

생물측정학-멘델주의 논쟁에 대한 통계학사적 고찰†

조재근¹⁾

요약

생물측정학-멘델주의 논쟁은 다윈이 주장한 생물 진화의 연속성과 자연선택을 두고 1890년대 중반부터 약 10여년간 치열하게 진행되었던 논쟁이다. 본 연구에서는 생물학이 아니라 통계학에 중점을 두고 논쟁의 진행과정을 살펴보았다. 특히 수학자이던 피어슨이 생물학 자료에 대해 통계학을 연구하게 된 배경, *The Grammar of Science*에 나타나는 그의 과학철학과 논쟁에서 나타나는 피어슨의 입장, 그리고 그의 입장과 1830년대 이후 영국 통계학 전통 사이의 관계도 살펴보았다. 결과적으로 이 논쟁을 계기로 피어슨의 많은 통계학 연구가 나오게 되었고 처음으로 연구자들이 결집하게 되었으며 새로운 학술지 *Biometrika*가 창간되었음을 알 수 있었다. 또한 이 논쟁은 우생학과 더불어 수학적인 통계학이 대학에 자리잡는 촉매제로 작용하였다.

주요용어: 상관관계와 인과관계; 멘델주의; 생물측정학; 진화와 유전학.

1. 서론

잘 알려져 있듯이 통계학의 역사에서 19세기 말부터 20세기 초에 이르는 영국 통계학의 역사는 20세기에 번창하게 될 현대 통계학이 탄생하는 역사이기 때문에 통계학의 역사를 연구하는 많은 학자들이 여러 가지 방향에서 연구해왔다 (스티글러, 2005; Porter, 1986; Hald, 1998). 그런데 당시는 아직 통계학이라는 학문이 대학을 비롯한 학문 세계에서 하나의 독립적인 분야로 자리잡지 못하고 있던 때였기 때문에 통계학은 다른 분야를 연구하는 과정에서 발달하게 되었고, 19세기말 영국에서 그러한 역할을 했던 분야는 유전학(genetics)이었다. 그런데 사실상 유전학 역시 수학이나 물리학 등과는 달리 당시에 새로 태어난 학문이었다. 즉 이 시기에 비로소 유전학의 기본적인 용어와 이론들, 연구 방법('genetics'라는 이름 자체가 이 논쟁의 주역 중 한 사람인 베이트슨에 의해 20세기에 들어와서야 만들어진 용어이다) 등이 등장했던 것이다.

그러한 과정 속에서 1900년을 전후한 십여년 동안 유전학과 진화이론을 둘러싼 논쟁이 영국을 중심으로 벌어진 적이 있는데 흔히 '생물측정학-멘델주의 논쟁(biometry-Mendelism controversy)'이라 불리는 논쟁이 바로 그것이다. 그 논쟁의 한쪽 편에는 피어슨(Karl Pearson, 1857-1936)과 웰던(W. F. R. Weldon, 1860-1906)을 중심으로 한 생

† 이 논문은 2005년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임 (KRF-2005-041-C00106).

1) (608-736) 부산시 남구 대연동 110, 경성대학교 정보통계학과, 교수. E-mail: jkjo@ks.ac.kr

물측정학자들(biometricians)이 있었고 그 반대편에는 베이트슨(William Bateson, 1861-1926)을 주축으로 한 멘델주의 유전학자들(Mendelian geneticists)이 있었다 (이들을 중심으로 논쟁에 관련되었던 여러 학자들의 계보를 보려면 Kim (1994) 37쪽의 그림 참조). 세 사람의 주역들 중 베이트슨과 웰턴은 케임브리지 대학에서 생물학을 함께 공부한 생물학자들이었다. 그런데 생물학과는 관계가 없는 응용수학 교수 피어슨이 웰턴의 동료로 논쟁에 가세하면서 통계학이 논쟁의 중요한 요소가 되었다.

처음 진화의 연속성이라는 문제를 두고 웰턴과 베이트슨이 대립하기 시작한 것은 1890년대 초부터였다. 그러던 중 오랫동안 주목받지 못했던 멘델(Gregor Mendel, 1822-1884)의 연구가 새롭게 발견되면서 논쟁은 새로운 단계를 맞게 된다. 잡종 교배에 대한 멘델의 논문 Mendel (1866)은 발표 이후 오랫동안 무시당하다가 1900년이 되어서야 생물학자들의 주목을 끌게 되었다. 수도원장이었던 멘델이 자신의 수도원 정원에서 재배한 완두콩을 가지고 어떤 형질의 우성 또는 열성을 나타내는 인자의 존재와 그 발현 비를 밝힌 논문을 보고 멘델 이론이 진화의 불연속성을 뒷받침한다고 생각한 베이트슨은 멘델주의를 적극 옹호하게 되었다. 그 결과 베이트슨으로 대표되는 생물학자들은 나중에 ‘멘델주의자(Mendelian)’라고 불리게 된다. 반면 생물측정학자들은 멘델의 이론이란 자신들이 주장하는 이론체계의 특수한 사례에 지나지 않는다고 주장했다. 따라서 원래 진화의 연속성 문제로 시작된 논쟁은 멘델주의에 대한 입장 차이가 더해지면서 더욱 격화되었다.

그들은 논문과 책을 통해 논전을 벌였을 뿐 아니라 공개석상에서도 치열하게 논쟁을 벌였다. 가령 1904년 영국학술진흥협회(British Association for the Advancement of Science)에서 “멘델주의가 주장하는 진화 과정은 거추장스럽고 입증할 수 없는 가설일 뿐이다” (웰턴); “멘델주의는 이전까지 형편없이 부적절하고 모순되는 수준에 머물던 유전이라는 사실을 조화롭게 설명함으로써 유전학을 엄청나게 발전시켰다” (베이트슨); “멘델주의자들은 숫자들을 제시하면서 그들이 보여주려하는 이론과 그 숫자들이 어떻게 부합하는지 증명할 생각은 하지 않는다” (피어슨) 등의 설전을 벌이다가, 마침내는 베이트슨이 갓 창간된 *Biometrika*를 손에 들고는 “생물측정학과가 얼마나 어리석은지 보여주는 명백한 증거가 여기 있다”고 면전에서 막말에 가까운 비난을 퍼붓기에 이르렀던 것이다 (MacKenzie, 1981, p.246).

당시로서는 상당히 긴 기간동안 치열하게 전개된 그 논쟁은 1918년 이후 피셔(R. A. Fisher)를 필두로 한 그 다음 세대 연구자들이 집단 유전학(population genetics)이라는 새로운 분야를 크게 발달시켜 양쪽의 주장을 결합하면서 마무리되었다. 나중에 피어슨이 사망한 직후 그의 아들 이곤 피어슨은 이 논쟁을 일컬어 “생물측정학과와 멘델주의자들 사이에는 근본적으로 상반된 것이 아무 것도 없었다. 같은 목표를 추구하였지만 서로 다른 길로 접근했는데 그 두 길은 서로 대립된 것이라기보다는 사실 서로 보완적인 것들이었다.” (Pearson, 1936, p.227)고 평가하였다. 또 미국의 저명한 유전학자는 유전학의 역사에서 “이 논쟁으로 인해 유전학의 상당 부분에서 통계학이라는 강력한 연구방법이 여러 해 동안 쓰이지 못하게 되었다.” (Sturtevant, 1965, p.58)고 아쉬워하기도 하였다. 이로부터 알 수 있듯 후세 사람들은 이 논쟁을 많은 부작용을 낳은 성과 없는 논쟁으로 평가하였다. 그럼에도 불구하고 그 논쟁은 유전학의 역사에서 중요한 논쟁 (Olby, 1966;

Provine, 1971; Bowler, 2003)으로서 과학사 연구자와 과학철학자들도 주목한 논쟁이었고 (Magnello, 1998; Magnello, 2005; Morrison, 2002), 의학의 역사를 연구하는 학자들 역시 연구주제로 삼았던 논쟁이었다 (Rushton, 2000). 또한 양쪽 연구자들의 정치, 사회적 입장과 논쟁의 진행과정은 과학사회학을 연구하는 연구자들에게도 중요한 소재가 되었다 (MacKenzie, 1981; Kim, 1994).

이 논쟁은 그 이름에서도 알 수 있듯이 통계학 전문가 또는 통계학적 방법을 옹호하는 생물학 전문가들이 중요한 역할을 했던 논쟁이었다. 사실 피어슨이 통계학에 대해 처음으로 논문을 발표한 것은 1893년이었는데 바로 그 시기에 피어슨을 통계학자로 만든 것은 이미 베이트슨과 대립중이었던 웰턴과의 만남이었다. 또한 그 이후 피어슨의 적지 않은 연구들이 바로 이 논쟁을 계기로 나온 것이었으므로 이 논쟁은 통계학의 역사에서 짚고 넘어갈 필요가 있는 논쟁이다. 당시의 역사를 연구한 학자들에 따르면 1900년에서 1914년 사이에 영국에서 통계학 이론을 연구하는 사람의 수는 26명이었고, 그들 중 어떤 식으로든 피어슨의 생물측정학과로 분류될 수 있는 사람은 모두 12명, 즉 거의 절반에 달했다고 한다 (MacKenzie, 1981, p.175). 것처럼 이 논쟁에는 20세기 영국 통계학의 선구자들이 거의 모두 관련되었음에도 불구하고 원래 생물학 문제를 둘러싸고 진행된 것이다보니 대부분의 통계학사 연구자들이 이 논쟁에 집중적인 관심을 기울이지 않았다.

따라서 이 연구에서는 통계학을 중심에 두고 이 논쟁을 살펴보려 한다. 그러므로 논쟁의 주역이었던 세 인물 가운데서 우리는 당연히 피어슨의 생각과 활동을 가장 주목할 것이다. 우리는 먼저 세 사람의 주역을 중심으로 논쟁의 진행과정을 개괄적으로 살펴볼 것이다. 그런 다음 피어슨과 베이트슨 사이의 논쟁, 그리고 피어슨과 피셔의 차이에 대해서도 살펴볼 것이다. 그 이후 우리는 수학자인 피어슨이 통계학을 무기 삼아 그의 전공 분야도 아닌 생물학 논쟁에 적극 참여한 배경을 *The Grammar of Science* (Pearson, 1892)에 나타나는 그의 과학철학적 입장을 통해 살펴볼 것이다. 즉 “과학은 설명이 아닌 묘사(science is not explanation, but description)”라는 주장으로 대표되는 그의 과학철학적 입장이 어떻게 해서 역시 그에게 낯선 분야였던 생물학을 통해 통계학 연구로 곧바로 이어졌는지 알아볼 것이다.

