Galton's blinding glasses. Modern statistics hiding causal structure in early theories of inheritance

Bert Leuridan

1 The Problem: Probability and Statistics as a Tool for Discovering Causal Patterns

More than a century ago, biological knowledge of the mechanism of heredity was rather scarce.[1] In that context, the influential scientist and statistician Francis Galton tried to analyze hereditary phenomena statistically. He discovered interesting phenomenological regularities and posited a theoretical or causal mechanism of hereditary transmission to explain them. But, as I will argue, his causal ideas were perniciously biased by the statistical techniques he used.

Nowadays it is commonplace to attribute to probability theory and statistical inference a central place in the philosophy of causality. Several authors, such as Pearl and Spirtes, Glymour & Scheines (Pearl, 2000; Spirtes et al., 2000), have built theories of causal inference and causal discovery on the basis of knowledge of (conditional) (in)dependence relations between variables, together with some graph theoretical theorems and a set of assumptions. Their theories are tightly linked with contemporary statistical techniques such as structural equation modeling (SEM).[2] To put it metaphorically: *probability and statistics are viewed as glasses through which we can see or detect causal relations*.

This raises a problem. If I succeed in showing that Galton's knowledge of statistics impeded him to successfully develop a biological theory of inheritance (and thus acted as *blinding glasses*), it could be called into

[1] I would like to thank Erik Weber, Joke Meheus and the other members of the Centre for Logic and Philosophy of Science (Ghent University) for their helpful comments and I would especially like to thank both anonymous referees for their very stimulating suggestions and their critical questions.

[2] 'Structural equation modeling' refers to many related techniques. Other, more or less equivalent labels are: 'covariance structure analysis', 'covariance structure modeling,' and 'analysis of covariance structures'. The term 'causal modeling' happens to be somewhat dated. (Kline, 2005, 9)

question whether contemporary statistical techniques are neutral with respect to their domain of application (i.e. with respect to the theory being developed or tested).

So here's the plan. In sections 2 and 3, I will discuss an important methodological difference between Mendel and Galton (the role of statistics in their scientific research) and I will shortly glance through the main characteristics of Mendel's theory.[3] Then, in section 4, I will show how probability and statistics generated two *explananda* which in their turn generated constraints for any theory of inheritance. In sections 4.1 and 5, I'll present Galton's *explanans* (his theory of heredity) and discuss the differences with Mendel's view on the matter. After parrying two possible counter-arguments to my reasoning (6.1, 6.2), I will summarize what this case-study shows us regarding the neutrality or non-neutrality of statistics in the work of Galton. Finally, in section 7, I will analyze the consequences for contemporary statistical techniques like SEM.

2 Gregor Mendel and Francis Galton: two different scientists

In the second half of the 19th century Francis Galton studied the processes of heredity. He started his research by considering 'hereditary genius' (Galton, 1869), but soon he turned to more easily observable characteristics. In *Natural Inheritance* (Galton, 1889) he bundled the results of more than twenty years of research on this topic. One of the most interesting aspects of Galton's work was the fact that he used and developed several modern statistical techniques, some of which are still used today (e.g. linear regression, which was mathematically developed by e.g. Karl Pearson).

At the time Galton started to work on the problem of heredity, Gregor Mendel had just finished a long series of experiments with pea plants (*P. sativum*). In 1866 he wrote his *Versuche über Pflanzenhybriden* (Mendel, 1933)[4] in which he meticulously presented his theory of inheritance. As this paper met little or no response in the biological community at the time, his ideas remained silent until they were rediscovered by Carl Correns (1900)

[3] My treatment of Mendel's theory of inheritance will closely follow the 'textbook' interpretation of his work, i.e. the interpretation that was formed in the beginning of the twentieth century by people like Bateson (1902) and that can still be found in modern genetics textbooks (Klug and Cummings, 1997). This interpretation in many ways strongly differs from Mendel's original theory, which was finely discussed by Onno G. Meijer (Meijer, 1983) and which I present in my (Leuridan, 200+). I think this choice is legitimate, since the biometricians rivalled textbook-Mendel, not the original one (cf. *infra*, p. 3).

[4] This was translated as *Experiments on Plant Hybrids* (see Stern and Sherwood, 1966, 1–48).

and Hugo de Vries (1900).[5] In the meantime, Galton independently developed his biometric *theory of ancestral inheritance*. After 1900, these theories would become vehement rivals and it would take several years or even decades before the dispute between the Mendelians (e.g. Bateson, 1902) and the biometricians (e.g. Karl Pearson, but also Weldon, 1902) was settled.

Contrary to Galton, Mendel made little use of statistical theory. He inferred in a rather intuitive way from particular observations to general regularities.[6] So, given the current dominant role of statistics in the special sciences, shouldn't we expect Galton to have found the most 'true' regularities? Of course, every one knows that Mendel is still considered the founding father of genetics, while Galton's name is now only associated with dubious disciplines such as phrenology and eugenics.

3 Mendel's Theory of Inheritance

All of the phenotypic characteristics that Mendel observed in *P. sativum* were qualitative and discrete. He studied seven pairs of contrasting traits, such as seed shape (round or wrinkled) and stem length (tall or dwarf). His *explanandum* consisted of some very straightforward empirical regularities (phenotypic distributions). The task was to explain why after crossing dwarf plants with true-breeding tall plants, all the off-spring in the first filial generation was tall; or why, after selfing this off-spring, 75% of the second filial generation was tall, while 25% was dwarf (see figure 1.1).

