al concepto de Correlación.

En el artículo *Correlaciones y sus mediciones, principalmente a partir de datos antropométricos*, leído en la *Sociedad Real* en 1888 aparece el término correlación por primera vez en nuestra materia. (1)

Co-relación o correlación de estructura es una frase muy usada en biología y no menos en la rama que se refiere a la herencia; la idea es todavía más común que la frase, pero no tengo conocimiento de algún intento de definirla claramente, de investigar su modo de acción o de mostrar como medir su graduación.

Dos órganos variables se dice que son correlacionados cuando la variación de uno es acompañada, en promedio, con variación más o menos del otro y en la misma dirección.

Se muestra mas explicito respecto al concepto en cuestión,

(1) Karl Pearson. Notes on the History of Correlation. Contenido en: E.S.Pearson, M.G. Kendall. Ibidem., p. 199.

cuando asevera que:

Es fácil ver que la co-relación debe ser la consecuencia de las variaciones de los órganos debidas en parte a causas comunes. Si se debiera completamente a causas comunes la co-relación sería perfecta, como es el caso aproximadamente de las partes dispuestas simétricamente del cuerpo humano. Si no se debiera de ningún modo a causas comunes, la co-relación sería nula. Entre estos dos extremos hay un número sin fin de casos intermedios y se mostrará, como la estrechez de la correlación en cualquier caso particular admite ser expresada por un número único.

El número al que se refiere seguramente es el valor del coeficiente de reversión o regresión. Por cierto que Darwin (1) habla sobre la correlación entre órganos en la estructura de los animales.

Por eso, si el hombre sigue eliquendo y por lo tanto aumentando alguna peculiaridad (de los animales bajo domesticación) es casi seguro que modificará sin querer otras partes de la estructura, debido a las misteriosas leyes de la correlación.

Paralelamente a sus estudios sobre la herencia e inclusive como parte de sus investigaciones sobre esta, Galton había desarrollado una disciplina que era imprescindible en un laboratorio dedicado a la obtención de datos sobre características humanas, nos referimos a la antropometría (es justo señalar a Quetelet como el iniciador de este campo). Desde la invención de utensilios hasta el desarrollo de técnicas elaboradas hubo necesidad de refinar para cuantificar, clasificar y registrar datos antropométricos. Es natural, por lo tanto, que se interesara por estudios de esta naturaleza. Alphonse

(1) Charles R. Darwin. El Origen de las Especies. México: Editorial Diana. 1983, (15a. imp.) p. 29.

Bertillon oficial francés interesado en la clasificación de criminales en prisión, ideaba un método para distinguir a los individuos en términos de un conjunto de mediciones de sus partes corporales. Pensó que era posible la plena identificación de una persona a través de valores como el diámetro craneal, longitud del pie, longitud del dedo intermedio del pie y de la mano izquierdos, etc. Cuando Galton ponderó el trabajo de Bertillon(1) en principio estuvo de acuerdo con él, pero sin embargo al poco tiempo cambió de parecer. Se tenían antecedentes (de Darwin, por ejemplo) de que las medidas debían ser dependientes entre sí, de tal forma que era inútil pensar en conjuntos únicos de mediciones para cada persona, sino más bien dichos conjuntos serían de tipos de personas. Procedió a tabular y graficar datos tal como lo hizo con las estaturas obteniendo coeficientes de regresión, a pesar de que los factores no estaban medidos en unidades iguales (digamos, la altura en metros y el volumen torácico en centímetros cúbicos). Al reflexionar sobre sus resultados concluyó que no era posible la propuesta de Bertillon bajo las condiciones de la correlación y concluyó que el parecido familiar era un caso particular del concepto más amplio de la correlación. El análisis crítico de la obra de Bertillon condujo a Galton al concepto de correlación por una senda diferente. (2)

Como respuesta alternativa al problema de la identificación personal Galton investigó las huellas dactilares encontrando ciertas regularidades que le sirvieron para elaborar el libro *Finger Prints* (1892) cuya influencia se dejó sentir

(1) Francis Galton. Personal Identification and Description. Nature, June 21, 1888. (Citado en Ruth Schwartz Cowan, Ibidem., p. 526).

(2) Ruth Schwartz Cowan. Ibidem., p.527.

al establecerse una oficina con el fin de recabar información al respecto, hecho que incluso se llevó a cabo en varios países.

Galton no avanzó, en lo que se refiere a la correlación, mas allá de lo presentado en su libro *Herencia Natural*. Se contaba ya con una medida de asociación entre dos variables con posibilidad de aplicarse en muchos y diversas situaciones; tarea que les correspondería a otros más jóvenes. El primer paso estaba dado. Al mismo tiempo que recibía honores otorgados por varias asociaciones científicas, se dedicó a ser consejero y tutor de quienes lo reconocían como tal.

Continuemos con el estudio de mayores contribuciones a la Estadística.

KARL PEARSON

(1857 - 1936)

El análisis de la obra de Galton estaría incompleto si no se continuara con la de otro científico, casi contemporáneo suyo llamado Karl Pearson quien ha sido considerado el "fundador de la ciencia de la estadística"(1). Y esto no solamente porque prosiguió de manera directa la obra y el avance logrado por Galton en el momento en que se necesitaban nuevos ímpetus, tanto en los contenidos específicos de las investigaciones como en el liderazgo y presencia en actividades iniciadas y auspiciadas por quien llegó a ser su guía admirado (escribió una monumental biografía de Galton en tres volúmenes). Sino también debido a que por méritos propios supo mejorar y acrecentar con su labor admirablemente prolífica, el campo de la estadística considerándole junto con Galton, los iniciadores de un gran primer período del desarrollo de esta disciplina.(2)

Nacido en Londres en 1857, tuvo como padre a William Pearson abogado consejero de la Reina y líder en la Cámara de la Cancillería, con cualidades tales como: una gran habilidad, energía mental y física excepcionales y un agudo interés en la investigación histórica, características que mostrara también su hijo Karl.(3) A lo largo de su preparación

-
- (1) Samuel Wilks. Karl Pearson: Founder of the Science of Statistics. Scientific Monthly. 53. (1941). 249.
 - (2) Helen M. Walker. The Contributions of Karl Pearson. Journal of American Statistical Association 53, (1958) 12.
 - (3) W. Kruskal, J.M. Tanur. International Encyclopedia of Statistics New York (A) 1978. p. 691.

académica, Pearson dió muestras de una amplia gama de intereses siempre con gran vigor intelectual y con frecuentes y placenteras polémicas dirigidas hacia las imposiciones autoritarias, haciendo gala de gran confianza en sí mismo. Ingresó al University College en 1866, al que tuvo que abandonar siete años después por razones de salud; regresando a la edad de 18 años, pero esta vez al King's College para estudiar matemáticas, obteniendo su título con honores cuatro años más tarde.