그런데 19세기 말까지만 하더라도 통계학이 수학적 학문이 아니었음을 생각해볼 때 그의 과학철학과 통계학 연구를 이어주는 연결고리를 찾기는 쉽지 않다. 우리는 그것을 찾기 위해 1830년대 영국 통계학의 역사로 거슬러 올라갈 것이다. 젊은 네이먼을 놀라게 했던 피어슨의 과학철학은 1890년대에 새로이 나타난 것이었고 (Reid, 1982) 그의 수학적 통계학 연구 역시 분명 새로운 역사를 여는 것이었다. 하지만 우리가 1830년대 영국에서 일어났던 통계 운동에 대해서 살펴보면 당시에 이미 통계학이 과학이 되기 위해서는 “원인과 결과를 찾아서 설명하는 대신 자연과 사회의 현상을 숫자를 써서 나타내야 한다”고 영국 과학계가 규정했었다는 사실을 알게 된다. 결국 피어슨이 ‘설명이 아닌 묘사하는 과학’으로 통계학을 선택한 것은 바로 그러한 전통의 영향을 받은 때문이었을 것이다. 즉 그가 수학적 통계학이라는 새로운 방향으로 나아간 것은 통계의 역할에 대해 반 세기 이상 지속된 영국적인 관념이 그의 과학철학과 결합된 결과라는 것이다. 이와 같이 피어슨의 과학철학과 그의 통계학 연구 사이의 관계를 살펴보고 난 뒤 우리는 이 논쟁의 쟁점을 통계학을 중

심으로 다시 생각하게 된다. 그 쟁점이란 “통계학적인 방법의 특징은 무엇이며 통계학적 방법으로 과학 연구에서 무엇을 할 수 있는가?”라는 질문이다. 어쩌면 피어슨은 논쟁 속에서 이 질문에 대한 답을 찾지 못한 것인지도 모른다. 그럼에도 불구하고 이 질문은 지금까지도 계속 논의되어야 할 질문이다. 따라서 통계학의 역사 속에서 이 논쟁을 살펴보는 작업은 백여년 전 현대 통계학이 출현하게 된 과정을 이해하는데 도움이 될 뿐 아니라 오늘날 통계학의 정체성을 살피는 데에도 도움이 될 것이다.

2. 생물측정학-멘델주의 논쟁과 통계학

2.1. 논쟁의 배경과 진행과정

생물측정학과와 멘델주의자들 사이에 벌어진 논쟁에서 문제가 되었던 것은, 간단히 말해서 다윈이 주장한 진화가 어떻게 이루어지는가라는 문제였다. 진화의 과정에 대한 다윈의 주장은 “Nature Non Facit Saltum(자연은 도약하지 않는다)”는 유명한 말로 요약할 수 있다. 그는 어떤 생물 종의 작은 변이(variation)들이 점진적이며 연속적인(gradual and continuous) 과정을 거쳐 축적되면서 하나의 종이 점차 새로운 종들로 분화된다고 생각했고 그 과정에 자연선택(natural selection)이 작용한다고 주장했다. 하지만 다윈은 자신의 주장을 뒷받침할 근거를 충분히 제시하지 못했기 때문에 <종의 기원>이 나오고 얼마 지나지 않아 그의 주장을 반박하려는, 혹은 뒷받침하려는 여러 연구와 학설들이 등장하게 되었다.

즉 진화의 연속성에 대한 논쟁이란 진화가 점진적인 과정인가 아니면 도약의 과정인가를 둘러싼 논쟁이었다. 여기서 골턴으로부터 시작하여 논쟁을 시간순으로 간략하게 살펴보자. 다윈의 <종의 기원>이 출판된 것은 1859년이고 골턴이 Natural Inheritance를 출판한 시기는 1889년이었다. 그 책에서 골턴은 다윈의 이론을 인간에 대해 입증하기 위해 사람의 유전을 연구하였고 그 중 키의 유전을 설명하면서 회귀와 상관이라는 개념을 처음 소개하였다. 한편 생물학자로서 역시 다윈의 이론을 입증하려 하였던 웰턴은 골턴의 책을 읽고는 통계학적인 방법이 진화와 유전을 연구하는데 적합하다고 믿게 되었다. 통계학적 방법을 적용하기 위해서는 대표본이 필요하다고 생각한 웰턴은 해변에서 많은 게와 새우들을 잡아 머리와 등딱지 폭, 몸통 길이 등을 측정하는 한편 그때부터 골턴과 교류하기 시작하였다. 1891년 University College London(UCL)에 동물학 교수로 부임한 웰턴은 생물학 자료를 수학적으로 연구하기 위해 같은 대학의 응용수학 교수이던 피어슨을 찾아갔고 그 이후 두 사람은 대단히 밀접한 관계를 유지하게 되었다. 1892년에 골턴과 피어슨이 만나게 된 것도 웰턴을 통해서였다.

피어슨의 통계학 연구에서 웰턴이 차지하는 역할은 매우 큰 것이었다. 물론 피어슨은 웰턴과 교류하기 이전인 1891년부터 일종의 시민 대상 강의인 그레삼 강좌(Gresham lecture)를 통해 통계학 강의를 하고 있었다. 하지만 1891년부터 1892년 사이에 “통계의 기하학(the geometry of statistics)”이라는 제목으로 그가 했던 강의는 여러 가지 통계 그래픽 방법에 대한 것이었다 (‘히스토그램’이라는 이름도 이때 그가 처음 붙인 것이다). 하지만 웰턴의 영향을 받은 이후인 1893년부터 그는 생물학 데이터의 빈도분포를 다루기 시작

하였다. 웰턴과의 교류로 인해 피어슨 자신이 새로운 연구를 개척하기 시작했다는 자신감은 “생물학의 역사에서 처음으로 생물에 대한 과학이 정밀 과학, 즉 수학적 과학이 될 기회가 왔다. 이제 사람들은 유전과 진화 문제를 새로운 견지에서 접근하고 있다” (Porter, 2004, p.238)라는 자신만만한 선언에서 잘 알 수 있다.

그 후 피어슨은 웰턴의 데이터를 비롯하여 여러 분야의 데이터를 분석하기 위한 수학적 통계학 이론들을 연구하여 1894년 (요약은 1893년)부터 ‘Contributions to the Mathematical Theory of Evolution’이라는 제목을 붙인 논문들, 그리고 나중에는 ‘Mathematical Contributions to the Theory of Evolution’이라고 제목을 약간 바꿔붙인 일련의 논문을 발표하기 시작하였다. 그 중 첫번째 논문은 뾰족뾰족한 빈도곡선을 적률을 써서 평균과 표준편차가 서로 다른 두 정규분포곡선으로 나타내는 것이었다 (Pearson, 1894, p.72). 피어슨은 하나의 빈도 곡선을 두 정규분포로 나눈 뒤 그 두 분포는 하나의 종이 자연선택에 의해 두 가지 종으로 분리되는 과정을 보여준다고 해석하였다. 즉 그의 주장에 따르면 많은 데이터를 모아 빈도분포들을 잘 살펴보면 다윈의 진화 이론을 입증할 수 있다는 것이었다. 그런데 그러한 작업에서는 원래 데이터 수도 많은데다 9차, 12차에 이르는 고차방정식도 풀어야하는 등 엄청난 계산이 필요했다. 따라서 피어슨은 여러 ‘인간 계산기(computers)’들을 동원하는 한편 1894년부터 ‘브룬스피가(Brunsviga)’라는 독일제 계산기(‘강철로 된 두뇌’라고도 불렸다)도 계산에 이용하였다. 이 사례는 통계학 계산을 위해 계산기를 이용한 매우 이른 사례에 해당할텐데, 심지어 그는 “직접 계산을 하지 않는 생물측정학자는 기생충과 다름없다”고 까지 주장하면서 계산의 중요성을 강조했다 (Porter, 2004, p.241). 이후 그가 발표한 아주 많은 연구를 전체적으로 평가한다면 거기에는 생물학적 성과보다는 통계학적 성과가 훨씬 풍성했다. 가령 적률 추정법, 피어슨의 곡선족, 피어슨의 곱적률 상관계수, 다중 상관계수, 그리고 피어슨의 가장 큰 업적으로 손꼽히는 카이제곱 적합도 검증 등이 거의 모두 같은 제목을 달고 발표된 연구에서 나온 것들이었다.

한편 같은 생물학자로서 베이트슨과 웰턴은 영국 과학진흥협회 동물학 분과 모임 등에서 진화의 연속성을 두고 일찍부터 대립하고 있었다 (헤니그, 2006, pp. 242-269). 그 대립이 표면화된 것은 1894년 베이트슨이 불연속적인 변이가 새로운 종을 생기게 한다는 주장을 담은 글 (Bateson, 1894)을 발표하면서부터였다. 그런데 그 글이 발표된 직후 웰턴이 베이트슨의 주장을 반박하는 서평 (Weldon, 1894)을 발표하였고 이때부터 두 사람은 웰턴이 1906년에 사망할 때까지 십여년 동안 극심하게 대립하였다. 피어슨 또한 논문을 통해 줄곧 웰턴의 주장을 옹호하게 되면서 베이트슨과 대립하였고 1906년 웰턴이 사망한 이후에는 생물측정학과의 대표로 논쟁을 이끌게 되었다.

그 와중에 1900년에 있었던 ‘멘델의 재발견’으로 말미암아 논쟁은 더욱 가열되었다. 1866년에 발표된 멘델의 잡종교배에 대한 논문이 멘델 사후 30여년이 지난 뒤 세 사람의 생물학자에 의해 재발견된 것이다. 그러자 베이트슨을 비롯하여 불연속적인 진화를 주장하던 사람들은 멘델의 논문이 자신들의 주장을 뒷받침하는 것으로 보고 이를 크게 환영하였다. 반면 생물측정학과는 멘델 유전학의 주장을 부정하기보다는 그 주장이 단순한 몇몇 유전 현상에만 적용될 뿐 보다 복잡한 유전 현상에까지 널리 쓰일만한 이론적인 모델은 아니라고 생각했다.