Mendel's *explanans* was based on a causal mechanism invoking material bearers of hereditary traits. In each pair of contrasting traits, one is dominant and the other is recessive (round is dominant to wrinkled, tall is dominant to dwarf). These traits are caused and carried over from parental plants to filial plants by *unit factors* ('*Factoren*', see Mendel, 1933, 24). A unit factor either codes for the dominant trait or for the recessive trait. They occur *pairwise* in the individual pea plants, but singly in the gametes.[7] Roughly half of the gametes carries (a copy of) the first factor. The other half carries (a copy of) the second factor. Later this would be called the '*principle of segregation*'. As would be emphasized by William Bateson in the beginning of the 20th century, no gamete carries an intermediate factor

[5] For a nice demystification of this mysterious rediscovery and of the preceding neglect, see Meijer (1983, sections 1 and 2) and also Darden (1991, section 4.3.).

[6] This statement is perhaps somewhat unfair and should be nuanced. Mendel *was* trained mathematically and statistically (see Meijer, 1983, 128), but not in the 'new' tradition of Quetelet, Galton, etc.

[7] At least, this is what classical Mendelism has taught us. However, there is every indication that according to Mendel factors occur pairwise in heterozygotes, but singly in homozygotes (Meijer, 1983; Leuridan, 200+); cf. footnote 3.

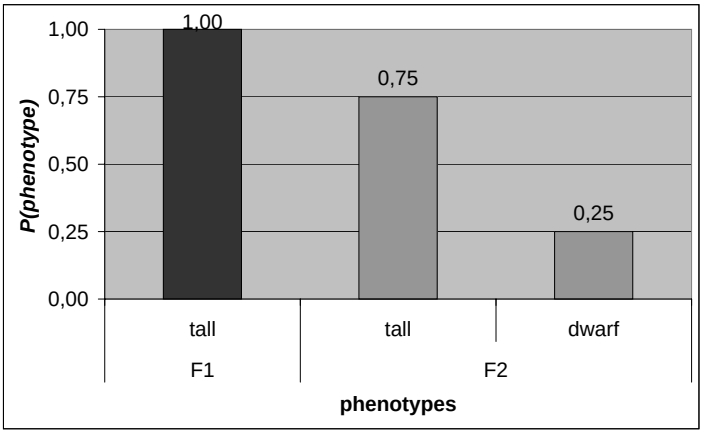

Figure 1.1. Mendel's explananda illustrated. The F1-generation resulted from crossing dwarf plants with true-breeding tall plants. The F2-generation resulted from selfing F1.

– this was dubbed '*the purity of the gametes*' in Bateson (1902, 108).[8]

A last point to be noted with respect to Mendel's theory is that the 'genotype'[9] of a pea plant is *screened off* or *d-separated* by the set of gametes that produced it (cf. Pearl, 2000, 16). Once it is established what unit factors are carried by the germinal cell and the pollen cell, the origin of these gametes plays no further role. Remote ancestry has no influence, conditional on the gametes.

To summarize, these are the main features of Mendel's theory that we should bear in mind in the following sections and which we should oppose to Galton's. Traits are grouped in pairs, in which one is dominant, the other recessive. They are caused and transmitted by unit factors that occur pairwise in individual organisms. Each gamete, however, contains only one factor, according to the principle of segregation. Moreover, the gametes are pure (they never carry intermediate factors), and they *screen off* the resulting individual from its ancestry.

[8] I will not discuss the 'principle of independent assortment', as Galton did not explicitly treat multi-hybrid crosses.

[9] I apologize for anachronistically using terms like 'genotype', 'phenotype' and their derivatives, but I think it's worth the trouble so as to avoid unnecessary circumlocution.

4 Statistics generating two *explananda* for Galton's Theory of Heredity

4.1 Preview of Galton's theory of *Ancestral Inheritance*

Like Mendel, Galton invoked bearers of hereditary traits to set up a causal mechanism of heredity.[10] He called them 'elements' or 'particles'.[11] But contrary to Mendel's unit factors, these did not occur in pairs. In each individual organism, an *indefinite* or *incalculable number of elements* responsible for the same phenotypic trait is present.

From a modern perspective, this may look weird, as chromosomes occur pairwise in diploid organisms. At the time Mendel and Galton developed their views on the mechanism of inheritance, however, cytological constraints were rather poor (Darden, 1991, chapter 7). This should be borne in mind if we want to assess the merits of Galton's work.

Mendel had a relatively clear view on the transmission of unit factors and on the possible genetic make-up of the gametes.[12] By contrast, the genetic make-up of gametes or individuals was never treated concretely in Galton's texts and the gametes played no inferential or predictive role in his theory. Although he stated, in the beginning of his *Natural Inheritance*, that "there is no direct hereditary relation between the personal parents and the personal child" and that "the main line of hereditary connection unites the sets of elements out of which the personal parents had been evolved with the set out of which the personal child was evolved" (both are quoted from Galton, 1889, 19), he directly predicted the traits of issue from the traits of its ancestors. These ancestors not only included the parental generation, but also the grand-parents, great-grand-parents, ... Galton was convinced that all ancestry may in principle have an influence on the set of elements from which the organism is built, and thus on its phenotype.[13] The influence is neither screened off by the parents, nor by their gametes. Therefore, Galton's theory can be called the *theory of ancestral inheritance*.

This suffices to show that the causal mechanisms proposed by Mendel and by Galton were really different. In the following sections, I will show

[10] I have found no explicit discussion of the concept of 'causation' either by Mendel or by Galton. Their writings, however, bespeak that both had a 'mechanistic' view on causation.

Note that in this respect they strongly differed from e.g. Pearson who had a Machian view on causation in which, first, causation is defined as perfect correlation and in which, secondly, the existence of causal relations in the empirical world is excluded *a priori* (Pearson (1900, chapter IV) or (1911, chapter V)).

[11] Galton also used the word 'element' to refer to the phenotypic traits themselves, instead of the particles that caused them (see section 5.1).

[12] See 'The reproductive cells of hybrids' (in Stern and Sherwood, 1966, 23–32).