Se interesó sobremanera por la cultura alemana; en todo aquello que se refiere al folklore, sociedad, pensamiento e historia del pueblo alemán; estuvo en Heidelberg y Berlín estudiando física y compenetrándose con el socialismo, leyendo a Martín Luther, Karl Marx, Maimonides y Spinoza. Estos estudios no lo distrajeron de las matemáticas y de la física. Al regreso a Inglaterra se ocupó de escribir sobre el movimiento de los fluidos y la teoría de la elasticidad. Durante su estancia en Cambridge recibió, según una pequeña nota autobiográfica, intensas influencias de sus profesores, de su tutor y examinadores, entre quienes se encontraba Isaac Todhunter autor del libro *Historia de la Teoría Matemática de la Probabilidad desde Pascal hasta Laplace* (1865) y cuyo examen al ser resuelto por Pearson, éste se hizo acreedor a una mención en un manuscrito que dejó incompleto Todhunter llamado *Historia de la Elasticidad* (1886-1893); gracias al cual, Pearson fué solicitado para completar dicha obra a la muerte de Todhunter.

En 1884 se inicia en la cátedra de Matemáticas Aplicadas y Mecánica en el University College de Londres, dedicándose a dar clases de matemáticas a los alumnos de ingeniería, además de sus escritos sobre elasticidad y conferencias sobre socialismo e historia alemana. Su primer libro

publicado fué *El Nuevo Werther* (1880) antecedente de otros dos: *La Etica del Libre Pensamiento* (1888) y *Los Elementos de la Ciencia* (1892). En estas obras se manifiesta su curiosidad filosófica, su deseo de hacer del mundo algo entendible y plantea sus primeras investigaciones sobre historia y religión. De su primer libro se ha extraído el siguiente párrafo:

No esperes que yo te hable de ningún sistema o método; el tiempo parece demasiado corto y el mundo demasiado vasto para eso. Corro de la ciencia a la filosofía, y de la filosofía a nuestros antiguos amigos los poetas; y entonces, cansado por el excesivo idealismo, se me antoja práctico volver a la ciencia. ¿Has pensado alguna vez que todo lo que existe en el mundo es digno de conocerse, que no hay nada en el mundo que no merezca ser estudiado?. Los gigantes de la literatura, los misterios de los espacios de muchas dimensiones, los intentos de Boltzmann y Cookes de penetrar en el laboratorio mismo de la Naturaleza, la teoría kantiana del universo, y los últimos descubrimientos de la embriología con sus maravillosos relatos sobre el desarrollo de la vida -que inmensidad más allá de nuestro alcance-. Mientras paseo aquí por las montañas, que sensación de vergüenza me llena, por no saber nada de botánica o geología, porque, aunque puedo contemplar con amor las flores y rocas, ignoro la historia de vida en la una, y la historia del mundo en la otra. Pero mis sentidos se turban; siento que nuestra naturaleza es demasiado pequeña, nuestras fuerzas son nada. Feliz es el hombre que no sólo puede decir, "no sé nada", sino que puede contentarse con esa nada.

Es a este problema ...la unión de lo divino con lo humano, el sentimiento de cuya existencia conduce a su simbolización en el hombre-dios, Cristo a quien deseo dedicar mi vida. Para esto debo estudiar lo que pensaron los grandes cerebros del mundo y aprender los motivos de su acción. (1).

En la segunda obra, *La Etica del Libre Pensamiento*, Pearson

(1) E.S. Pearson. Ibidem. p. 3.

resume varias conferencias dictadas por él entre 1883 y 1887. En ellas discurre en cuánto a interrogantes que lo van conduciendo a una formación ético religiosa que regirá su futura labor científica.

El eterno por qué. ¿Por qué estoy aquí?, ¿Qué relación guardo yo, una parte con el todo, con la suma de todas las cosas materiales y espirituales?, ¿Qué conexión tiene lo finito con lo infinito?. Para contestar este cuestionario se desarrolló la religión. Pero el gran peligro de las relaciones existentes es este. No concuerdan con nuestro real conocimiento de la relación de lo finito a lo infinito, ocultan nuestra anterior ignorancia con ayuda de la imaginación. El mito reemplaza al verdadero conocimiento...

La ciencia y la filosofía se nos presentan continuamente con perspectivas más amplias de la relación del hombre con la naturaleza, del pensamiento individual con el pensamiento abstracto, y cada religión debiera por ello estar en un estado de desarrollo. Pero ellas, difícilmente se encuentran en ese estado, insisten en explicar por mitos, por dogmas persistentes. El rechazo de toda explicación mística, la franca aceptación de todas las verdades descubiertas con respecto a la relación de lo finito con lo infinito es lo que yo denomino *Libre pensamiento* o verdadero conocimiento religioso. En otras palabras, el libre pensador en mi sentido del término, posee más religión real que cualquier creyente en mitos...; su verdadero conocimiento lo hace, en el más elevado sentido de las palabras, un hombre religioso. Para ser un libre pensador no es suficiente desechar todas las formas de dogmatismo, ni atacarlas con burda sátira: ésta es sólo una acción negativa. El verdadero libre pensador debe poseer los mayores conocimientos de su época; debe pertenecer a su siglo y tomar nota de lo que ha alcanzado el pasado y de lo que está logrando el presente; aún mejor si él mismo trabaja por el incremento del conocimiento humano o por su difusión entre sus compañeros- tal hombre puede verdaderamente ser llamado sumo sacerdote del libre pensamiento...

Los más grandes poetas, filósofos, naturalistas, hombres tales como Goethe, Spinoza y Darwin han sido todos libres pensadores; ellos se esforzaron independientemente de las creencias dogmáticas y armados con los más elevados conocimientos y pensamientos de su tiempo, por echar luz sobre el problema de la vida.

Se ha querido prolongar la cita anterior con el fin de mostrar el terreno fértil en el cual caerán, como vemos, las semillas de grandes pensadores. En esta misma obra se tienen otras conferencias de las cuales las últimas son "Las Bases Morales del Socialismo" y "Los Derechos de la Mujer", ambos ejemplos de su pensamiento 'liberal' que incluso le llevó a sufrir severas críticas pero que no le impidieron seguir discutiendo cuestiones sociales y sexuales 'científica e imparcialmente' como él sostenía. Por ejemplo cuando se planteaba si la capacidad intelectual de los padres se transmitía por igual entre los hijos hombres y las mujeres, respondía que la respuesta debía postergarse para cuando se pudiera llevar a cabo una investigación seria.

Pearson abandona su cátedra del University College para tomar una sobre Geometría en el Gresham College en 1890, cuyo programa estaba sujeto a la voluntad del conferenciante. El primer curso lo tituló 'el objeto y los conceptos de la Ciencia'. Sus notas servirían más tarde, para publicar 'Los Elementos de la Ciencia'; obra que según Samuel Wilks(1) perdurará por largo tiempo como obra clásica en la filosofía de la ciencia.