멘델의 재발견 이후 치열하게 전개되던 논쟁은 1906년 웰던이 46세라는 아까운 나이로 갑자기 사망하면서 그 열기를 상당 부분 잃어버리게 되었다. 한편 논쟁이 치열해지자 양쪽의 주장을 결합하려는 시도들이 드물게나마 나타나기도 하였다. 하지만 그러한 시도들은 거의 현실적인 효과를 거두지는 못하였고 양 진영의 주장을 통합하게 되는 피셔의 결정적인 연구가 나온 것은 1918년이 되어서였다. 표본으로부터 모집단에 대해 추론하는 ‘통계학적 추론’과 이론적 구조의 측면에서 모집단을 해석하는 ‘과학적 추론’을 구분할 때, 피셔의 논문 Fisher (1918)은 멘델주의 유전학이라는 이론적 구조의 측면에서 생물측정학과의 상관관계를 어떻게 해석할 것인지 보여주는 가장 대담한 ‘과학적 추론’이라고 평가받는다 (Aldrich, 1995, p.373). 그리하여 1910년대 말 이후 양쪽 주장이 ‘종합’된 결과는 피셔, 그리고 할데인, 라이트에 의해 집단유전학(population genetics)의 발달로 이어졌다 (Fisher, 1930). 비로소 수학적 연구가 생물학에서 중요한 한 부분으로 자리잡게 된 것이다.

2.2. 피어슨과 베이트슨의 대립

논쟁의 과정에서 피어슨은 다윈의 이론을 계량적인 방법으로 뒷받침하려는 생물학적인 목적 외에 자신이 처음으로 개척하고 있는 통계학이라는 학문에 대해 대단히 주목할 만한 발언들을 하게 된다. 그 대표적인 예를 골턴에게 보낸 편지에 쓴 다음 구절에서 잘 볼 수 있다: “유전학 연구는 첫째도 통계학 문제이고 둘째도 통계학 문제이며 단지 셋째가 겨우 생물학 문제이다” (1897년 2월 12일 피어슨이 골턴에게 보낸 편지 (Pearson, 1914-1930)). 그는 이러한 주장을 여러 곳에서 펼쳤는데 이로부터 우리는 유전학은 물론 통계학에 대해서도 별로 알지 못했던 피어슨이 짧은 시간만에 얼마나 변했는지 알 수 있다. 그리고 무엇보다도 이러한 발언들이 베이트슨을 비롯하여 통계학적인 방법을 사용하지 않는 많은 생물학자들을 얼마나 자극했을지도 짐작할 수 있다. 당연히 이에 대해 생물학자인 베이트슨은 일종의 ‘영역 침범’이라는 반응을 보였다. 앞서 논쟁이 시작될 무렵에 발표한 책에서 베이트슨은 골턴의 정규분포를 소개하고 불연속적인 변이를 설명하기 위해 봉우리가 둘 있는 도수분포 그림도 책에 실은 바 있다 (Bateson, 1894; Bateson, 1928b에 재수록, p.261, p.263). 1890년대 후반부터 피어슨이 비대칭 분포라든가 여러 가지 상관계수를 비롯한 통계학적 방법을 생물학 데이터에 적용하여 연속 변이에 대한 주장을 펼쳤을 때 베이트슨은 자신이 그러한 방법을 이용하지는 않았지만 그렇다고 해서 피어슨이 이용한 통계학적 방법 자체에 대해서 반박하는 경우는 없었다. 그러다가 피어슨과 베이트슨 사이의 논쟁은 1901년 *Biometrika*가 발행되기 시작하여 피어슨이 발언할 공간이 크게 확대된 이후 본격적으로 이루어졌다.

그 대표적인 경우를 *Biometrika*가 창간되는 직접적인 원인이 되었던 피어슨의 논문 (Pearson, 1901), 그에 대한 베이트슨의 반박논문 (Bateson, 1901), 또 그에 대한 피어슨의 재반박논문 (Pearson, 1902)에서 볼 수 있다. 피어슨은 처음 논문에서 나뭇잎이나 동물, 인간의 여러 특성을 측정하고 상관계수를 계산하여 그것들이 서로 비슷한 값을 갖는 것을 근거로 유전의 연속성을 주장하였다. 이에 대해 베이트슨은 상관계수라는 방법 자체를 문제 삼지 않고 그것들로부터 피어슨이 결론을 유도하는 과정이 과학적이지 못하다고 비판하였다. 그러자 피어슨은 *Biometrika* 제1권에서 베이트슨의 반박을 곧바로 재반박하면서

통계학을 사용하지 않는 베이트슨의 생물학은 낡은 생물학이라고 비판하였다:

생물학 연구에 통계학이 필요한 이유는 진화 문제가 오로지 수명, 번식력, 건강 그리고 질병에 대해 다루는 동태통계(vital statistics)의 문제이기 때문이다. 진화를 연구하는 사람으로서 이런 통계 없이 연구를 진행하기는 불가능하다. 이는 마치 Registrar-General (1830년대에 만들어져 영국 전역에 걸친 통계 업무를 담당하던 정부기관)에서 인구를 헤아리지도, 질병을 분류하지도, 통계 이론에 대한 지식도 없이 전국적인 사망률에 대해 말하는 것과 같다. 내가 보기에 베이트슨씨로 대표되는 생물학자들이 하려고하는 연구가 바로 그런 것이다 (Pearson, 1902, pp. 320-321).

그 글에서 피어슨은 통계학에서는 변이, 상관, 유전 등의 개념을 명확히 양적으로 정의하는데 그러한 통계학을 이용한 생물학이 새로운 생물학이라고 주장하였다. 진화에 대한 연구가 통계학적인 문제라는 주장은 피어슨 혼자만의 주장이 아니었다. 똑 같은 주장을 웰턴의 글에서도 찾아볼 수 있다:

동물 진화 문제가 본질적으로 통계학의 문제라는 것은 아무리 강조해도 지나치지 않다. 어떤 특성에 있어서 일정한 정도로 비정상적인 모습을 나타내는 동물의 백분율, 한 기관에서 비정상적인 정도에 따른 사망률 차이, ..., 이런 것에 대한 지식은 종들의 과거 역사와 미래 운명을 연구하는 유일하게 타당한 기초이다 (Weldon, 1893, p.329).

이러한 주장에 대한 베이트슨의 비판은 다음과 같은 인용문에서 잘 볼 수 있다:

유전학의 문제를 해결하려면 정밀한 실험분석을 하는 길 외에는 다른 길이 없다. 유전이 이루어지는 생리학적 과정에 대해 알지도 못한 채 대규모 현상을 포괄적인 통계학적 방법으로 연구하는 것이 그럴 듯 해 보일지 모르지만 그런 방법으로는 아무 진전도 이룰 수 없다 (Bateson, 1904a, p.130).

거친 통계학적 방법은 유전 현상에 활용할 만한 합당한 도구가 아니다. 생물측정학자들은 마치 프로크루스테스처럼 상관 표에다가 분석하지도 않은 데이터를 억지로 끼워맞추려하는데 제대로 훈련 받아 분별력을 지닌 사람이라면 그런 방법을 쓰지 않는다. 엄밀한 실험 검사를 거쳐, 연구 대상인 특정 식물과 동물의 습성이나 성질에 대해 완전히 정통한 사람이 사실에 대해 세밀하게 분석해야만 진리를 찾을 수 있다 (Bateson, 1904b; Bateson, 1928a에 재수록, p.240).

또한 베이트슨은 유전현상에 대해 골턴이 주장한 ‘조상 유전의 법칙’(law of ancestral heredity, 즉 어떤 사람의 형질은 부모로부터 1/2, 조부모로부터 1/4, 증조부모로부터 1/8, 등등이 더해진 것이라는 일종의 중회귀모형)을 멘델의 법칙이 재발견되기 전에 나왔던 여러 이론 중 하나로 언급한 다음 골턴과 생물측정학과를 다음과 같이 비판하였다.

골턴의 방법은 분석이 부족하기 때문에 실패하였다. 그가 만든 식은 어떻게 보더라도 유전 법칙의 하나라기보다는 진짜 유전 법칙의 부차적인 결과 중 하나일 뿐이다. 골턴의 분석적이지 못한 방법을 연장시키고 피어슨 교수와 생물측정학과로부터 유발된 소위 유전학 연구는 ... 궁극적으로 통계학 이론의 발달에 이바지할지 모르지만 유전학 문제에 적용되었을 때는 결과적으로 그 연구가 애써 밝히려 했던 질서를 감추는 역할만 하였다 (Bateson, 1913, pp. 6-7).

눈 색깔의 유전을 밝히기 위해 골턴은 광범위하게 혈통을 수집한 바 있고 피어슨교수와 그의 조수들은 유사한 종류의 생물측정학적 연구를 발표하였다. 통계학적인 종류의 다양한 결론이 그러한 데이터에 바탕을 두고 있었지만 정밀한 분석은 하나도 없었기 때문에 그렇게 얻은 결과에서 유전학적인 중요성이라고는 찾을 수 없다 (Bateson, 1913, p.107).

말 색깔의 유전 문제는 피어슨 교수와 그의 조수들이 많이 연구한 주제이다. 그들은 말의 혈통 기록들에 생물측정학 방법들을 적용하여 표로 만들고 연구하였다. 그 결과 여러 가지 주장들이 나왔는데 그 가운데에는 말 색깔의 경우에는 멘델의 원리가 전혀 들어맞지 않는다는 주장도 있었다. ... 피어슨교수의 상관계수표는 6000마리가 넘는 말의 기록을 편집하여 만든 것인데 ... 유전의 진짜 모습은 발견하지 못했다 (Bateson, 1913, pp. 124-125).

그런데 한편으로 생물학 연구를 위해 통계학적 방법이 과연 타당한 방법인가를 두고 피어슨과 베이트슨이 벌인 논쟁은 사실 새로운 것은 아니었다. 우리는 베이트슨의 비판, 즉 많은 데이터를 바탕으로 한 통계학적 방법이 아닌 정밀한 실험을 통해 생리학적 과정을 연구하는 것만이 올바른 방법이라는 단호한 비판으로부터 그 비판이 나오기 약 40년전에 등장했던 프랑스의 저명한 의학자 클로드 베르나르의 비판을 떠올리게 된다. 베르나르는 19세기 중반 통계학적 방법이 의학 연구에 사용되는 것을 보고 통계학적 방법이 아닌 '실험의학'을 옹호하는 책 (베르나르, 1985)에서 다음과 같은 주장을 펼친 바 있다.