[13] Cf. *infra*, p. 11.

in what ways this difference can (partly) be explained by laying bare the role played by Galton's statistical knowledge. His statistics generated two *explananda* which in their turn imposed influential constraints on any would-be *explanans*.[14]

4.2 The first *explanandum*: The normal distribution

Contrary to Mendel, Galton mostly studied continuous traits. In his *Natural Inheritance*, he paid a lot of attention to the schemes of distribution and the schemes of frequency of e.g. human strength, stature, span of arms, weight, breathing capacity, etc. (Galton, 1889, 35–50 and 200). All these characteristics, Galton noted, are normally distributed.

In the 19th century, the normal distribution played a very important role, not only in astronomy, but also in the biological and the social sciences. Adolphe Quetelet, the Belgian statistician and sociologist, discovered that measurements of phenomena such as human stature, birth ratios and crime rates were all normally distributed. The term 'Quetelismus' refers, then, to the exaggeration of the dominance of the normal distribution, i.e. to the view that "*all* naturally occurring distributions of properly collected and sorted data follow a normal curve" (Stigler, 1986, 201, my italics; see also 203–205). Galton explicitly acknowledged the influence of Quetelet and stated that the latter introduced the idea that the Law of Error (i.e. the normal distribution) might be applicable to human measures (e.g. Galton (1877, 493) and (1889, 54–55)).

Galton knew several ways to represent distributions of data. One way was to represent them graphically. Another way was to cite a series of eleven percentiles (the 5th, 10th, 20th, ..., 80th, 90th and 95th). But the most economical method, applicable in case the data were normally distributed (as seemed mostly the case), was to cite just two numbers: $\mathtt{M}$ and $\mathtt{Q}$. $\mathtt{M}$ was the mean or median. $\mathtt{Q}$ he called the 'Prob. Deviation' and it was defined as one half of the interquartile range: $\mathtt{Q} = \frac{1}{2}(\mathtt{Q}_2 - \mathtt{Q}_1)$, where $\mathtt{Q}_2$ and $\mathtt{Q}_1$ are the third and the first quartile respectively. It conveyed the dispersion of the distribution and thus played a role similar to the standard deviation σ.[15] Once $\mathtt{M}$ and $\mathtt{Q}$ were known, all percentiles could be calculated and the scheme of distribution could be drawn.

The normal distribution of Human Stature played a tremendously important role in Galton's biometrical work. He wrote:

> In particular, the agreement of the Curve of Stature with the Normal Curve is very fair, and forms a mainstay of my inquiry

[14] In fact, Galton's statistics generated at least three explananda (see footnote 32).

[15] The standard deviation σ is by definition larger that $\mathtt{Q}$, since $\mathtt{M} \pm \sigma$ includes about 68% of the observations, while $\mathtt{M} \pm \mathtt{Q}$ includes only 50% of them.

> into the laws of Natural Inheritance. (Galton, 1889, 57)

Several sets of data, collected by Galton himself, indicated that the median male stature `P` was 68.25 inch (Galton used the symbol `P` to refer to the median in the context of Stature) and that `Q` = 1.7 inch.

Obviously, the normal distribution generated a major constraint on Galton's theory of heredity. Whatever mechanism one was to propose to explain the processes of inheritance, it had to be able to explain why inherited traits are normally distributed. As Galton stated in 1877,

> The conclusion is of the greatest importance to our problem. It is, that the processes of heredity must work harmoniously with the law of deviation, and be themselves in some sense conformable to it. (Galton, 1877, 512)

What causes variables to be normally distributed? In 1810, Pierre Simon Laplace first introduced what was later to be called the *'Hypothesis of Elementary Errors'*: that the joint action of a multitude of independent 'errors' produces a normal distribution (Stigler, 1986, 201–202). Laplace's main topic of interest was the distribution of astronomical observations,[16] but his ideas forcefully influenced Quetelet and later also Galton. Galton wrote:

> The Law of Error finds a footing wherever the individual peculiarities are wholly due to the combined influence of a multitude of "accidents" [...]. (Galton, 1889, 55)

So the constraint imposed by the normal distribution was the following: Galton had to introduce a '*variety of petty influences*' in his biological theory (Galton, 1889, 16–17).

4.3 The second *explanandum*: Regression towards the mean

Perhaps the most important one of Galton's contributions to modern statistical theory was his concept of '*regression towards the mean*'. It would later result in the theory of linear regression (elaborated by Galton's protégé Karl Pearson). What exactly did Galton mean by this regression?

In the 1860's, he first tackled the topic of inheritance by studying human genius or talent and the way it was distributed within families. In *Hereditary Genius* (1869), Galton observed that, generally speaking, the relatives of gifted men (such as Johann Sebastian Bach or Jacob Bernouilli) are gifted

[16] If every measurement is the aggregate of many independent component measurements, each of them subject to small errors, the normal distribution of astronomical data is explained.

too, but less so. And the more remote a relative is, the less he is talented (Stigler, 2002, 176–177). In other words, Galton found a 'regression towards mediocrity'.

Of course, genius or talent is not easily observable, let alone measurable, but other characteristics are.[17] In 1884, Galton gathered hundreds or even thousands of records about the human population, called the *R.F.F. Data* (*Record of Family Faculties*). They comprised information about the Stature, Eye-colour, Temper, Artistic Faculty, ... of whole families (spanning several generations). The major part of *Natural Inheritance* concerned Human Stature, and the main research question was whether it is inheritable or not, i.e. whether the offspring of tall people is – on average – tall, or not.[18] But what do we mean by 'tall'?

As I stated earlier, Galton knew that male stature was normally distributed with $\mathtt{P} = 68\frac{1}{4}$ inch and $\mathtt{Q} = 1.7$ inch (see page 6). Women are slightly smaller than men, but this difference disappears if their statures are multiplied or transmuted by 1.08. These figures suggested a very straightforward criterion for 'tallness'. People are tall *iff* they have a stature larger than P. Human stature could be written as the sum of two components:

$$\mathtt{Stature} = \mathtt{P} \pm \mathtt{D}, \tag{1.1}$$

in which D is the individual's deviation from the mean (Galton, 1889, 51–52 and Chapter VII). So the sign and size of D indicated whether and to what extent an individual was tall or small and Galton's research question could be rephrased as follows: does the issue of large parents with a large D itself also have a large D?