En un segundo curso dado entre 1891 y 1894, llamado en esta ocasión 'La Geometría de la Estadística y las Leyes del Azar', encontramos un cambio inusitado respecto a sus múltiples áreas de interés. Algo había sucedido para que Pearson abordara el estudio de la probabilidad y la estadística, los incluyera en su programa de clases y les otorgara gran importancia para la investigación científica. ¿Qué o quiénes influyeron en él para desviar su atención de los temas y aplicaciones tradicionales de las matemáticas

(1) Helen M. Walker. Ibidem., p. 250.

hacia lo que sería su aportación a la ciencia, y por lo cual, ahora nos ocupa su vida y su obra. Recordemos que en 1889 Karl Pearson se daba tiempo tanto para su cátedra en el University College como para sus escritos y conferencias antes mencionados, cuando tienen lugar dos acontecimientos que definirán su trayectoria futura.

El primero es la publicación de *Herencia Natural* de Galton en ese año. Y segundo, un año más tarde, W.F.R. Weldon es asignado a la cátedra de Zoología, también en el University College.

Respecto a la influencia de Galton en quien será un digno sucesor, tenemos que en la introducción de *Herencia Natural*, Pearson había leído:

Es de sumo interés (el tema de la obra) en sí mismo. Nos familiariza con la medición de la variabilidad y con leyes curiosas del azar que se aplican a una amplia diversidad de materias sociales. Puede decirse que esta parte de la investigación corre por un camino a alto nivel, que ofrece amplias perspectivas e inesperadas direcciones y desde el que pueden hacerse fáciles descensos hacia metas completamente distintas de las que tenemos que alcanzar ahora. (1)

lo cual, le hizo reflexionar, provocando la siguiente reacción:

'camino a alto nivel', 'amplias perspectivas en inesperadas direcciones', 'fáciles descensos hacia metas completamente distintas' -había aquí campo para un vagabundo aventurero... Interpreté que esas frases de Galton significaban que había una categoría más amplia que la causalidad, a saber la correlación de la cual, la causalidad era sólo el límite y que esta nueva concepción de la correlación traía a la

(1) E.S.Pearson, Ibidem., p. 45.

psicología, antropología, medicina y sociología, en gran parte, dentro del campo del tratamiento matemático. Fue Galton quien primero me liberó de mi prejuicio de que las matemáticas sólo podían ser aplicadas a fenómenos naturales dentro de la categoría de la causalidad. Había aquí por primera vez una posibilidad, no quiero decir una certeza, de alcanzar conocimientos -tan válidos como se creía entonces que lo eran los conocimientos físicos- en el campo de las formas vivientes y sobre todo en el campo de la conducta humana. (1)

Las consecuencias de las ideas de Galton no se hicieron esperar. A partir de 1891, Pearson dió conferencias que incluían, Probabilidad, Variación y Correlación. Inició la publicación de una serie de cuatro memorias llamada, *Contribuciones a la Teoría Matemática de la Evolución*, en las cuales aparecerán varios de sus logros más importantes. Resulta muy interesante conocer el desarrollo de sus primeras clases sobre Estadística.

Inicia con lo que es Estadística, Geometría de la Estadística e historia del término. Considera la contribución de la escuela Británica de Aritmética Política, la Alemana de Ciencias Políticas y la Francesa de Probabilidad, para acuñar el término en cuestión. Entre 'el estudio de datos numéricos relacionados con las ciencias físicas, biológicas o sociales' y 'el estudio numérico de hechos sociales', escoge el primero como el más acertado para definir a la estadística. Continúa con la ejemplificación de por lo menos once gráficas con sus variantes para mostrar geoméricamente las estadísticas, señalando ventajas y desventajas en cada caso, así como el uso de curvas cuando las variables son de tipo continuo. La segunda parte del curso la dedica a las leyes del azar asociadas con los fundamentos del conocimiento. Nombra a Laplace, Quetelet,

(1) E.S.Pearson. Ibidem, p.45.

de Morgan, Stanley Jevons, Boole, Venn y Edgeworth y recomienda a sus lecturas. En la introducción enfatiza el valor de las experiencias o registros de hechos pasados como la base para llevar "expectativas de la subjetividad, al plano de la objetividad"; primero se cree en la rutina basada en la experiencia estadística, para luego poder inferir; se tienen diferentes grados de creencia proporcionados por el número de hechos ocurridos registrados, y como extremos habrá esperanza subjetiva y objetiva (donde la esperanza es "el registro de experiencias pasadas alcanzadas por la observación".)

Cuanto más amplio es el conocimiento y la experiencia de un hombre, más íntimamente se aproxima la esperanza subjetiva al valor objetivo.(1)

Muestra la regularidad estadística de los grandes números con lanzamientos de monedas, extracción de bolas y resultados de la ruleta.

El azar, afirma Pearson(2), reside en nuestra ignorancia de una serie de causas finalmente equilibradas y complejas y éste no tiene que ver con una libertad objetiva de la naturaleza sino con la forma subjetiva en que clasificamos una cantidad de hechos similares en la experiencia pasada.

Para deducir las leyes del azar en clase, discute dos caminos el teórico y el empírico. En el primero analiza las probabilidades 'a priori' advirtiendo la necesidad de recurrir a alguna forma de experiencia para mostrar su validez como en el caso de los valores 'igualmente probables', en donde las frecuencias observadas mostrarán de forma burda su probabilidad equitativa. Siguiendo el segundo camino,

(1) Ibidem., 220.

(2) Ibidem.

distingue entre experimentos artificiales (juegos de azar) y fenómenos naturales (alturas de soldados italianos, variaciones de ciertas razones en estadísticas vitales, medidas de órganos de ciertos grupos de animales, etc.), y va deduciendo las leyes del azar. Obtiene deducciones experimentales con lanzamientos de monedas, ruletas de MonteCarlo y loterías Danesas; encuentra una estrecha relación entre lo práctico y lo teórico (comparando las frecuencias relativas de los resultados al lanzar 2,400 veces 10 monedas a la vez con las probabilidades esperadas según el modelo binomial). Observa la distribución de las divergencias de la media y establece: a) que las observaciones se agrupan alrededor de la media y, b) se disponen simétricamente por encima y debajo de ella, surgiendo dos preguntas: ¿Cuál es la ley, según la cual se distribuyen las divergencias de la media?, ¿Cómo son influenciados los resultados al incrementar el número de casos observados?. Para esta última pregunta propone analizar el error medio o divergencia.

Pearson no fué el único cautivado por la medida de variabilidad, alguien más se vió involucrado en los inicios de la correlación junto con él. Nos referimos a W.R. Weldon, quien sería imposible de omitir en este estudio.