1. 생물의 물리화학적 현상을 계산으로 환원할 수 없다.
2. 생리학적 현상은 지극히 복잡해서 명확하게 나타내거나 서로 비교할 수 없다. 생리학과 의학에서 평균을 쓰는 것은 그러한 생물학적 특징을 파괴해버림으로써 결과가 겉보기에 정확해보이도록 한다. 의사가 한 질병에 대해 아주 많은 개인들을 관찰하여 평균을 낸 다음 그 결과로 각 개인 환자들에 대한 처방을 내린다면 그 의사는 실제로는 존재하지도 않는 환자에 대해 처방을 내린 셈이다.
3. 통계학이 점점 널리 쓰이는 이유는 사실들을 비교할 수 있게 해준다는 점이다. 그러나 의학에서는 통계학의 기초가 되는 병명, 사망 원인의 진단과 기록이 정확하지 않다. 기록이 정확하다 할지라도 성, 연령, 기질 등이 완전히 닮은 환자를 찾을 수 없으므로 비교가 불가능하다.

4. 설사 그 모든 문제들이 해결된다 할지라도 통계학은 확실성(certainty) 대신 단지 확률(probability)만을 알려줄 뿐이다.

베르나르의 견해에 따르면 과학적 법칙은 확실성 위에, 절대적 결정론에 입각해야 하므로 통계학에서 법칙이 나올 수 없다는 것이다. 즉 의학에서 필요한 것은 무수한 관찰을 거듭하면 큰수의 법칙에 따라 어느 비율로 치유될 것이라고 말하는 것이 아니며, 아주 많은 경우가 아닌 환자 개인이 치유될 것인가에 답할 수 있는 방법이라는 것이다. 그런데 통계는 개개의 경우에 대해서는 아무 것도 가르쳐주지 않으므로 통계학으로부터는 확실한 법칙에 따르는 현상을 설명하는 실험 과학이 나올 수 없다는 것이다. 그러므로 베르나르는 관찰의 나열에 불과한 통계학에 입각하고 있는 한, 의학은 참다운 과학, 즉 확실한 과학이 못 되고 영구히 억측과학(conjectural science)에 머물 것이라고 주장했다. 의학과 생물학의 차이가 있기는 하나 베이트슨의 비판은 베르나르의 비판과 매우 비슷한 것이다. 그리고 통계학적 방법에 대한 그와 같은 비판은 19세기 후반뿐만 아니라 지금까지도 호소력을 가진 비판 중의 하나이다. 이러한 비판들은 구체적인 통계학 이론이나 방법을 문제삼는 것이 아니라 의학이나 생물학 같은 특정 분야에서 그 방법의 효용을 부정하면서 관찰이 아닌 실험 연구의 필요성을 강조하고 있다.

베이트슨이 볼 때 통계학적 방법에만 의지해서 생물학에 대한 주장을 하는 피어슨은 뛰어난 수학자일지는 몰라도 생물학에 대해서는 아마추어에 지나지 않았던 것이다. 피어슨은 자신의 주장이 전혀 베이트슨을 설득시키지 못하자 “베이트슨과 내가 변이, 불연속, 상관, 유전 등의 용어를 같이 쓰고 있지만 그 정의가 서로 다른 것 같다... 불행하게도 우리는 서로 완전히 다른 언어로 말하고 있다.” (Pearson, 1902, p.321)라고 비난 (또는 불평)했다. 결국 두 사람 사이의 논쟁이 정밀한 실험연구와 대표본 관찰연구 사이의 간격, 나아가 생물학과 통계학 사이의 간격을 좁히지는 못했던 것이다.

2.3. 피어슨과 피셔

이 논쟁에서는 현대 통계학의 두 거인이라 할 피어슨과 피셔가 모두 중요한 역할을 하였다. 특히 논쟁에 참가하지는 않았으나 Fisher (1918, 1930) 등의 연구로 논쟁을 마무리한 사람 가운데 하나인 피셔의 위치는 독특하다. 무엇보다 그가 양쪽의 주장을 종합할 수 있었던 이유는 그가 유전학자이자 통계학자였기 때문이다. 1890년 생인 피셔는 1909년에 입학한 케임브리지대학(Gonville and Caius College)에서 수학을 전공하기는 하였으나 우생학에 큰 관심을 갖고 대학 내에 우생학 학생 조직인 Cambridge University Eugenics Society를 주도적으로 만든 사람이었다 (MacKenzie, 1981, pp. 184-188). 그는 우생학 운동에 매우 열심이었는데 그가 1912년부터 1920년 사이에 발표한 97편의 글 가운데 92편이 서평을 비롯하여 우생학에 관련된 글이었을 정도였다 (Stigler, 2005, p.33). 즉 스티글러의 표현을 따르자면 1921년 이전까지 피셔는 통계학에 대해서는 단편적인 논문을 몇 편 썼을 뿐 그야말로 우생학에 폭 빠져있었던 (“Prior to 1921 Fisher had been immersed in eugenics”, (Stigler, 2005, p.47)) 것이다.

피셔가 두각을 나타내기 시작하던 1920년대 이전까지의 시기는 통계학사에서 상관과 곡선 적합의 시대, 곧 피어슨의 시대였다 (Yates와 Mather, 1963, p.98). 그러한 경향은

1911년에 초판이 나온 이후 1950년대까지도 가장 정평있는 교과서였던 율(G. U. Yule)의 *An Introduction to the Theory of Statistics* 가운데 1932년에 나온 제10판 (Yule, 1932)의 약 400쪽 가운데 상관과 곡선 적합에 관한 내용이 거의 절반을 넘게 차지할 정도인 데서 잘 알 수 있다. 1910년대까지는 피셔 역시 상관과 곡선 적합에 대한 논문을 조금 쓰다가 1920년대에 접어들면서 최대가능도 추정법, 충분통계량 등을 소개하면서 그러한 분위기를 크게 바꾸어놓았고 피어슨을 멀리 앞질러가게 되었던 것이다.

피어슨 역시 20세기에 접어들면서 검증법 등을 발표하였다. 하지만 우리가 아래에서 살펴볼 그의 과학철학을 감안해볼 때 피어슨은 기본적으로 통계학의 역할은 데이터를 잘 묘사하는 것(description)이라고 생각하였음을 알 수 있다. 사실 상관계수에 대한 연구에서 오늘날 우리가 ‘피어슨의 적률 상관계수’라고 부르는 것을 유도한 뒤 피어슨은 그것을 ‘상관계수의 가장 좋은 값(the best value)’이라고 불렀는데 (Pearson, 1896, pp. 264-265) 그 유도 과정을 보면 피어슨은 이변량 정규분포에서 모집단 상관계수에 대한 최대가능도 추정량(MLE)을 구한 셈이었다. 그런데 피어슨은 거기서 멈추고 말았고 피셔는 1922년 논문에서 MLE를 자신의 중요한 이론적인 연구 가운데 하나로 만들었다. 피어슨은 누구보다 통계학의 위상을 높게 생각하고 중요한 과학 분야라고 믿었지만 그의 통계학적 성과가 피셔와 다른 이유에는 그가 생각한 과학과 피셔가 생각한 과학 사이의 차이도 분명 있었을 것이다. 그러므로 수학적 능력으로 두 사람의 연구를 설명하기보다는 그러한 철학적인 견해로 두 사람을 비교해보는 것도 향후의 연구주제 가운데 하나일 것이다.

한편 긴 논쟁의 주역이었음에도 불구하고 피어슨의 연구가 유전학의 역사에서 언급되지 않은 반면 피셔는 유전학자로 뚜렷이 역사에 남았다. 두 사람이 모두 수학을 전공하였지만 웰턴과의 만남으로 비로소 생물학에 관심을 갖고 항상 통계학을 생물학보다 우위에 두었던 피어슨에 비해 피셔는 대학 시절부터 우생학 운동을 통해 지속적으로 생물학에 관심을 갖고 있었다. 따라서 두 사람의 차이에는 유전학에 대한 전문성의 차이도 작용하였다. 그런데 20대에 교수 자리를 얻은 피어슨에 비해 피셔는 상당히 늦게 대학에 자리를 잡았다. 그가 1919년부터 1933년까지 Rothamsted Experimental Station에서 통계학자로 일한 후 대학의 교수로 일한 경력을 보면 “Professor of Eugenics” (1933-1943, University College London), 이어서 “Professor of Genetics” (-1957, Cambridge)가 모두였다. 즉 그는 피어슨과 달리 한 번도 “통계학 교수”였던 적이 없었던 것이다. 또한 피셔 자신의 말에 따르면 그는 해마다 멘델보다 더 많은 식물을 재배했으며 (Fisher, 1952, p.9), 스스로를 수학자나 통계학자보다는 자연을 연구하는 자연과학자라고 생각했기 때문에 수학이나 통계학은 그에게 있어서 과학 연구를 위한 도구였다고 한다. 그의 딸인 Joan Fisher Box가 아버지에 대한 책의 제목을 R. A. Fisher, the Life of a Scientist라고 달았던 것도 그러한 이유 때문이라는 것이다 (Edwards, 2003, p.311). 유전학자로서 피셔는 비록 독특한 인물이기는 하였으나 유전학의 주류에 해당하는 염색체나 유전의 메커니즘에 대해서는 별로 관심이 없었으므로 결코 통계학자로서의 피셔만큼 지배적인 위치를 차지하지 못한 것으로 평가된다 (Yates와 Mather, 1963, p.120). 결론적으로 왜 피어슨은 논쟁을 종결시키지 못하였고 피셔는 성공했는가라는 문제는 향후 더 많은 연구가 필요한 주제인데 이 절에서 언급했듯이 두 사람의 과학에 대한 철학적 견해 차이, 유전학에 대한 전문성의 차이 등도 그 이유

로 들 수 있을 것이다.