The stature of the issue of unlike parents does not depend on the specific statures of the father and of the mother, as Galton's data revealed. It depends only on their *average* stature. Therefore, Galton introduced the concept of the 'Mid-Parent', which is defined as "an ideal person of composite sex, whose Stature is half way between the Stature of the father and the transmuted Stature of the mother." (Galton, 1889, 87) The Mid-Parental Statures are normally distributed with $\mathtt{P} = 68\frac{1}{4}$ inch and $\mathtt{Q} = 1.21$ inch.[19]

[17] Galton took great pains, however, to classify men according to their talent, both on basis of their 'reputation' and on basis of their 'natural ability' (Galton, 1869, 37–38).

[18] The *R.F.F. Data* contained the statures of 205 pairs of parents and their 930 adult children. Other data sets also reported on Stature: the *Special Data* covered about 783 brothers from 295 families and the *Measures at the Anthropometric Laboratory* consisted of about 10,000 data (Galton, 1889, 71–82).

[19] Note that 1.21 is a theoretically predicted number, as $\frac{1.7}{\sqrt{2}} = 1.21$. According to the *R.F.F. Data*, the Mid-Parental Q is 1.19. Galton considered the agreement between these numbers to be excellent (Galton, 1889, 92–94 and 208).

$$\texttt{Stature Mid-Parent} = \frac{\texttt{Stat.Father + Transm.Stat.Mother}}{2}, \quad (1.2)$$

The *R.F.F. Data* revealed that the off-spring of tall Mid-Parents is on average taller than P and that the relation between the Mid-Parental Stature and the *average* Stature of the Son is stable (where the Son refers to both the sons and the transmuted daughters). If the Mid-Parental deviation is D, then the filial deviation is on average $\frac{2}{3}$ D:

> I call this ratio of 2 to 3 the ratio of "Filial Regression." It is the proportion in which the Son is, on the average, less exceptional than his Mid-Parent. (Galton, 1889, 97)

So Galton could now describe with numerical precision what he had discovered years before: the inheritance of characteristics is subject to 'regression towards the mean'.

$$\texttt{Stat.Son} = \texttt{P} \pm \frac{2}{3}\ \texttt{D}, \text{ where D is the Mid-Parent's deviation.} \quad (1.3)$$

Take for example a Mid-Parent that is very tall, say 71.25 inch (D = 3 inch). Equation 1.3 predicts and figure 1.2 illustrates that its issue will, on average, be 70.25 inch (D = 2 inch = $\frac{2}{3} \times 3$ inch).[20]

So here we are presented with a second set of constraints. Every theory of heredity should be able to explain this filial regression. In the following section, I will show how Galton's theory of Ancestral Inheritance did this.[21]

[20] Figures 1.1 and 1.2 nicely illustrate the difference qua explananda between Mendel's theory and Galton's.

[21] 'Filial Regression' was not the only kind of regression that Galton discovered (Galton, 1889, 99–110). Related notions which will prove to be relevant in the following sections are 'Mid-Parental Regression' (equation 1.4), 'Parental Regression' (equation 1.5) and 'Fraternal Regression' (equation 1.6):

$$\texttt{Stat.Mid-Parent} = \texttt{P} \pm \frac{1}{3}\ \texttt{D}, \text{ where D is the deviation of the Son,} \quad (1.4)$$

$$\texttt{Stat.Son} = \texttt{P} \pm \frac{1}{3}\ \texttt{D}, \text{ where D is the deviation of one of his Parents,} \quad (1.5)$$

$$\texttt{Stat.Brother} = \texttt{P} \pm \frac{2}{3}\ \texttt{D}, \text{ where D is the deviation of a known man.} \quad (1.6)$$