Walter Frank Raphael Weldon (1860-1906), biólogo interesado en Zoología, realizaba investigaciones sobre todo en especies marinas, estimulado por el ambiente creado por el darwinismo y su teoría de la Evolución. Entre los problemas que se proponía investigar estaba por ejemplo, el siguiente:

Al establecerse un nuevo conjunto de características de un adulto que conducen a la evolución de una nueva familia,

siempre se ha acompañado de una evolución de un nuevo conjunto de características larvarias que conducirán a la formación de un tipo de larvario peculiar a la familia recientemente establecida; los dos conjuntos de características no tienen todavía conexión entre ellos demostrable. (1)

Este y otros problemas dejaron de serlo o por lo menos encontró el camino para resolverlos cuando Galton propuso el método para cuantificar la correlación entre órganos. Efectivamente, Weldon comprendió que era necesario abandonar el enfoque morfológico para concentrarse en la cuantificación de las variaciones de los miembros de una especie animal resultando que en *Las Variaciones que ocurren ciertos Decapod Crustacea I. Crangon Vulgaris* (1890), Weldon aplicó por primera vez los métodos de Galton a tipos zoológicos además del hombre. En este artículo se muestra que las mediciones hechas sobre varias razas locales de camarones dan una distribución de frecuencias que sigue aproximadamente la ley normal. En otro escrito que le siguió, *Sobre ciertas variaciones correlacionadas en Crangon Vulgaris*, Weldon calculó los primeros coeficientes de correlación para dos órganos o caracteres en el mismo individuo. En opinión de Pearson, estos dos artículos marcan una época en la historia de la ciencia, llamada posteriormente biometría. (2)

Galton hizo algunas modificaciones al primero de los dos escritos y con esto se inició una fructífera amistad entre ambos.

(1) K. Pearson. N.F.R. Weldon 1860-1906 (artículo incluido en: E.S. Pearson, M.G. Kendall, Ibidem. p. 278,

(2) Ibidem., 282.

Se dice que el nombramiento de Weldon como profesor de Zoología en el mismo College de Pearson había resultado crucial para el 'fundador de la estadística' debido a que para ambos catedráticos, la teoría de evolución planteaba un análisis estadístico. En el caso de Weldon 'el nuevo cálculo' de Galton abría posibilidades de darle respuesta a las interrogantes que él ya se venía formulando decidiendo abandonar los viejos métodos para probar y comprobar lo que veía como hipótesis potenciales en las teorías darwinistas. Y tratándose de Pearson, él era quien poseía una sólida preparación matemática, había solicitado la cátedra de Gresham de Geometría en 1890 involucrando la estadística y la probabilidad en el programa que proponía y quien solía conversar con Weldon a la hora del almuerzo, aprovechando, ambos para sugerirse problemas y proponer soluciones. Se puede tomar el siguiente párrafo extraído del prólogo del primer volumen de *Biometrika*, como el planteamiento de los objetivos que se propusieron a raíz de su encuentro en el University College y que constituiría el trabajo simbiótico que desafortunadamente, no pudieron continuar durante mucho tiempo debido a la muerte prematura de Weldon:

Hace pocos años, todos aquellos problemas que dependían para su solución, de un estudio de las diferencias entre los miembros individuales de una raza o especie, fueron olvidados... El punto inicial de la teoría de Darwin de Evolución, es precisamente la existencia de aquellas diferencias... el primer paso en una investigación sobre el posible efecto de un proceso de selección sobre cualquier característica de una raza, debe ser una estimación de la frecuencia con la cual, los individuos exhibiendo cualquier grado de anomalía con respecto a dicha característica, aparecen. Estos y otros problemas involucran la colección de datos estadísticos en gran escala...

En un segundo editorial se abunda más sobre tales

lineamientos:

...el problema de la evolución es un problema estadístico... debemos voltear a las matemáticas de los grandes números, a la teoría de los fenómenos a gran escala, a interpretar sin peligro nuestras observaciones,

...nos podemos preguntar cómo sucedió que el fundador de nuestra teoría moderna de descendencia apeló tan poco a la estadística. La inclinación propia de la mente de C. Darwin le llevó a establecer la teoría de descendencia sin concepciones matemáticas; asimismo la mente de Faraday trabajó en el caso de electromagnetismo. Pero así como cada idea de Faraday permite definirla matemáticamente y demanda análisis matemático..., igualmente cada idea de Darwin -variación, selección natural... parece enseguida ajustarse a definiciones matemáticas y demandar análisis estadístico. Que la concepción estadística misma no fué requerida enteramente en el trabajo de Darwin... que la mente de Darwin no trabajó fácilmente cuestiones matemáticas... es quizá mejor evidenciado en una carta de 1857. Pero que él haya comprendido la importancia del método estadístico para sus investigaciones es evidente no solamente en este pasaje sino en varios otros...* El biólogo, el matemático y el estadístico han tenido hasta el momento campos de trabajo ampliamente diferenciados... El día llegará en que encontremos matemáticos quienes sean biólogos competentes y biólogos quienes sean matemáticos competentes.(1)

Pearson reconce en Weldon un factor importante en el desarrollo de la estadística aplicada a la biología. A partir de 1890 las contribuciones de Pearson consistirán en 35 memorias en Proceedings y Philosophical Transactions of the Royal Society. Es fácil comprender el porqué se le ha llamado el 'creador de la estadística' después de revisar sus aportaciones, tomando en cuenta que antes de él se manejaban solamente la distribución binomial y normal, sobre todo relacionadas con la teoría de los errores; fuera de

* Schrodinger afirmó que Darwin fué el primer científico en estar consciente del papel vital de la estadística. 'Su teoría, dijo, está basada en la Ley de los grandes números', (Ver On the History of Statistical Method in Biology, O.B. Sheynin. Archive for History of Exact Sciences. V. 22, 1980), 362.

[1] O.B. Sheynin Newton and the Classical Theory of Probability Archive for History of Exact Sciences. Vol. 7. No. 3, 1971. p. 240.

esto se tenfan solamente algunas excepciones con Cournot, Quetelet, y desde luego Galton. Con Weldon y Pearson, el desarrollo y aplicación de los métodos estadísticos estuvieron dirigidos principalmente a tratar problemas de evolución y herencia. Respecto a lo primero, se buscaba una teoría sobre la influencia en la correlación y variabilidad de los órganos. Y en lo segundo se trataba de ensayar la bondad de la ley de Herencia Ancestral de Galton. Precisemos algunos de los avances logrados: antes de 1893 la variación se medía con el error medio o gráficamente por medio de la ojiva de Galton; en el primero de las contribuciones matemáticas a la teoría de la evolución, Pearson propone mediarla en términos de la desviación estándar introducida aquí, junto con el símbolo σ , por primera vez.