2.4. 피어슨의 과학철학과 논쟁

앞에서 살펴본 바에 따르면 피어슨은 1890년대 초에 학문적으로 큰 변화를 겪었다. 아주 짧은 시간에 응용수학 교수가 통계학의 역할을 적극 옹호하는 생물측정학 연구자로 변신했기 때문이다. 그가 진화와 유전에 대해 관심을 가진 배경에 대해서는 MacKenzie (1981)가 우생학과 그의 연구를 연계하여 상세히 살핀 바 있다. 그런데 통계학의 역사에서 더 중요한 것은 피어슨이 생각한 통계학이 어떤 통계학인가라는 문제이다. 사실 19세기 영국에서는 나중에 Royal Statistical Society가 되는 Statistical Society of London이 이미 1834년에 만들어져서 1838년부터 Journal of the Statistical Society of London을 발간하고 있었다. 그런데 우리가 주목해야 할 것은 피어슨이 세계 유일의 통계학 교수로서 영국뿐 아니라 세계 최고의 권위와 명성을 누리면서도 평생 단 한번도 Royal Statistical Society의 회원이 아니었다는 사실이다 (Journal of the Royal Statistical Society에 실린 그의 부고 기사 (Journal of the Royal Statistical Society, 1936)는 바로 이 점부터 불만스럽게 짚고 있다). 즉 피어슨은 자신이 사회 조사로 대표되는 19세기 중반 영국의 통계학 전통을 잇는 것이 아니라 새로운 통계학, 즉 수학적인 통계학을 세우고 있다고 생각하였던 것이다. 여기서 우리는 과연 그의 통계학이 영국의 통계학 전통으로부터 것처럼 단절된 것인지 알아보려 한다. 이를 위해 우리는 그가 통계학이라는 분야를 갑자기 평생의 연구분야로 삼게된 배경을 그의 과학철학과 연계하여 살펴볼 것이다. 그의 과학철학이 가장 잘 드러나는 것은 1892년에 초판이 나온 The Grammar of Science이다. 피어슨은 생물학과 통계학에 대한 연구가 축적됨에 따라 이 책을 약 10년 간격으로 고쳐 펴냈는데 1900년에 제2판, 1911년에 제3판을 냈다.

2.4.1. 인과와 상관

논쟁에서 양 진영 사이에는 유전이라는 현상에 대해 무엇을 연구해야 할 것인가라는 근본적인 문제에서부터 차이가 있었다. 간단히 말해서 멘델주의자들이 무엇을 작용때문에 유전이라는 결과가 일어나는지 그 생리학적인 인과관계를 보려했다면 생물측정학과는 유전에 관계되는 여러 변수들 사이의 상관관계를 찾으려 하였다. 이러한 연구 목적의 차이는 당연하게 연구방법의 차이를 동반할 수 밖에 없었는데, 생물측정학과들은 수천 마리에 이르는 새우나 계와 같은 대규모 데이터를 대상으로 길이, 모양, 색깔 등과 같은 외양을 측정하고 세대별로 그 분포를 찾아 비교하려 하였다. 반면 멘델주의자들은 많은 개체보다는 단 하나의 개체에 대해서라도 길이 아니라 생물 내부로부터 유전 현상 배후에 있는 매커니즘을 알아내려고 하였다.

논쟁이 시작되기 전부터 피어슨은 The Grammar of Science에서 과학연구의 목적이 법칙화된 인과관계를 밝혀내는 것이 아님을 강력하게 주장하였다 (Pearson, 1892, 제4장). 또한 피어슨이 골턴에 대해 쓴 방대한 전기 중 제14장에서도 아래의 인용문에서 보듯 인과관계의 한계를 지적하고 상관을 강조한 내용을 많이 찾아볼 수 있다.

1889년 (상관에 대한 내용이 포함된 골턴의 Natural Inheritance가 발표된 해) 까지만하더라도 과학자들은 인과관계에 대해서만 생각하였다... 튀르고와 콩도르세가 쓴 것을 보면 보물 상자를 열기 위한 열쇠 (즉 상관)를 갖고 있지 못했기 때문에 그들이 사회 현상에 수학을 활용하는데 거의 아무런 진보도 이루지 못했음을 알 수 있다... 케틀레의 연구 중에 많은 부분과 초기 인류학자들의 연구가 별다른 소득이 없었던 것도 다 같은 이유 때문이다... 인과라는 개념은 부서져 가루가 되기 시작했다... 이제 우주에 대한 철학적인 관점은 변량들의 상관관계 시스템이라는 관점이 되었으며 그 시스템에서는 완전한 상관관계, 즉 (설사 물리학적인 현상들에서라 할지라도) 절대적인 인과관계에 접근은 하겠지만 결코 도달하지는 못한다 (Pearson, 1914-1930, 3A, pp. 1-2).

피어슨은 또한 물리학자들이 주장하는 인과관계란 그들의 필요에 따라 가정하는 순전히 이론적인 극한, 즉 완전한 상관관계일 뿐이라고 주장했다. 그러므로 이전까지 생물, 경제, 사회 현상에 대한 연구가 엄밀한 인과적인 법칙에 따르는 물리학과 비교할 때 과학의 반열에 오르지 못했던 것도 상관관계의 강도가 물리학에서의 강도보다 낮기 때문이므로 과학의 기준을 인과관계가 아닌 상관관계로 바꾼다면 그들도 모두 엄연한 과학이 될 수 있다는 것이다.

이와 같은 입장에 따라 그는 여러 가지 상관계수를 만들었을 뿐 아니라 데이터를 분석할 때에도 상관관계를 지나칠 정도로 내세웠다. 즉 피어슨에게 있어서 상관관계란, 그리고 그것에 바탕을 둔 통계학이란, 데이터를 분석하기 위한 하나의 도구나 방법 정도가 아니었다. 그보다는 모든 과학을 쇄신하고 과학의 범위를 크게 확장시켜주는 새로운 메타과학이었다. 그러했기 때문에 피어슨은 “유전과 진화의 문제란 무엇보다 통계학 문제”라고 과감한 주장을 할 수 있었던 것이다. 그의 생각으로는 생물학에서 제기되는 문제들은 아주 많은 변인들이 얽힌 것이기 때문에 이전과 같이 (그리고 멘델주의자들처럼) 인과적인 과학관으로 접근한다면 생물학은 과학으로서의 지위도 가질 수 없으므로 반드시 통계학이라는 새로운 과학관으로 접근해야만 한다는 것이다. 이러한 주장은 전통적인 생물학 교육을 받은 멘델주의자들이 받아들이기에는 너무나 지나친, 어쩌면 허황된 주장에 지나지 않았을지도 모른다. 따라서 논쟁은 별다른 소득도 남기지 못한 채 격화되기만 하였던 것이다.

2.4.2. 1830년대 영국의 통계와 피어슨

여기서 우리는 1890년대 초에 피어슨이 드러낸 과학관, 그리고 통계학에 대한 생각의 근원을 1830년대 영국통계학에서 찾아보려 한다. 영국에서는 이미 1660년대부터 과학자들의 단체로 왕립협회(Royal Society)가 존재했다. 그러나 이 단체가 마치 원로원처럼 고답적인 모습을 벗어나지 못하는데 불만을 가진 과학자들이 1831년에 영국과학진흥협회(British Association for the Advancement of Science: BAAS 또는 BA)를 만들게 되었다. 이 협회는 여러 분과를 두고 매년 영국 각 도시를 돌며 연구발표회를 가졌는데 창립 당시부터 큰 호응을 얻었다 (170여년이 지난 지금까지도 활발하게 활동하고 있으며 BA가 해마다 여는 Festival of Science는 영국에서 가장 큰 과학 전시 행사라고 한다). 통계학은 1833년에 BA의 한 분과(section F)가 되었다. 당시 BA의 분과로는 수학과 물리학, 화학

과 광물학, 지질학과 지리학, 자연사 (동물학과 식물학), 의학이 있었으므로 통계학은 여섯 번째 분과가 되었다.

그런데 통계학과 같은 신생 학문이 과학으로 인정받아 BA의 한 분과가 되는 것이 간단하지만은 않았다. 무엇보다 기존의 과학계에서 이 신생학문이 과학으로서 자격이 있는지 심사하고 그 학문의 역할과 미래도 규정하려 했기 때문이었다. 1833년 케임브리지에서 열린 BA 회의의 의장은 저명한 지질학자인 세지윅(Adam Sedgwick, 1785-1873)이었다. 그리고 케틀레의 조언을 받아 통계학 분과 설립 제안을 제출한 사람은 수학자로서 프로그램할 수 있는 컴퓨터에 대한 아이디어를 내놓기도 하였던 배비지(Charles Babbage, 1792-1871)였다. 의장으로서 세지윅은 배비지의 제안에 대해 다음과 같이 답했는데 그의 발언에는 당시 과학자들이 통계학을 어떻게 규정했는지 잘 나타나있다.

과학이란 기본적으로 측정하고 계산할 수 있는 것이어야만 한다. 따라서 통계학이 사실과 수치적 결과라는 형식을 가질 때에만 BA에 들어올 수 있다. 사실과 수치적 결과들은 정치경제학과 정치철학의 원자료(raw material) 역할을 한다. 그런데 정치 그 자체는 부당한 열정과 당파적인 적개심을 불러일으킨다. 그러한 것들은 BA가 학문의 올림피아로서 역할을 하는데 가장 중요한 화합을 깨트린다. 통계학자들이 자신의 영역에서 벗어나 정치라는 음산한 쪽과 어울리기 시작하는 순간, 불화의 악마가 학문의 에덴동산으로 들어온다 (Morrell과 Thackray, 1981, pp. 291-292).

통계학이 처음으로 과학으로 인정받아 BA의 한 분과가 된 것은 바로 이러한 조건하에서였다. 즉 단지 통계학은 사실과 숫자일 때에만 과학에 속하며, 그것이 어떤 정치, 경제, 사회적 입장이나 주장을 뒷받침하는데 이용된다면 즉각 과학의 영역을 벗어난다는 것이다. 그 규정에 따르면 통계 자료를 가지고 원인, 결과를 논해도 역시 과학이 될 수 없으므로 통계학의 역할이란 현상에 대한 순수한 묘사(description)에 국한된다.

BA에서만 통계학의 역할을 이처럼 규정한 것이 아니었다. BA에 통계학분과가 생긴 이듬해인 1834년에는 런던통계협회 (오늘날의 왕립통계협회의 전신)가 창립되었다. 당시 그 협회에서는 문장(seal)에 “Aliis exterendum”이라는 문구를 새겨넣었는데 이 라틴어 문구는 영어로 ‘to be threshed out by others’라고 번역된다. 그 뜻은 ‘데이터를 다른 사람들이 해석하도록 하자’는 것으로서 물론 여기서 해석이란 의견과 이론들을 동원하여 인과 관계를 따지는 것을 말한다. 결국 이 협회의 모토에 따르면 통계전문가들이 할 일은 데이터를 조사하고 모으는 일에 국한된다 (왕립통계협회의 문장에 지금도 나오는 밀집단이 바로 이것을 뜻한다). 뿐만 아니라 런던통계협회에서 1838년부터 발행하기 시작한 Journal of the Statistical Society of London (JRSS의 전신)의 첫 권 맨 앞에는 “통계는 원인을 거론하지 않으며 결과 역시 거론하지 않고 단지 사실만을 수집, 정리, 비교하는 과학이다. 사회, 정치적 문제에 대해 올바른 결론을 내릴 수 있는 바탕은 오로지 통계뿐이다” (Statistical Society of London, 1838, p.1)라는 주장이 실려있다.