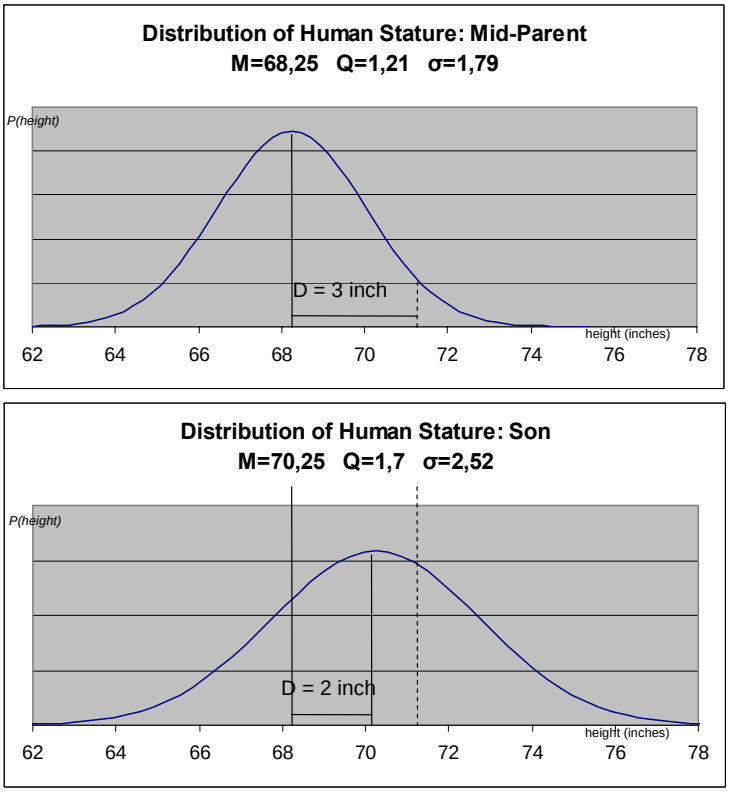

Figure 1.2. Filial Regression illustrated

5 Galton's Theory of Ancestral Inheritance as an *explanans*

5.1 Particulate Inheritance and the Hypothesis of Elementary Errors

Mendel's causal mechanism responsible for the processes of heredity consisted of unit factors in pairs. In Galton's theory, the situation was less clear-cut. He called it the theory of '*particulate inheritance*' and used the words '*elements*' and '*particles*' several times. But sometimes these elements seemed to denote (elements of) phenotypic traits, other times they might have referred to carriers of hereditary traits. Nonetheless, as Galton was influenced by August Weismann's theory of the germ-plasm and Charles Darwin's concept of Pangenesis, a 'material' interpretation of the elements or particles is certainly justified (Galton, 1889, 7–9 and 192–193).

Galton's hereditary particles are transmitted from parents to offspring,[22] but in principle every ancestor may contribute to an individual's elements. So the parents don't screen off the offspring from the rest of its ancestry. The influence of remote ancestry is obviously smaller than the parental influence. Nevertheless, it exists and played an important role in Galton's predictive inferences. The separate contribution of each ancestor follows a very simple rule, which would later be called the '*Law of Ancestral Heredity*' by Karl Pearson:[23]

> [...] the influence, pure and simple, of the Mid-Parent may be taken as $\frac{1}{2}$, and that of the Mid-Grand-Parent as $\frac{1}{4}$, and so on.

[22] Since phenotypic traits sometimes seem to skip a generation, he distinguished between *personal* elements (causing the traits they code for to be present), and *dormant* or *latent* elements ('unused' elements, having no phenotypic influence). Prima facie, this strongly resembles the Mendelian distinction between dominant and recessive traits or factors. There is an important difference, however. The relation of dominance/recessiveness is fixed for each pair of contrasting traits (round seed shape is always dominant to wrinkled seed shape in *Pisum*). By contrast, Galtonian elements that are latent in one organism can be personal in another. Galton had no definite answer to the question what determined whether an element would be latent or personal. He thought there were three possible answers:

> [...] first, that in which each element selects its most suitable immediate neighbourhood, in accordance with the guiding idea in Darwin's theory of Pangenesis; secondly, that of more or less general co-ordination of the influences exerted on each element, not only by its immediate neighbours, but by many or most of the others as well; finally, that of accident or chance [...]. (Galton, 1889, 19)

[23] Galton explicitly discussed the validity of this 'Law of Ancestral Heredity' for the inheritance of personal elements. At the end of his *Natural Inheritance* he hypothesized that it would also apply to the latent elements (Galton, 1889, 187–191).

> Consequently the influence of the individual Parent would be $\frac{1}{4}$, and of the individual Grand-Parent $\frac{1}{16}$, and so on. (Galton, 1889, 136)

Taken together, the set of all ancestors fully determines the Son's set of elements, as

$$(2 \times \frac{1}{4}) + (4 \times \frac{1}{16}) + (8 \times \frac{1}{64}) + \ldots = 1. \tag{1.7}$$

At first glance, this picture seems paradoxical. Although all hereditary influence passes through the parents, there is still room for the influence of the grand-parents, great-grand-parents, etc. This semblance of paradox is dissolved if we distinguish between *personal allowance* and *ancestral allowance*. Personal allowance is the allowance 'pure and simple' and it is governed by the Law of Ancestral Heredity (so that e.g. the father's personal allowance is $\frac{1}{4}$). The ancestral allowance comprises all influence that is just passed through an ancestor. An individual's total allowance thus is the sum of its personal and its ancestral allowance.[24]

From this picture it is easy to explain why characteristics like human stature are normally distributed. In principle, an infinite number of ancestors influences the set of hereditary particles of a man. And his 'genotype' consists of an indefinite or incalculable number of elements. This makes sure that the inheritance of traits is determined by a '*variety of petty influences*' (Galton, 1889, 16–17).[25] If it is assumed that these are to some degree independent of one another (as Galton did), a physical basis is provided for the Hypothesis of Elementary Errors. So the theory of Ancestral Heredity is capable of explaining the normal distribution of phenotypic traits (Galton, 1889, 84–85).

5.2 The Law of Ancestral Heredity, Dilution and Taxation

Now what is the cause of Filial Regression (equation 1.3)? Why is it that offspring tends to be more mediocre than its parents? Galton proposed an answer in his paper "Regression towards mediocrity in hereditary

[24] Galton did not use the concepts 'personal allowance' and 'ancestral allowance' in *Natural Inheritance*. They appear one time in his "The Average Contribution of each several Ancestor to the total Heritage of the Offspring" (Galton, 1897, 441).

[25] Galton incorporated two more sources of 'petty influences' in his theory. First, whether or not an element will be personal or dormant depends on very numerous influences (Galton, 1889, 22). Secondly, a trait such as Human Stature is not one element, but "a sum of the accumulated lengths or thicknesses of more than a hundred bodily parts" (Galton, 1889, 83–84), and each element or length of a body part is subject to errors or environmental effects.

stature" (Galton, 1886), which he recapitulated in *Natural Inheritance* (Galton, 1889).

Suppose some Mid-Parent has Stature $\mathtt{P} \pm \mathtt{D}$ and call $\mathtt{D}$ her peculiarity. From equation (1.4), Galton stated, it follows that the peculiarity of the Mid-Grandparent is $\frac{1}{3}\mathtt{D}$, that of the Mid-Great-Grand-Parent $\frac{1}{9}\mathtt{D}$, etc. If each generation would contribute its whole peculiarity, we should expect the Son to inherit $\mathtt{D}(1 + \frac{1}{3} + \frac{1}{9} + \&c.) = \mathtt{D}\frac{3}{2}$. This contradicts the expected Filial Regression of $\frac{2}{3}$ (Galton, 1889, 134).