En este mismo artículo se convierte la curva de los errores en la curva 'normal' y se introduce el método de momentos para ajustar una curva teórica a datos experimentales; a sugerencia de Weldon, también se trata el problema de la descomposición de una distribución de frecuencias en dos componentes normales; finalmente Pearson presenta una tabla de ordenadas de la curva normal. En otro artículo de la serie anterior, *Regresión, Herencia y Panmixia* (1895), se aborda el tema de la correlación, poniendo como antecedentes a Bravais, Galton y Edgeworth, se generaliza a problemas de correlación múltiple y se demuestra que

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}}$$

es el mejor estimador del coeficiente de correlación poblacional (fórmula que conocemos como producto-momento o la r de Pearson). Derivó otras dos fórmulas equivalentes, dió

una rutina simple para el cálculo de r que podía ser seguido por una persona sin mucho entrenamiento matemático, estableció una teoría general de la correlación para tres variables y dió los coeficientes de la ecuación de regresión múltiple en términos de los coeficientes de correlación de orden cero.

Para manejar datos correspondientes a características como color de pelo, ojos y de piel, introduce el método de momentos para calcular correlaciones en *Sobre la correlación de caracteres no medibles cuantitativamente y Sobre la Herencia de caracteres no sujetos a mediciones cuantitativas* en 1900. En este mismo año surge una de las más importantes aportaciones de Pearson a la estadística: la Ji cuadrada.

Pearson dedicó largo tiempo a la derivación de curvas de frecuencias a partir de una sola ecuación diferencial con parámetros obtenidos por el método de momentos. Al contar con varias curvas de frecuencia teóricas se preguntaba si se ajustaban o no y en qué medida, cuando se pretendía aplicarlas a distribuciones empíricas. Por tanto era necesario contar con un criterio o medida de discrepancias entre ambos conjuntos de valores, Quetelet y Weldon en tales circunstancias decidían la bondad de ajuste por simple observación. Sheppard propuso hacer tales comparaciones considerando la diferencia observado-esperado como una variable binomial, sin pensar en que estas variables eran correlacionadas. Sin embargo, Pearson, gracias al concimiento de la teoría de la normal multivariada supo dar como solución a la χ^2 (Ji cuadrada). Para esta función, cuyo valor aumenta en la medida en que el ajuste empeora, calculó sus valores, obtuvo su distribución y entendió el concepto de 'grados de libertad' solamente en un sentido:

sostenía que para k categorías el número de 'grados de libertad' debía ser $k-1$. (1) (Sin considerarlos como el número de observaciones en la muestra menos el número de restricciones o bien, menos el número de parámetros estimados).

En lo que se refiere a la Ley de Herencia Ancestral, Pearson se propuso aplicarla y ensayarla utilizando gran cantidad de datos que provenían de características en los humanos como: estatura, índice encefálico, color de ojos, fertilidad y longevidad. En caballos, se interesó por el color de pelo y fecundidad; en perros, el color de pelo; en mariposas, las marcas en las alas, etc. Pearson consideraba a la Ley de Galton como una ley general, a la cual le faltaban evidencias experimentales para definir algunos componentes de ella; por ejemplo, había que ver hasta dónde, los coeficientes de correlación tomaban valores fijos o constantes para distintas características y razas. Weldon y él pensaban que las características de la descendencia, ya sea de plantas o animales diferentes entre sí, caían en tres casos:

- 1) las características de los hijos estaban en un nivel intermedio entre las características del padre y de la madre, o
- 2) las características de los descendientes se podían separar por zonas, algunas de las cuales se identificaban con influencia de uno de los padres y las demás, herencia del otro, o
- 3) las características del hijo se debían completamente a uno de sus progenitores.

Para el primer caso, llamado también herencia mezclada, se habían concluido que para ciertas características:

a) estas se distribuyen normalmente, b) se tenía una correlación entre las diferentes características y c) que una correlación similar se observaba también con características correspondientes entre hijos, padres, abuelos, etc., d) además, cuando se observaban grupos apareados al azar, se daba un cierto grado de constancia en los valores de los coeficientes de correlación, e) asimismo, parecía existir evidencias sobre mortalidad selectiva en ciertos casos sobre las características heredadas.

Para los dos casos restantes se reconocía la falta de investigación y la confusión imperante, sin embargo cabe mencionar el punto de vista de lo que sería la 'escuela biométrica' sobre herencia. El enfoque de los grandes números:

En primer lugar debemos liberar nuestra mente en el presente estado de nuestro conocimiento del mecanismo de herencia y la reproducción, de toda esperanza de alcanzar una relación matemática que exprese el grado de correlación entre progenitores individuales y descendientes individuales...

Debemos ir de la herencia en general a la herencia en categorías cada vez más limitadas, en lugar de intentar obtener reglas generales de la observación de casos particulares. (1)

En el transcurso del año 1900 se redescubren los trabajos de Mendel, dieciséis años después de su muerte. Tres biólogos, Hugo de Vries (Holanda), Carl Correns (Alemania) y

(1) E.S. Pearson, Ibidem., p. 67.

Erich Tschermak (Austria) encuentran cada uno de forma independiente los mismos resultados de Mendel(1) publicados tiempo atrás. El reconocimiento al biólogo checoslovaco se hace patente con la aceptación de su teoría por parte de un buen número de científicos renombrados, entre quienes se encontrarán los principales críticos al enfoque biométrico de Galton y Pearson. De aquí en adelante se inició una confrontación a lo largo de varios años entre los biometras y los mendelistas, sobre el tema de la herencia surgiendo una serie de diferencias, algunas insalvables. Tratemus de hacer un balance de los inconvenientes de ambos enfoques señalados por los contrarios.

Quienes se opusieron a las leyes mendelianas se apoyaban en la teoría de la herencia por mezcla, que era la que prevalecía en su tiempo y en la cual estaba basada la Ley de Herencia ancestral; para un biólogo mendeliano esta ley le resultaba innecesariamente complicada, las matemáticas en la biología aún no eran totalmente aceptadas (la revista *Biometrika* surgió por la negativa por parte de la Royal Society de publicar trabajos que no mostraran claramente la separación entre lo biológico y lo matemático). Los biometras encontraban inaceptable que se descartasen a los híbridos de la primera generación en los escritos de Mendel, por carecer de interés refiriéndose exclusivamente a los primeros descendientes de un cruzamiento que se parecían solamente a uno de sus padres; por otro lado, el que las características parentales se separaran íntegramente después de estar asociados en los híbridos de la primera

(1) Marcel Blanc Gregor Mendel: La Leyenda del Genio Desconocido, Mundo Científico No. 34, (1984) 277.

generación negaba la mezcla de caracteres aceptada por muchos. (1) En el caso de la fórmula de Galton se incorporaban las contribuciones genéticas de los antepasados (padres, abuelos, bisabuelos, etc.,) a los descendientes, lo cual explicaba según los biometras, el atavismo* observado en muchos casos. Mendel, sin embargo, tal como propuso sus leyes originalmente, éstas involucraban solamente a los padres del descendiente. Otra dificultad en la aplicación de las leyes mendelianas, o mejor dicho, un punto débil que señalaban sus oponentes, residía en la transmisión de caracteres de tipo continuo; por ejemplo, si se cruzan dos plantas una de flores grandes y otra de flores pequeñas, los híbridos deben ser de talla mediana. Mendel manejó solamente características discontinuas; en este sentido la Ley de Herencia Ancestral concordaba muy bien con la transmisión de caracteres de tipo continuo, puesto que involucraba cantidades.