물론 이처럼 선언적으로 규정된 통계학만이 영국에서 나타났던 것은 아니고 사회개혁 운동을 위해 통계자료를 활용하려는 사람들도 적지 않았다. 하지만 19세기말 새로운 수학적 통계학이 발달하기 이전까지 영국통계학은 1830년대의 모습에서 크게 달라지지 않았

다. 그렇다면 결과적으로 볼 때 피어슨이 자신의 과학철학과 어울리는 학문으로 통계학만 한 것을 찾기는 어려웠을 것이다. 피어슨이 *The Grammar of Science*를 통해 “과학의 통일성은 무엇을 대상으로 한 것인가에 있지 않고 단지 과학의 방법에 있다” (Pearson, 1892, p.16)고 했을 때 그의 과학은 인과관계의 규명보다는 묘사의 과학이었다.

묘사의 과학으로서 피어슨이 통계학을 선택한 것은 상당히 짧은 시간동안 이루어진 것 같다. 그는 수학적인 통계학을 연구하기 이전에 통계학을 최초로 강의한 사례, 즉 1891-1892년에 있었던 그레삼 강좌에서 다른 아닌 통계그래픽을 강의했다. 이는 통계학에 대한 피어슨의 초기 입장이 묘사에 중점을 둔 것이었음을 시사해준다. 또한 웰턴과 교류하게 된 이후 생물학 데이터를 분석하기 시작하면서 그는 점점 묘사의 과학이란 다른아닌 통계학이라는 생각을 굳혔던 것 같다. 피어슨이 발표한 초기의 통계학 논문에서 묘사의 과학으로서 통계학의 면모가 잘 드러나기 때문이다. 특히 기운곡선들을 소개한 두번째 논문 (Pearson, 1895)에서 피어슨은 나타난 그대로의 빈도분포를 수식으로 표현하려고 하였고 그의 곡선족(families of curve)은 바로 그러한 시도의 결과였다. 그의 통계학 연구가 생물학 연구에서 사실상 시작된 것 역시 피어슨의 과학에 대한 견해를 생각할 때 갑작스러운 것은 아니었다. 사실 본격적인 연구가 시작되기 이전에 이미 그는 정밀과학(exact science)에 대비되는 범주로서 생물학을 지질학과 함께 ‘descriptive science’라고 부르기까지 하였던 것이다 (Pearson, 1895, p.15).

이처럼 짧은 시간만에 피어슨이 생물학 데이터를 중심으로 통계학을 묘사의 과학으로서 선택했다는 것은 이미 그가 관념 속에서 통계학에 대한 이미지, 즉 1830년대 이후 영국 통계학의 전통으로부터 오는 이미지를 갖고 있었기 때문일 것이다. 이와 같이 볼 때 생물측정학-멘델주의 논쟁에서 생물측정학과의 기본적인 입장, 즉 설명보다는 묘사를 중요시하고 인과관계가 아닌 상관관계를 내세우는 입장은 새로운 것이 아니라 1830년대 이후 영국통계학의 흐름 속에서 나왔다고 보아야 할 것이다. 사실 피어슨의 통계학으로 대표되는 19세기말 이후 수학적 통계학의 발달은 사회통계조사가 주류를 이루던 1830년대 이후 영국의 통계학과 상당히 단절된 것으로 인식되기 쉽다. 하지만 우리가 살펴본 바에 따르면 오늘날의 수학적 통계학은 그 시작에 있어서 1830년대 통계학의 전통과 공유하는 것이 적지 않으며 일정 부분 그 연속성상에 있다고 할 것이다. 또한 피어슨이 대표본에만 집중한 것도 그가 여전히 19세기 중반 영국 사회통계의 영향을 받고 있었음을 말해준다. 1830년대에 맨체스터, 런던 등 영국 각 지역에서 생긴 통계협회들은 많은 비용과 시간을 들여 노동자들의 생활상태 등을 조사하였다. 당시만 하더라도 표본추출에 대하여 알지 못했기 때문에 그들은 항상 그 지역안에 있는 모든 가구를 방문하여 전수조사하려고 하였고 결과적으로 여러 개월에 걸친 조사 끝에 수천에 이르는 데이터를 모았던 것이다.

2.5. 통계학이 거둔 성과들

이 논쟁이 생물학 분야의 논쟁이기는 했지만 피어슨이 논쟁의 주역 가운데 한 사람이었고 또한 논쟁이 십여 년에 걸쳐 상당히 오랫동안 진행되었기 때문에 논쟁이 줄곧 생물학 논쟁에만 그치지 않고 통계학이 새로운 모습을 갖추는 과정에도 적지않은 영향을 미쳤다. 1894년 통계학 논문을 처음으로 발표한 이후 피어슨이 기운 곡선들을 이용하여 생물학 데

이더에 맞는 곡선을 찾는 연구를 하고 또한 상관계수를 가지고 유전에 대해 연구할 때만 하더라도 그의 관심은 생물학쪽에 더 기울어 있었던 것 같다. 하지만 연구가 거듭되면서 피어슨의 연구는 생물학만을 위한 이론이 아니라 보다 보편적으로 적용될 수 있는 통계학 이론에 대한 연구로 확산되어 갔다. 대략 1900년에 접어들 시기부터 생물학보다는 통계학 이론이 더 우선인 연구들이 나타나게 되는데 Barnard (1993)같은 사람은 1900년에 발표된 카이제곱 적합도 검증에 대한 연구 Pearson (1900)가 그러한 연구의 최초라고 주장하기도 하였다.

논문을 통한 통계학 연구 성과와 함께 이 논쟁으로 인해 통계학이 거둔 성과로는 논쟁의 주역으로 부각된 피어슨을 구심점으로 하여 여러 연구자들이 생물측정학과라는 하나의 집단으로 결집하게 되었다는 것이다. 피어슨은 1894년에 생물측정학 연구실(biometric laboratory)을 만든 바 있다. 처음에 그 연구실에는 응용수학을 공부하는 학생들만 모였으나 피어슨의 명성이 높아지면서 다양한 학생들이 모여들었고 이 학생들은 연구소를 떠난 뒤 통계학을 널리 전파하는 역할을 하였다 (Hilts, 1981, p.579). 당연히 연구실에서는 많은 공동연구들이 신속히 이루어졌고 얼마 지나지 않아 영국 국내 뿐만 아니라 여러 외국으로부터 점점 많은 연구자들이 새로운 학문을 공부하기 위해 연구실로 모여들게 되었다. 머지않아 이러한 성과는 1911년 최초의 통계학과 설립으로 이어지게 되었다.

한편 이 논쟁의 직접적인 결과로 1901년 최초의 수학적 통계학 학술지 *Biometrika*가 등장하게 되었다. 이 학술지가 생기기 전까지 거의 영국 왕립협회의 기관지에 논문을 싣던 피어슨은 1901년부터는 거의 모든 논문을 새로 창간한 *Biometrika*에 싣기 시작하였다. 이 새로운 학술지는 생물측정학-멘델주의 논쟁때문에 생긴 것이었다. 논쟁이 가열되면서 상대편의 주장을 반박하는 논문들이 양쪽에서 계속 나오게 되었고, 특히 왕립협회에 제출한 피어슨의 논문을 베이트슨이 심사하고 공격하는 일이 벌어지자 1900년 가을에 웰던이 생물측정학과만의 독자적인 학술지가 필요하다고 주장하기에 이르렀다. 이에 따라 1901년 10월부터 *Biometrika*가 나오게 되었던 것이다.

창간호 맨 앞에 편집자들이 실은 ‘The Scope of *Biometrika*’라는 글에 따르면 새로운 학술지의 목적은 “다른 학술지에서 체계적으로 수집, 발표되지 않고 있는 생물학 데이터들을 한 군데에 모으는 수단뿐만 아니라 그러한 데이터들을 과학적으로 다루는데 필수적인 통계학 이론을 전파하는 수단” (*Biometrika*, vol. 1, p.1)이라고 되어있다. 또한 이어지는 ‘The Spirit of *Biometrika*’에서는 “우리가 그 나라의 출생, 사망, 수명, 혼인과 출산에 대한 통계를 모르고서는 그 나라의 성장을 알 수 없듯이 어떤 생명체의 동태통계를 모르고서는 그 생명체의 변화를 알 수 없다. 따라서 진화를 연구하는 사람은 넓은 의미에서 모든 생명체들의 호적등기소장이 되어야한다”라고 생물학 연구에서 통계의 역할을 분명히 강조하였다.

이제 사실상 심사 과정도 없이 직접 발언할 수 있는 매체가 생겼으므로 피어슨의 독설은 거칠 것이 없게 되었다. 예컨대 그는 ‘On the Fundamental Conceptions of Biology’라는 글에서 베이트슨을 직접 거명하면서 멘델주의자들을 비판하고 생물측정학을 옹호하였다. 그는 베이트슨을 비롯한 생물학자들이 ‘혼란스럽고 정의조차 되지 않은 개념밖에 모르는’, ‘구닥다리 생물학과(the old school of biologists)’라고 부르는 한편 통계학을 이용

하는 학자들은 ‘아주 명확하고 계량적으로 분명한 개념’을 지녔다고 썼다 (Pearson, 1902, p.321). 또한 같은 글에서 그는 ‘베이트슨 씨가 생물학에 수학을 적용하기를 거부한다면 결코 진화를 설명할 수 없을 것’이며 ‘수학적 지식은 지금 물리학자에게 중요한 만큼 곧 생물학자에게도 중요한 것이 될 것’이라고 주장했다 (Pearson, 1902, p.344).

그런데 Biometrika는 영국왕립통계협회에서 발행하는 Journal of the Royal Statistical Society (오늘날 JRSS-A의 전신)가 오래전인 1838년부터 존재했었는데도 불구하고 생긴 통계학 학술지였다. 피어슨을 비롯한 생물측정학과가 새로운 통계학 학술지가 필요하다고 생각한 이유는 Journal of the Royal Statistical Society의 경우 출생, 혼인, 사망 등의 동태통계라든가 빈곤, 보건, 노동 계급의 생활 조건, 센서스 등과 같은 사회조사통계가 중심이었고 데이터에 대한 수학적 접근법은 거의 담고있지 않았기 때문이었다. 이는 영국 통계협회만 그랬던 것이 아니고 1888년부터 발행된 미국통계협회 학술지 Journal of the American Statistical Association의 경우도 마찬가지였다. 즉 19세기가 끝나고 20세기가 시작될 무렵까지만 하더라도 수학적 바탕이 되는 통계학 연구는 대학이나 학술지에서 찾아보기 어려웠던 것이다. 피어슨 이후의 통계학은 사실의 수집에 머무르지 않는 과학적인 추론이라는 점에서 그 이전의 통계학과 매우 다른 모습의 새로운 학문이라 할 수 있다. 비록 ‘biometry’라는 용어가 피어슨과 웰턴의 희망대로 진화와 유전 연구의 대명사가 되지는 못했으나 그들이 만든 Biometrika는 통계학의 새 시대를 여는 독보적인 역할을 맡았다. 1890년대에 통계학은 1830년대 이후 영국통계학에 비해 비약적인 진화를 겪었고 생물측정학-멘델주의 논쟁은 그 진화를 촉진하는 중요한 촉매제가 되었던 것이다.