So Galton considered the possibility that the bequests of the successive generations are somehow taxed or diminished. His data did not allow to directly measure the size of this tax, but he had two limiting hypotheses. On the one hand, if the bequest of every generation is taxed just once, the tax rate has to be $\frac{4}{9}$.[26] On the other hand, if the tax is repeated at each successive transmission, the rate should be $\frac{6}{11}$.[27] Galton's data did not allow to choose between these hypotheses. But as the both values differed but slightly from $\frac{1}{2}$, he decided that this would be a very good approximation (Galton, 1889, 134–136).

Galton's reasoning lacked logical rigour and can be challenged from several sides.[28] Nevertheless, one should not consider it as totally *ad hoc*. Both in 1886 and in 1889, Galton concluded that, as the tax rate should be estimated to be $\frac{1}{2}$, the Mid-Parent contributes half of his peculiarity, the Mid-Grand-Parent one quarter, etc. That is, he combined it with the Law of Ancestral Heredity (see page 11).[29] In 1897, Galton published an extra argument for this law. If each ancestor *may* contribute to the heritage of the offspring ('as is shown by observation'), if remote ancestry contributes less than near ancestry ('as is well known'), if the contribution of the parents to the children is the same as that of the grand-parents to the parents, etc. ('as is reasonable to believe'), and if the total amount contributed equals 1 ('as is necessarily the case'), then only the series of $\frac{1}{2} + (\frac{1}{2})^2 + (\frac{1}{2})^3 + etc.$ can describe the share of the Mid-Parent, the Mid-Grand-Parent, etc. (Galton, 1897, 403). Michael Bulmer deems it very plausible that Galton had this argument in mind in 1886 (Bulmer, 2003, 246).[30] If he is right, as I think he is, Galton's choice for the tax rate of $\frac{1}{2}$ was not totally *ad hoc*.[31]

---

[26] $\mathtt{D}\frac{2}{3} = \mathtt{D}\frac{3}{2} \times \frac{4}{9}$

[27] $\mathtt{D}\frac{2}{3} = 1\mathtt{D} \times \frac{6}{11} + \frac{1}{3}\mathtt{D} \times (\frac{6}{11})^2 + \frac{1}{9}\mathtt{D} \times (\frac{6}{11})^3 + etc.$

[28] For a crushing discussion of Galton's derivation, see Bulmer (1998) and Bulmer (2003, 243).

[29] Note that the Law of Ancestral Heredity matches only the second of both limiting hypotheses, namely that the tax is repeated at each successive transmission.

[30] In the bibliography of Bulmer (2003), however, this article is wrongly dated to 1885.

[31] Bulmer's conviction is based on the following quote by Galton: "These and the foregoing considerations were referred to when saying that the law might be inferred

So now we see how Galton's theory explains regression towards the mean. Since the peculiarity D of the Mid-Parent is mixed with the smaller peculiarities of more remote ancestry, the Son's deviation from P is smaller than D. Galton used a very powerful metaphor to illustrate this:

> [The] effect resembles that of pouring a measure of water into a vessel of wine. The wine is diluted to a constant fraction of its alcoholic strength, whatever that strength may have been. (Galton, 1889, 105)

The exceptionality of the parents is diluted by the mediocrity of the rest of the ancestry (hence, I call this the *Dilution Theory*), so that their offspring is more mediocre too. But, as we have seen, the Dilution Theory needed to be completed with the concept of *Taxation* to get the correct ratio of Filial Regression.

6 Two possible objections and a conclusion

Before I turn to the conclusion, I want to anticipate two possible objections to my thesis that statistics played a blinding role in the development of Galton's theory of heredity. I will show that neither the observational nature of the data on Human Stature, nor the continuous or blending nature of the observed characteristics can be cited as an alternative explanation for Galton's failure to arrive at the 'right' theory of heredity.

6.1 Observational versus experimental data

One of Mendel's major merits was that he paid a lot of attention to his experimental set-up. By carefully selecting a well-suited organism (pea plants) and manageable pairs of opposing characteristics, and by meticulously planning the right monohybrid as well as multihybrid crosses, he was able to confirm his causal theory of inheritance – a theory that still is considered the basis of modern genetics, although it has been subject to a vast amount of changes, specifications and additions (see Darden, 1991).

It is certainly true that Galton preferred observational data about the human population. He preferred data about humans because he considered them more interesting or relevant. He was thrown back on observational data because experimenting on human beings would have been (and still is!) rather problematic.

In the 1870's, however, after having published *Hereditary Genius* (1869), but before collecting the *R.F.F. Data*, Galton did experiment with plants

with considerable assurance *à priori* [...]." (Galton, 1897, 403) Galton had indeed stated in 1886 that his law might have been deductively foreseen (Galton, 1886, 253).

(*sweet peas*, not to be confused with *P. sativum*). He weighted thousands of seeds to determine their size and then selected several sets for planting. Each set consisted of seventy seeds, divided in seven packets of ten seeds of exactly the same weight $(K, L, \ldots, Q)$. The K-class contained very heavy seeds, L the next heaviest, and so on. He sent these sets to his friends throughout the United Kingdom and asked to plant them according to very minute instructions and to collect the produce of each class separately. Seven experiments (with in total $7 \times 7 \times 10 = 490$ parental seeds) succeeded (see Galton (1877, 512–514) and (1889, 79–82 and 225–226)).