Como resultado de las críticas hechas a sus planteamientos, Pearson realizó revisiones y aclaraciones respecto a la Ley.

En todos los casos como los de hombre, caballo y perro, en que padres de caracteres idénticos no producen idéntica descendencia, la teoría de la estadística no muestra que puede alcanzarse una predicción más aproximada, tomando en cuenta muchos antecesores en lugar de unos cuantos.

La ley de Herencia Ancestral en su forma más general no es de ninguna manera una hipótesis biológica, es simplemente un teorema fundamental en la teoría estadística de la correlación múltiple aplicada a un tipo particular de

(1) Darwin escribió: 'por regla general al cruzar dos razas, los caracteres propios de cada una de ellas tienden a confundirse de forma íntima'.

* Atavismo: características físicas heredadas de parientes más lejanos que los padres o abuelos.

estadísticas. Si las estadísticas mismas de la herencia son buenas, los resultados deducidos de este teorema serán verdaderos cualesquiera sea la teoría biológica de la herencia que se proponga... (1)

Con el propósito de mostrar que los enfoques en pugna podían ser complementarios puesto que en uno se explicaba lo que con el otro no se lograba, Pearson señalaba que el punto de vista biométrico sobre la herencia no negaba teoría fisiológica alguna, sino por el contrario, sería para refutarla o afirmarla.

Un trabajo que no logró terminarse fue aquel en que Weldon se proponía desarrollar, con el apoyo matemático de Pearson, una teoría que mostrase la necesidad de un modelo que incluyese al mendelismo en un extremo de la escala y la herencia mezclada en el otro. La conciliación entre las dos aproximaciones tuvo que esperar hasta que otro brillante estadístico y genetista pudiera demostrar la consistencia de los principios mendelistas con las correlaciones observadas entre padres e hijos. (2) Nos referimos a Ronald A. Fisher.

Las controversias dieron como resultado algunas contribuciones por parte de Pearson, junto con la ayuda de Weldon y sus experimentos de procreación:

- comprensión de la teoría de la correlación múltiple;
- el ensayo de la bondad de ajuste de la Ley de la Herencia Ancestral, como fórmula descriptiva por medio de una extensa colección y análisis de datos sobre herencia en el

(1) E.S. Pearson. Ibidem., p. 75.

(2) W. Kruskal, J.M. Tamur. Ibidem., p. 356.

hombre y los animales;

- la aproximación a las leyes de la herencia y la evolución por medio del estudio de los efectos de conjunto en poblaciones bajo condiciones naturales.
- introducción de mediciones precisas en el estudio de poblaciones raciales (Antropometría).

Aún cuando todavía no era posible dar conclusiones respecto a los estudios de herencia de caracteres físicos, ya que se pensaba en la adquisición de mayor información, Pearson extendió sus investigaciones a la herencia de caracteres mentales. Recordemos que Galton afirmaba que la inteligencia era hereditaria, pero sin embargo, hacía falta un soporte experimental a tal afirmación. A profesores de escuelas londinenses se les pidió que efectuaran medidas y estimaciones físicas y mentales de sus alumnos. Como características físicas se consideraron la salud, color de ojos, color y rizado de pelo, índice encefálico, longitud, ancho y altura de la cabeza y potencial atlético. Y como características mentales se incluían vivacidad, seguridad, introspección, popularidad, escrupulosidad, temperamento, habilidad y carácter en la letra. La medida que se utilizó para cuantificar sus variables fue el coeficiente de correlación. En el caso de caracteres no mensurables cuantitativamente, obtuvo los coeficientes a partir de tablas que desarrolló; por ejemplo, el método tetrárquico en tablas 2x2 para el caso de tuberculosis en el padre y el niño. En donde el cálculo de r implicaba que la variable en este caso, la diátesis, (predisposición a la tuberculosis) fuera de tipo continuo y con distribución normal pero no se hacía hincapié en estas condiciones, sino únicamente en el concepto de la medida de r de asociación entre variables.

P A D R E				
	Tuberculoso	No Tuberculoso	Total	
N	Tuberculoso	63	172	235
I				
N	No Tuberculoso	309	4,862	5,171
O	T o t a l	372	5,034	5,406

En una memoria aparecida en *Biometrika* se compararon dos tablas de correlaciones fraternas para nueve caracteres físicos y ocho mentales; para cada caso se tomaron tres coeficientes hermano-hermano, hermano-hermana y hermana-hermana, observándose como coeficientes medios, en caracteres físicos:

hermano-hermano 0.54; hermana-hermana 0.51; hermano-hermana 0.51; para caracteres mentales:

hermano-hermano 0.52; hermana-hermana 0.51; hermano-hermana 0.52. Tales cantidades le llevaron a las siguientes conclusiones:

Estamos obligados, creo que literalmente obligados, a aceptar que los caracteres físicos y psíquicos en el hombre son heredados, en líneas generales, en la misma forma y con la misma intensidad...

Pero no hemos llegado al fin de nuestras conclusiones. Suponiendo nuestra distribución normal para los caracteres psíquicos, hemos encontrado resultados no sólo consistentes en sí mismos, regresión lineal, por ejemplo, como en el caso de la herencia de la inteligencia, sino que hemos encontrado el mismo grado de semejanza entre caracteres físicos y psíquicos. Esta igualdad implica algo más.

Implica una herencia igual de los padres. El grado de semejanza entre niños y padres para los caracteres físicos en el hombre puede ser aplicado al grado de semejanza entre

niños y padres para caracteres psíquicos. Heredamos los temperamentos de nuestros padres tal como heredamos su estatura, antebrazo y dimensiones. (1)

Omitiendo otros aportes no menos importantes, llegamos a las primeras dos publicaciones de una *Serie biometrika* que financió la Draper Company en 1904-1905 cuyos contenidos se pueden considerar como los avances más importantes, después de la χ^2 .