이상의 논의를 정리해본다면 논쟁을 거치면서 피어슨과 웰턴은 베이트슨의 공격에 맞서 통계학의 역할을 더욱 강조하게 되었다. 그 결과 통계학은 유전학으로부터 구분되는 학문으로서의 독립적인 위상을 점점 굳혀가게 되었다고 할 수 있다. 결과적으로 볼 때 이 점이 이 논쟁에서 통계학이 거둔 가장 큰 성과라고 할 수 있다. 물론 유전학 역시 마찬가지였다. 논쟁의 초기에 피어슨은 유전학의 모든 주제는 통계학적인 문제라고 하면서 통계학의 역할을 과장해서 생각하기도 하였으나 20세기에 접어들어 유전학은 독자적으로 비약적인 발전을 거듭하여 통계학 못지 않은 뚜렷한 분야로 성장하였다. 논쟁의 성과를 통계학을 중심으로 정리하자면 다음과 같다:

1. 피어슨을 중심으로 수학적 통계학 연구자들이 최초로 결집하게 되었고,
2. 상관관계와 곡선 적합에 대한 연구가 크게 발달하는 계기가 되었으며,
3. 수학적 통계학을 위한 최초의 학술지 Biometrika가 탄생하는 데 직접적인 역할을 했을 뿐 아니라,
4. 우생학 운동과 더불어, 수학적 통계학이 연구소와 학과의 형태로 대학에 자리잡는데 중요한 역할을 했다.

3. 결론

지금까지 우리는 생물학 문제를 두고 벌어진 이 논쟁을 통계학을 중심으로 살펴보았다. 특히 논쟁에서 생물측정학파가 취했던 기본적인 입장의 배경을 살피기 위해 1830년대 이후 영국 통계학의 전통으로까지 거슬러 올라가 보았다. 그로부터 우리는 통계학과 무관하던 피어슨이 아주 짧은 시간만에 수학적인 통계학자로 변모한 데에는 1830년대 이래 영국 통계학의 전통이 작용하였을 것으로 추론해 보았다. 즉 우리가 살펴본 바에 따르면 19세기 말까지도 영국통계학은 통계학이 처음으로 과학으로 인정받을 때 영국 과학계로부터 요구 받은 성격을 지니고 있었다. 즉 통계학이란 인과관계를 따지는 것보다는 현상을 묘사하기만 하는 과학이라는 것이다. 피어슨 역시 그러한 이미지를 지니고 있었을 것이며 그와 같은 통계학은 ‘묘사 과학’인 생물학 연구에 매우 적합하다고 믿었을 것이다. 이에 따라 그는 “모든 생물학 문제는 통계학 문제이다”라고 주장하며 논쟁에 적극 나서게 되었고 통계 그래픽을 거쳐 기운 곡선과 상관에 대한 연구를 발표하게 되었을 것이다.

한편 이 논쟁을 촉매 삼아 통계학은 새로운 통계학으로 태어나게 되었다. 이 논쟁을 거치면서 여러 연구자들이 집단적인 학파를 이루어 연구를 진행한 결과 대학 내에 연구소와 학과가 생기고 수학적인 통계학을 위한 학술지까지 생기는 등, 그 이전까지 수학 연구자들의 산발적인 연구들과는 비교할 수도 없을 만큼 큰 진전을 이루게 되었던 것이다. 학파를 이루고, 학과와 연구실을 만들어 공동으로 연구하고, 그 성과물을 항상 발표할 학술지를 갖는 것이 신생 학문의 역사에서 얼마나 중요한가는 피어슨과 에지워스를 비교하면 잘 알 수 있다. 에지워스는 피어슨에 앞서 일찍이 1880년대부터 통계학 연구를 시작하여 1920년대까지 상관계수와 다변량 분포 등에 대한 연구결과를 발표하였다. 하지만 피어슨의 경우와 달리 그의 연구는 당대 사람들로 부터 제대로 이해받지 못한 것은 물론 거의 아무런 후계자도 얻지 못한 채 고립된 것에 머물렀다 (스티글러, 2005, pp. 620-627). 그 중요한 이유는 에지워스가 옥스포드대학의 경제학교수로서 생물측정학파의 일원이 아니었고 따라서 이 논쟁에도 참여하지 않았다는 데서 찾을 수 있을 것이다. 또한 우리가 영국 바깥으로 눈을 돌려 덴마크의 틸(T. N. Thiele, 1838-1910)의 연구와 피어슨의 작업을 비교해볼 수도 있다. 틸은 피어슨보다 앞서 그가 ‘half-invariant’라고 부른 누울(cumulant), 기운 곡선, 선형 모형 등에 대한 적지 않은 연구를 남겼음에도 불구하고 피어슨만한 명성을 전혀 누리지 못하였다 (Hald, 1981, 2000). 비록 그가 통계학과 보험수학을 연구하기는 하였으나 코펜하겐 대학의 천문학 교수로서 통계학을 연구하였고 피어슨처럼 통계학의 제도화를 위해 노력하지는 않았던 것이다.

그런데 한편으로 피셔가 이 논쟁을 종합시켜서 마무리했다고 하지만 몇 가지 자료들을 조금 자세히 들여다보면 피셔가 양쪽을 공평하게 종합시킨 것은 아니었다. 피셔는 일찍이 1924년에 이 논쟁을 일컬어 “과학의 역사에서 가장 쓸모없는 논쟁 중 하나” (Fisher, 1924, p.189; Morrison, 2002, p.40에서 재인용)라고 불렀는데 얼핏 보면 피셔가 논쟁에 참가한 양쪽 편을 모두 싸잡아 비난하고 있는 것 같지만 그렇지는 않았다. 그는 여러 군데에서 그보다 앞선 세대의 통계학자들, 특히 피어슨으로 대표되는 생물측정학파들을 (학문적인 비난이라기 보다는 거의 인신공격에 가까울 정도로) 비난하였다. 가령 피어슨이 죽은 지 한참 지난 1951년이 되어서까지도 피셔는 베이트슨을 기념하는 어느 강연에서 이 논쟁을 언

급하면서 피어슨을 비난하였다. 그는 논쟁에서 베이트슨이 상대해야 했던 것은 “변론에 능한 지칠 줄 모르는 논쟁가인 칼 피어슨이 멘델의 업적을 깎아 누르기 위해 썼던, 절반은 날조에 지나지 않는 통계학적 주장” (Fisher, 1952, p.5)이었다고 말했다. 즉 그는 피어슨을 아예 ‘통계학자’라고 부르지도 않으면서 피어슨이 논쟁에 이용했던 통계학 이론을 형편없는 것들이라고 몰아세운 것이다. 이러한 피셔의 혹평은 통계학의 역사에서 피어슨과 생물측정학과의 역할을 평가절하시키는 한편 통계학의 역사에서 이 논쟁이 별로 언급되지 않게 만들기도 하였을 것이다. 또한 멘델주의를 완전히 받아들이지 않고 도리어 멘델주의자들과 험악한 논쟁을 벌인 사실 때문에 생물측정학과의 평가는 유전학의 역사에서도 거의 불쾌한 기억 정도로 치부될 수밖에 없었다.

한편, 생물측정학과의 대부 격인 골턴이 거의 평생에 걸쳐 우생학 운동의 주역이었으며 피어슨을 비롯한 생물측정학과 역시 우생학에 대한 골턴의 입장을 어떤 식으로든 물려 받은 사람들이었다. 그런데 사회주의자였던 피어슨이 우생학에 상당히 동조적이었던 데서 알 수 있듯이 100여년 전 영국에서 우생학은 지금 우리가 생각하는 우생학과 매우 달랐다. 적어도 19세기말에서 20세기 초반까지 우생학은 거의 ‘건전한 상식’이었다고 해도 과언이 아닐 정도였다. 예컨대 JSTOR을 이용해 통계학과 관련 있는 학술지만 검색해 보면 20세기 초에 나온 *Biometrika* 뿐 아니라 *Journal of the Royal Statistical Society*, *Publications of the American Statistical Association* 등에 ‘eugenics’라는 단어가 굉장히 자주 등장했음을 확인할 수 있다. 피어슨이 UCL에서 운영한 연구실로는 생물측정학 연구실 외에 우생학 연구실도 있었으며 초기 통계학과는 재정이나 조직면에서 우생학과 밀접한 관련을 맺고 있었다.

사실 피어슨은 영국 우생학 운동의 주역 가운데 한 사람에 그치지 않고 죽기 2년 전인 1934년에는 히틀러의 독일 민족 개조 실험을 높이 평가하는 연설을 남기기까지 했다 (Klein, 1997, p.185). 비록 1936년에 사망한 피어슨이 나치의 학살을 목격하지는 못했지만 자신의 우생학 신념이 그런 목적으로 이용되는 것에 동의했을 리도 없을 것이다. 하지만 피어슨이 연설에서 “독일의 우생학 실험이 실패한다면 그것은 오로지 갓 시작된 독일의 수학적 통계학 연구 수준이 낮기 때문일 것” (1934년 피어슨의 연설 (Klein, 1997, p.185)에서 인용)이라고 분명히 밝혔듯이 1890년대부터 시작된 그의 수학적 통계학 연구에 우생학을 발전시키려는 목적도 뚜렷이 들어있었음은 부정할 수 없다. 이런 측면에서 생물측정학과는 비윤리적인 사이비 과학과 가까운 사람들로 또한 비난받게 되었다.