The experimental data showed that large seeds beget large seeds. But, as was the case with hereditary genius, '*Reversion*' could be observed (the label 'regression towards the mean' was not yet introduced in 1877). In 1877, he gave no exact value for the regression coefficient, stating only that it is constant. In "Regression towards mediocrity in hereditary stature", which dates nine years later, this lacuna is removed:

> It will be seen that for each increase of one unit on the part of the parent seed, there is a mean increase of only one-third of a unit in the filial seed; and again that the mean filial seed resembles the parental when the latter is about 15.5 hundredths of an inch in diameter. (Galton, 1886, 259) (see also (1889, 225))

This suffices to show that the difference between Mendel's and Galton's scientific practice should not be sought in the experimental data of the former and the observational data of the latter.

6.2 Alternative inheritance versus blended inheritance

Can't we explain the difference between Galton and Mendel by looking at the variables, i.e. the phenotypic traits, they studied? After all, Human Stature and the Size of sweet peas are continuous variables, while Mendel observed pairs of discrete, opposing characteristics.

It is certainly true that Galton paid a lot of attention to continuous traits and it is equally true that that was the best way to discover regression-phenomena.[32] In *Natural Inheritance*, however, he took great pains to argue that the theory of Ancestral Inheritance could encompass the transmission

[32] As I phrase it here, it looks as if Galton did not come upon the phenomenon of regression by merely analyzing his data, but that he actively sought for it. In fact, this is true. Galton needed 'Regression towards the Mean' to explain another statistically inspired *explanandum*.

In 1877, Galton published "Typical Laws of Heredity" (Galton, 1877). One of the main explananda in that paper was the fact that the distribution of characteristics (specifically the Size of sweet peas) remained more or less constant in each successive generation. Using the terminology of 1889, it was to be explained why both its `M` and `Q` remained constant.

of both *blended heritages* and *alternative heritages* (Galton, 1889, 12–14 and 138–153). To prove this point, Galton studied the transmission Eye-Colour:

> If notwithstanding this two-fold difference between the qualities of Stature and Eye-colour, the shares of hereditary contribution from the various ancestors are alike in the two cases, as I shall show they are, we may with some confidence expect that the law by which those hereditary contributions are found to be governed, may be widely, and perhaps universally applicable. (Galton, 1889, 139)

How could the Law of Ancestral Heredity be used to predict the distribution of Eye-colours in issue, conditional on the Eye-colours of its parents, grand-parents, etc.? Galton distinguished between three types of Eye-Colour: light, hazel and dark, and then treated the problem as if it concerned Stature.

Suppose you want to predict the stature of some man, S, but that you only have information about one of his parents, F (having peculiarity $\mathtt{D}$). By equation (1.2), both parents of F have on average the peculiarity $\frac{1}{3}D$, while his grand-parents (i.e. the great-grand-parents of S) have $\frac{1}{9}D$, ... From the Law of Ancestral Heredity it follows that F transmits only $\frac{1}{4}$ of his peculiarity to S; his parents transmit $\frac{1}{16}D$, etc. So the total calculable or predictable heritage that is transmitted through F is[33]

$$\mathtt{D}\{1 \times 1 \times \frac{1}{4} + 2 \times \frac{1}{3} \times \frac{1}{16} + 4 \times \frac{1}{9} \times \frac{1}{64} + \&c.\} \approx \mathtt{D} \times 0.30,$$

consisting of F's known personal allowance (0.25 $\mathtt{D}$) and its predictable ancestral allowance ($(0.30 - 0.25)\mathtt{D} = 0.05$ $\mathtt{D}$). By analogy, two parents have a known total allowance of 0.60, "leaving an indeterminate residue of 0.40 due to the influence on ancestry about whom nothing is either known or implied" (Galton, 1889, 149). This residue is a direct consequence of the fact that the ancestral influence is not screened off by the parents or the parental gametes in Galton's theory.

These results can be easily re-interpreted in the context of Eye-colour. We only need to interpret the personal allowance or ancestral allowance as

---

In the absence of regression, the dispersion $\mathtt{Q}$ would increase from generation to generation. The offspring of tall men would on average be as tall as its parents, some of it would even be taller; in the next generation, even taller issue would result, ... But this would contradict Galton's data about sweat peas, as well as the findings of Quetelet, Galton and others on human characteristics (Venn, 1889, 415). Regression was the perfect candidate to solve this problem, as it would act as a counterbalance to this dispersive tendency.

[33] See Galton (1889, 148–149). Note that Galton's formula on p. 149 contains a printing error.

fractions of the total number of children in a family that will inherit some specific trait (Galton, 1889, 149–150). If a parent has dark eyes, that will cause 30% of his children to have dark eyes. Two dark-eyed parents will cause 60% of their children to be dark-eyed. Of the residue, 40% in this case, Galton assigns 28% to dark eyes and 12% to light eyes, proportionally to their overall ratio in the population. This gives rise to the following table, from which it is easy to predict the distribution Eye-colours in issue, conditional on the Eye-colours of its parents and grand-parents.[34]

Contribution to the heritage from each	Data limited to the eye-colours of the					
	2 parents		4 grand-parents		2 parents and 4 grand-parents	
	I.		II.		III.	
	Light	Dark	Light	Dark	Light	Dark
Light-eyed parent	.30				.25	
Hazel-eyed parent	.20	.10			.16	.09
Dark-eyed parent		.30				.25
Light-eyed grandparent			.16		.08	
Hazel-eyed grandparent			.10	.06	.05	.03
Dark-eyed grandparent				.16		.08
Residue, rateably assigned	.28	.12	.25	.11	.12	.06

For example, from the premise that in a family there are two light-eyed parents, three light-eyed grand-parents and one hazel-eyed grandparent, you can calculate that on average 91% of the children will have light eyes. The rest, 9%, will be dark-eyed.[35]

Thus we can conclude that, even if Galton paid heavy attention to continuous variables, he also tried to incorporate the transmission of discrete characteristics.

6.3 Conclusion: the neutrality of Galton's Statistics

I have shown that, in the second half of the 19th century, Galton's views on the mechanism of heredity were so much constrained by his statistical explananda, that he failed to discover the Mendelian scheme.[36] From this it

[34] Note that only the distribution of dark eyes and of light eyes is calculated. This table is reproduced from Galton (1889, 213).