En la primera memoria, *Sobre la Teoría de la Contingencia y su Relación con la Asociación y Correlación Normal*, propone a ψ^2 (contingencia media cuadrática), como una medida de asociación entre dos variables, en caso de que los caracteres no estén clasificados en escalas cuantitativas, es decir, atributos cualitativos (ojos azules, negros, maculino, femenino, etc.) y demuestra: a) cómo se puede usar $N \psi^2 = \chi^2$ para ensayar la independencia de los dos caracteres, y b) cómo, bajo ciertas condiciones $C = \sqrt{\psi^2 / (1 + \psi^2)}$ tiende al valor del coeficiente de correlación r .

La segunda memoria titulada *Sobre la Teoría General de Correlación Asimétrica y Regresión no lineal*, incluía: a) las propiedades de la razón η de correlación (se emplea para determinar la estrechez de la relación entre dos variables cuando la regresión no es lineal) y la diferencia entre η y r en la medida de la desviación de la linealidad de regresión, b) el error probable de η , el cual se demostró posteriormente, ser inexacto, c) la clasificación de distintos tipos de variabilidad y de formas de líneas de regresión que surgen cuando dos variables no pueden ser representadas por la normalidad bivariada, d) fórmulas para

(1) E.S.Pearson, *Ibidem.*, p. 80.

líneas de regresión parabólicas, cúbicas y cuárticas; poniendo además expresiones para sus constantes en términos de momentos y momentos-producto de las observaciones, e) y finalmente, las condiciones que debían satisfacer las ecuaciones de regresión lineal, parabólica y cúbica para ser adecuadas. Quizá convendría, en este momento, recordar lo que decía Pearson cuando comentaba *Herencia Natural* de Galton: " la correlación traía a la psicología, antropología, medicina y sociología, dentro del campo del tratamiento matemático"; solamente es cuestión de revisar el contenido de la bibliografía de estadística para cada una de las disciplinas mencionadas para confirmar sus palabras.

Siguiendo con la *Serie Biométrica* encontramos otro problema de interés para Pearson, el Albinismo. Teniendo como antecedentes a Darbishir y Suchester en Oxford, junto con Weldon y sus investigaciones de cruces experimentales de ratas para determinar hasta dónde podía considerarse el albinismo como carácter hereditario, Pearson recopiló gran número de datos. algunos ya publicados y otros obtenidos con ayuda de médicos, correspondientes a seres humanos y perros; realizó inclusive investigaciones históricas, reuniendo datos de sociedades y culturas remotas (Africa, India, Sudamérica). Pero pronto se dió cuenta de la complejidad del tema al registrar la característica de varias generaciones de perros que se criaron para tal propósito, empezando por el no poder clasificar a los individuos como albinos y no albinos, y por el considerar la existencia de diferentes grados de intensidad de dicha característica que impedía su análisis. Finalmente se limitó a advertir lo insuficiente del conocimiento de su tiempo para problemas de tal naturaleza, así como a refutar o hacer patente las limitaciones del enfoque mendeliano para el caso en cuestión. Haciendo una analogía entre el

albinismo y los estados patológicos en el hombre, puntualizaba los peligros y lo aventurado que sería la aplicación de las leyes de Mendel a problemas de tipo social debido a las posibles conclusiones derivadas de ellas, en el sentido de permitir, por ejemplo, matrimonios con antecedentes anormales. (1) En 1904, Galton había instalado una Oficina de Registro Eugénésico cuyo objetivo era el de promover el estudio científico de la Eugenesia, tiempo después bajo la dirección de Pearson funcionó en conexión con el Laboratorio Biomédico y a partir de 1911 se unieron formando el Departamento de Estadística Aplicada del University College, fungiendo Pearson como su primer profesor. Desde 1907 el laboratorio empezó a publicar artículos muy controvertidos: tuberculosis pulmonar, alcoholismo, deficiencia mental y locura. Un ejemplo más del tratamiento estadístico dirigido a nuevas investigaciones por parte de Pearson y su equipo de laboratorio, es el que se refiere a cuestionar medidas de control de la tuberculosis adoptadas después del descubrimiento del bacilo de Koch.

Al analizar datos provenientes de hospitales y de historiales familiares de criminales se concluyó que la predisposición a la tuberculosis (diátesis) era hereditaria en el hombre en la misma proporción que otros caracteres físicos. Y el hecho de que la correlación encontrada para la diátesis entre padres e hijos era mucho más elevada que entre marido

(1) La Oficina Americana de Registros Eugénésicos declaraba en 1912 que el defecto mental era casi seguramente un carácter mendeliano recesivo y recomendaba que 'la debilidad en alguna cualidad debería casarse con la fuerza en la misma característica' y que 'la fuerza podía casarse con la debilidad'.

(E.G. Pearson. Ibidem. p. 106.)

y mujer, favorecía al factor hereditario puesto que de otra manera, si se tratase de acuerdo con la teoría de la infección pura debía esperarse que hubiese tanta probabilidad de que marido y mujer se infectasen como la correspondiente entre padres e hijos. Por lo anterior era de esperarse que librasen debates y discusiones en torno al problema con los médicos quienes defendían el tratamiento clínico para erradicar la enfermedad. Más aún, estudios actuariales realizados no pudieron encontrar evidencias claras de que las tasas de mortalidad comparadas entre pacientes tratados en sanatorios y fuera de ellos resultarán más bajas en el primer grupo.

Otra polémica tuvo lugar cuando Pearson trató el problema del alcoholismo. ¿Se tiene evidencias de que el alcoholismo de los padres influye sobre la mentalidad y físico de los hijos? No fué la conclusión a la que llegaron M. Elderton y K. Pearson después de buscar correlaciones entre diferentes anomalías de niños y sus antecedentes paternos; advirtiéndole que su investigación fué limitada y que las conclusiones podían cambiar el análisis del fenómeno bajo diferentes circunstancias. Otras dos memorias sobre el mismo tema fueron editadas por el Laboratorio Eugénico sobre casos de alcoholismo extremo en adultos. En ellas se afirma que el alcoholismo es consecuencia de defectos mentales y no al contrario entonces, puesto que la locura se consideraba hereditaria 'habría que separar a los niños mentalmente defectuosos como un primer paso en el tratamiento del alcoholismo y otras enfermedades'. Para los integrantes del laboratorio se tenían evidencias de que la tendencia a defectos patológicos eran heredados en la misma manera que sucedía para los caracteres físicos. Hoy en día sigue siendo el alcoholismo un tema de permanente investigación. Nuevas

investigaciones en el campo han demostrado que en gran porcentaje de los casos esta enfermedad puede ser ocasionada por factores biológicos y genéticos que pueden ser detectados, antes de que los daños causados por la enfermedad sean irreversibles. (1)

La producción de trabajos de Pearson, varios de ellos realizados junto con investigadores del laboratorio biométrico o de eugenesia, continuaron casi ininterrumpidamente, el analizar a cada uno de ellos está fuera de nuestro alcance, pero podemos referirnos a muchos de ellos en términos de su clasificación más general. Una buena parte de sus estudios estuvieron motivados por sus ideas eugenésicas. Pensaba que:

Ningún plan de educación por amplio que fuere convertiría en la escala de inteligencia, la debilidad hereditaria en vigor hereditario. El único remedio, si es que hay alguno, es alterar la fertilidad relativa de los buenos y malos linajes en la comunidad. (2)

La creencia en el factor hereditario como determinante en la formación del hombre integral, más que el factor ambiental (Naturaleza sobre Crianza) dió como resultado investigaciones con temas afines a problemas sociales principalmente de tipo racial.