지금까지 살펴본 것처럼 19세기말 영국에서 현대적인 수학적 통계학이 발달하기 시작한 역사에서는 유전학이나 우생학처럼 수학이나 통계학 이외의 요인들이 크게 작용하였다. 우선 유전학과 같은 분야들은 통계학 이론이 발달할 수 있는 많은 데이터와 질문들을 제공하였다. 뿐만 아니라 그 분야의 문제를 둘러싼 논쟁을 통해 통계학이 다른 분야들과 구별되는 적절한 정체성을 마련하는 데에도 도움을 주었다. 물론 논쟁의 주역이었던 한편 새로운 통계학의 방향을 찾아야 했던 피어슨과 같은 인물들은 여러 가지 시행착오를 겪을 수밖에 없었다. 이 글의 서론에서 우리는 “통계학적인 방법의 특징은 무엇이며 통계학적 방법으로 과학 연구에서 무엇을 할 수 있는가?”라는 질문을 제시한 바 있다. 비록 20세기에 통계학은 대단히 많은 분야와 손잡고 넓고 깊이 발달하였지만 아마 그 확산 과정이 “생물학

문제란 통계학 문제”라고 장담했던 호기어린 피어슨의 방식은 아니었을 것이다. 하지만 통계학이 순수수학으로 환원되지 않는 한, 그리하여 이러한 질문의 유효성이 사라지지 않는 한, 피어슨의 시대는 여전히 통계학사에서 기록될 가치가 충분해 보인다.

참고문헌

- 베르나르 클로드 (1985). <실험의학방법론>, 유석진 옮김, 대광문화사, 서울; Bernard, C. (1865). *Introduction à l'étude de la médecine expérimentale*.
- 스티글러 스티븐 (2005). <통계학의 역사>, 조재근 옮김, 한길사, 파주, 경기도; Stigler, S. M. (1986). *The History of Statistics*. Belknap Press of Harvard University Press.
- 헤니그 로빈 (2006). <정원의 수도사>, 안인희 옮김, 사이언스북스, 서울; Henig, R. M. (2000). *The Monk in the Garden: The Lost and Found Genius of Gregor Mendel, the Father of Genetics*. Houghton Mifflin.
- Aldrich, J. (1995). Correlations genuine and spurious in Pearson and Yule, *Statistical Science*, **10**, 364–376.
- Barnard, G. A. (1993). Introduction by G. A. Barnard, in *Breakthroughs in Statistics: Volume 2: Methodology and Distribution*, Edited by Samuel Kotz and Norman L. Johnson, Springer, New York, 1–10.
- Bateson, W. (1894). *Materials for the Study of Variation Treated with Especial Regard to Discontinuity in the Origin of Species*, Macmillan & Company, London, (Reprinted in Bateson, 1928b, 211–312).
- Bateson, W. (1901). Heredity, differentiation and other conceptions of biology: A consideration of professor Karl Pearson's paper 'On the principle of homotyposis', In *Proceedings of the Royal Society of London*, **69**, 193–205, (Reprinted in Bateson, 1928b, 404–418).
- Bateson, W. (1904a). Experimental studies in the physiology of heredity, *Reports to the Evolution Committee of the Royal Society*, II, 1–131.
- Bateson, W. (1904b). Presidential address to the zoological section, british association: Cambridge meeting, 1904, in Beatrice Bateson, (Reprinted in Bateson, 1928a), 233–259.
- Bateson, W. (1913). *Mendel's Principles of Heredity*, Cambridge University Press, Cambridge.
- Bateson, W. (1928a). *William Bateson, F. R. S. Naturalist: His Essays & Addresses Together with a Short Account of His Life by Beatrice Bateson*, Cambridge University Press, Cambridge.
- Bateson, W. (1928b). *Scientific Papers of William Bateson*, Edited by R. C. Punnett, **1**, 29–68, Cambridge University Press, Cambridge.
- Bowler, P. J. (2003). *Evolution: The History of an Idea*, 3rd edition, University of California Press, California.
- Edwards, A. W. F. (2003). R. A. Fisher twice professor of genetics: London and Cambridge or 'A fairly well-known geneticist'. *The Statistician*, **52**, 311–318.
- Fisher, R. A. (1918). The correlation between relatives on the supposition of mendelian inheritance, *Philosophical Transactions of the Royal Society of Edinburgh*, **52**, 399–433.

- Fisher, R. A. (1924). The biometrical study of heredity, *Eugenics Review*, **16**, 189–210.
- Fisher, R. A. (1930). *The Genetical Theory of Natural Selection: A Complete Variorum Edition*, Oxford University Press, Oxford.
- Fisher, R. A. (1952). Statistical methods in genetics, *Heredity*, **6**, 1–12.
- Hald, A. (1981). T. N. Thiele's contribution to statistics, *International Statistical Review*, **49**, 1–20.
- Hald, A. (1998). *A History of Mathematical Statistics from 1750 to 1930*, John Wiley & Sons, New York.
- Hald, A. (2000). The early history of the cumulants and the Gram-Charlier series, *International Statistical Review*, **68**, 137–153.
- Hilts, V. L. (1981). *Statist and Statistician*, Arno Press, New York.
- Journal of the Royal Statistical Society (1936). Obituary: Professor Karl Pearson, F.R.S., 1857–1936, *Journal of the Royal Statistical Society*, **99**, 424–428.
- Kim, K. M. (1994). *Explaining Scientific Consensus: the Case of Mendelian Genetics*, The Guilford Press, New York.
- Klein, J. L. (1997). *Statistical Visions in Time: A History of Time Series Analysis 1662–1938*, Cambridge University Press, Cambridge.
- MacKenzie, D. A. (1981). *Statistics in Britain 1865–1930: The Social Construction of Scientific Knowledge*, Edinburgh University Press, Edinburgh.
- Magnello, M. E. (1998). Karl Pearson's mathematization of inheritance: From ancestral heredity to mendelian genetics (1895–1909), *Annals of Science*, **55**, 35–94.
- Magnello, M. E. (2005). Karl Pearson and the origins of modern statistics: An Elastician becomes a statistician, *The Rutherford Journal*, **1**, <http://www.rutherfordjournal.org/>.
- Mendel, G. (1866). *Experiments in Plant - Hybridisation*, Harvard University Press, Cambridge.
- Morrell, J. and Thackray, A. (1981). *Gentlemen of Science: Early Years of the British Association for the Advancement of Science*, Clarendon Press (Oxford University), Oxford.
- Morrison, M. (2002). Modelling populations: Pearson and Fisher on Mendelism and biometry, *British Journal for the Philosophy of Science*, **53**, 39–68.
- Olby, R. C. (1966). *Origins of Mendelism*, Constable, London.
- Pearson, E. S. (1936). Karl Pearson: An appreciation of some aspects of his life and work, *Biometrika*, **28**, 193–257.
- Pearson, K. (1892). *The Grammar of Science*, Walter Scott, London (2nd ed., 1900; 3rd ed., 1911; Everyman's Library ed., 1937).
- Pearson, K. (1894). Contributions to the Mathematical Theory of Evolution, *Philosophical Transactions of the Royal Society of London, Series A*, **185**, 71–110.
- Pearson, K. (1895). Contributions to the mathematical theory of evolution-II. Skew variation in homogeneous material, *Philosophical Transactions of the Royal Society of London, Series A*, **186**, 343–414.
- Pearson, K. (1896). Mathematical contributions to the theory of evolution, III. Regression, heredity and panmixia, *Philosophical Transactions of the Royal Society of London, Series A*, **187**, 253–318.

- Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, *Philosophical Magazine*, series 5, 157–175 (reprinted in *Breakthroughs in Statistics: Volume 2*, edited by Samuel Kotz and Norman L. Johnson, Springer, New York, 1993, 11–28).
- Pearson, K. (1901). Mathematical contributions to the theory of evolution IX: On the principle of homotyposis and its relation to heredity, to the variability of the individual and to that of the race. Part I. homotypos in the vegetable kingdom, *Philosophical Transactions of the Royal Society, Serise A*, **197**, 285–379.
- Pearson, K. (1902). On the fundamental conceptions of biology, *Biometrika*, **1**, 320–344.
- Pearson, K. (1914–1930). *The Life, Letters, and Labours of Francis Galton*, Cambridge University Press, London.
- Porter, T. M. (1986). *The Rise of Statistical Thinking 1820–1900*, Princeton University Press, Princeton.
- Porter, T. M. (2004). *Karl Pearson: the Scientific Life in a Statistical Age*, Princeton University Press, Princeton.
- Provine, W. B. (1971). *The Origins of Theoretical Population Genetics*, University of Chicago Press, Chicago.
- Reid, C. (1982). *Neyman - From Life*, Springer, New York.
- Rushton, A. R. (2000). Nettleship, Pearson and Bateson: The biometric-Mendelian debate in a medical context, *Journal of the History of Medicine*, **55**, 134–157.
- Statistical Society of London (1838), Introduce, fourth annual report of the council of the ststistical society of London, *Journal of the Statistical Society of London*, **1**, 1–13.
- Stigler, S. M. (2005). Fisher in 1921, *Statistical Science*, **20**, 32–49.
- Sturtevant, A. H. (1965). *A History of Genetics*, Harper and Row, New York.
- Weldon, W. F. R. (1893). On certain correlated variation in *Carcinus maenas*, *Proceedings of the Royal Society of London*, **54**, 318–329.
- Weldon, W. F. R. (1894). The study of animal variation, *Nature*, **50**, 24–26.
- Yates, F. and Mather, K. (1963). Ronald Aylmer Fisher 1890–1962: *Biographical Memoirs of Fellows of the Royal Society of London*, available form *the R. A. Fisher Digital Archive*, University of Adelaide.
- Yule, G. U. (1932). *An Introduction to the Theory of Statistics*, 10th ed., Charles Griffin and Company, London.

[2008년 2월 접수, 2008년 3월 채택]

The Biometry-Mendelian Controversy in the History of Statistics[†]

Jae Keun Jo¹⁾

Abstract

From mid-1890's, biometricians and Mendelians debated over Darwin's evolutionary theory. Biologist W. Weldon and Mathematician K. Pearson were leaders of the biometric school and biologist W. Bateson led Mendelian school. In this paper topics of the controversy such as causation vs. correlation, frequency distribution are considered. And in relation to the tradition of British statistics, we consider the philosophy of Karl Pearson revealed in this debate. Besides many statistical methods and concepts by Karl Pearson, the newly born mathematical statistics got a new journal *Biometrika*, a department in university, and a school of researchers from this controversy.

Keywords: Causation and correlation; Mendelism; biometry; evolution and genetics.

[†] This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2005-041-C00106).

1) Professor, Department of Informational Statistics, Kyungsung University, 110 Daeyeun-Dong, Nam-Ku, Busan 608-736, Korea. E-mail: jkjo@ks.ac.kr