Para la difusión de tales trabajos (algunos ya publicados antes), así como para proclamar a la Eugenesia como ciencia, funda la revista 'Los Anales de Eugenesia'. En ella se

(1) Desmistificación del Alcohólistmo: Enfermedad causada por factores biológicos y genéticos: Excélsior 24 de noviembre de 1985.

afirma que el fundamento de dicha ciencia es la Probabilidad. Para esto se realizan estudios que incluyen mediciones de correlación entre caracteres físicos y genéticos. Consecuentemente se tiene razonamientos como el siguiente:

Difficilmente se puede hablar de selección en el hombre, excepto por la selección basada en caracteres somáticos. Tendría sentido hablar de evolución siempre y cuando estuvieran correlacionados los caracteres somáticos y genéticos. Puesto que esto último es indudable, entonces la Eugenesia toma gran importancia para la sociedad.

De las últimas publicaciones de Pearson tenemos una colaboración de *Biometrika* No. XVIII, 1936 'Tablas de función Beta Incompleta' aparecida en 1934, también 'Tablas para Estadísticos y Biometras', 'Tratados para calculistas' y dejó incompleta otra tabla de probabilidades del coeficiente de correlación.

Después de la renuncia a su cátedra de la Universidad en 1933, se divide el Laboratorio de Estadística Aplicada y se crean tres cátedras: Eugenesia, Estadística y Biometría. Karl Pearson fallece el 27 de abril de 1936 dejando lo que podríamos llamar la 'Escuela Inglesa de Estadística' a R.A. Fisher, su hijo E.S. Pearson, J. Newyman, M. Kendall, G. Yule, W.S. Gosset (student) y otros.

C O N C L U S I O N E S

La creación de técnicas específicas para resolver problemas específicos de su tiempo, es el común denominador de la labor de F. Galton y K. Pearson. En este trabajo se vieron algunos de los principales problemas que trataron de resolver. Se expuso cómo intentaron darles solución y qué se logró con sus investigaciones. Se analizó la peculiaridad de los estudios y el deseo de sujetarlos al tratamiento estadístico. Desde luego, hubo quien, con ingenio y creatividad, les precedió. Hombres con el gusto y la habilidad para registrar gran cantidad de datos, acumularon información de contingencias de la vida humana, descubriendo regularidades en esas masas de información, sin sospechar en principio, la similitud con otros fenómenos.

Enseguida hubo quien, gracias a esa regularidades o patrones de conducta, pudieron entrever la factibilidad de explicar tales descubrimientos en términos de modelos probabilísticos y más tarde aplicarlos en temas biológicos. La aparición de la teoría de la Evolución tuvo grandes implicaciones, una de ellas fué precisamente que Galton y Pearson se enfrascaran en la búsqueda de su demostración matemática. El primero de ellos por motivos que podrían ir desde una ideología propia de su clase conservadora, hasta el genuino interés científico decide trabajar en un punto central de la evolución: la herencia. El segundo, cuya formación matemática le permitió ir más allá del coeficiente de regresión y correlación propuesto por Galton, continuó con el análisis escrupuloso y analítico de la información.

Se vio que en el momento en que no sabían el porqué de las cosas optaron por tabular los datos. Por ejemplo, Galton

tabuló los resultados de sus experimentos con chicharos; surgiendo de ahí sus ideas sobre correlación. Asimismo, por querer comprender el mecanismo de la herencia y pretender demostrar la Selección Natural, registraron los cambios para hallar indicios de las leyes involucradas utilizando la ley de los grandes números.

No desecharon el estudio de la 'unidad' para conocer el proceso fisiológico de la herencia, pero discutieron la utilidad de conocer el comportamiento a nivel poblacional de la variable de estudio, para complementar o inclusive rechazar posibles teorías erróneas de él.

Conceptos como raza, linaje o élite de inteligencia superior han sido desechados, sin embargo no así los métodos que se utilizaron para ponderarlos cuando se consideraban vigentes. Actualmente diversas disciplinas recurren a los métodos estadísticos, en algunos casos, como la única opción para el análisis de sus materias de estudio, obteniendo resultados positivos. Apoyándonos en las herramientas ideadas por Galton y Pearson tenemos la alternativa de describir situaciones reales complejas con el fin de poder dilucidar las causas de los efectos; muchas veces sin éxito pero en ocasiones se llega a la antesala de dicho nivel de conocimiento. Con el concepto de correlación podemos orientar estudios, los cuales al ser mas profundos, nos es posible conocer la interdependencia entre las variables.

Pudimos apreciar el papel fundamental que tuvo el modelo normal en las primeras aplicaciones del razonamiento estadístico. La aplicación novedosa de dicho modelo en cuestiones diferentes de la teoría de los errores y la astronomía, trae consigo una serie de implicaciones que la colocan como primordial en la Estadística Aplicada.

Se vió como Quetelet había sugerido, su amplia utilización en campos como la psicología, medicina, antropometría y sociología. Y Galton, debido a la influencia del astrónomo belga, la propone para mostrar la distribución de la 'inteligencia'. Pearson, a su vez, generaliza de la distribución bivariada a normal multivariada, estableciendo la teoría de la correlación múltiple. Esto seguramente gracias al conocimiento de la teoría de las probabilidades lograda por Gauss, Laplace y otros.

Por la gran trascendencia de las aportaciones al conocimiento, no resulta tan fácil poner en su exacta dimensión a sus creadores. La falta de fuentes confiables así como cierta parcialidad en algunas apreciaciones también son motivo de ponderaciones no muy bien logradas. A pesar de lo anterior, se ha puesto empeño en mostrar cómo fueron los inicios de la Estadística en lo que se refiere a lo que podemos llamar la 'Escuela Inglesa' y en particular a los dos autores citados.

La obra de Galton y Pearson es solamente un pequeño eslabón en la cadena del desarrollo del conocimiento científico. Darwin toma de Malthus ideas para su teoría. Galton resulta excitado por la obra de Darwin. Pearson encuentra en Galton motivos y apoyo para realizar investigaciones, y así sucesivamente.