[35] $2 \times 0.25 + 3 \times 0.08 + 1 \times 0.05 + 0.12 = 0.91$,
$2 \times 0.00 + 3 \times 0.00 + 1 \times 0.03 + 0.06 = 0.09$ (see also Galton, 1889, 215, Table 19).

[36] By this I do of course not mean that Mendel's theory was either unconditionally true or that it is still used in its original form. I only mean that he laid the fruitful basis of modern genetics, even if the development of genetics involved a lot of changes to his

follows that statistics *can* bias scientific research. Moreover, Galton's case can be supplemented with examples from outside biology. In his utmost interesting book, *The Taming of Chance* (1990), Ian Hacking has argued that the practice of descriptive statistics by the bureaucracies of nation states affected the way people conceived of society, of other people and of themselves. So statistics have biased people's world views and they have (partly) constrained theory development in the social sciences too.[37] So clearly, probability and statistics were *not* always *neutral* scientific tools.

7 Consequences for Contemporary Statistics

In the literature on causal discovery, probability and statistics play an important role, as several flourishing research programmes use conditional (in)dependence relations as indicators of the presence and absence of causal relations and maintain strong ties to statistical techniques such as SEM (Pearl, 2000; Spirtes et al., 2000). In the SEM-literature, however, one is frequently warned not to draw causal inferences from structural models too quickly. Causal inference is only justified on basis of models that fit the data well. But good model fit is not enough. It may indicate that the model accurately reflects reality, but not necessarily so. It leaves open the possibility that the model is equivalent to one that corresponds to reality but itself is incorrect, or that it fits the data from a nonrepresentative sample but has poor fit in the population, or that it has so many parameters that it cannot have poor fit (Kline, 2005, 321). In addition to fitting the data well, the model should also correctly describe the causal relations between its variables.

A SEM-model has to be specified independently before it is confronted with the data. In the SEM-literature, heavy stress is laid on the need for reliable background knowledge or theory in this process.[38] But this

basic tenets (cf. Darden, 1991).

[37] In the *The Taming of Chance*, Hacking also discussed Galton's work. His interpretation of the relation between the normal distribution, regression towards the mean, and Galton's hereditary mechanism, however, strongly differs from what I have argued in this paper (Hacking, 1990, chapter 21). I plan to elaborate the differences between his interpretation and mine in the nearby future.

[38] The following quote illustrates this and shows at the same time that, contrary to what Pearl has argued, the exclusion of causal interpretations from SEM cannot be completely put down to SEM-practitioners seeking respectability by keeping causal assumptions implicit, or to the unsuitability of algebraic language to making or expressing causal assumptions (Pearl, 2000, 137–138).

> It is only from a solid base of knowledge about theory and research that one can even begin to address [the] requirements for inferring causation from correlation. Although facility with the statistical details of SEM is essential, it is not a substitute for what could be called wisdom about one's research

presupposes that this background knowledge is not perniciously biased by statistics itself.

So what about the neutrality of contemporary statistics? I would like to make two statements in this respect. First of all, we should not throw out the baby with the bathwater. One of statistics' fundamental strengths is that it has much thought for the presuppositions and internal limitations of each of the techniques developed (presuppositions with respect to the distribution of the data or the relation between variables, robustness against missing values and outliers, requisites concerning sample sizes, etc.). Add to this that tests have been developed for most of these presuppositions (tests which were not available to Francis Galton). Secondly, however, we should not be blindly optimistic. Galton's problem still is relevant today. Specifying a SEM-model involves that we prestructure our domain of interest. Examples of such prestructuring can be found at many places in Kline's introductory work (Kline, 2005). SEM-modeling involves marking out variables, specifying their possible values and developing and including methods for measuring them. It involves fixing covariance relations and perhaps imposing constraints on (covariances between) disturbance variables. It involves specifying the directionalities of presumed causal effects ...

This shows that, in practice, SEM (and like methods) should be used very carefully, i.e. one should always bare in mind the possibly *blinding* influence of the techniques themselves.

area. (Kline, 2005, 95)

BIBLIOGRAPHY

Bateson, W. (1902). *Mendel's Principles of Heredity. A Defense*. Cambridge University Press, London.

Bulmer, M. (1998). Galton's Law of Ancestral Heredity. *Heredity*, 81:579–585.

Bulmer, M. (2003). *Francis Galton. Pioneer of Heredity and Biometry*. The John Hopkins University Press, Baltimore, London.

Correns, C. (1900). G. Mendel's Laws concerning the behavior of progeny of varietal hybrids. In Stern and Sherwood (1966), pages 119–132.

Darden, L. (1991). *Theory Change in Science: Strategies from Mendelian Genetics*. Oxford University Press, Oxford.

de Vries, H. (1900). The law of segregation of hybrids. In Stern and Sherwood (1966), pages 107–117.

Galton, F. (1877). Typical laws of heredity. *Nature*, 15:492–5, 512–4, 532–3.

Galton, F. (1886). Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute*, 15:246–263.

Galton, F. (1889). *Natural Inheritance*. Macmillan and Co., London.

Galton, F. (1892 (1869)). *Hereditary Genius*. Macmillan and Co., London, 2nd edition.

Galton, F. (1897). The average contribution of each several ancestor to the total heritage of the offspring. *Proceedings of the Royal Society*, LXI:401–413.

Hacking, I. (1990). *The Taming of Chance*. Cambridge University Press, Cambridge.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. The Guilford Presss, New York, 2nd edition.

Klug, W. and Cummings, M. (1997). *Concepts of Genetics*. Prentice-Hall, New Jersey, 5th edition.

Leuridan, B. (200+). Mendel's Bayesian net or the causal structure of classical genetics. unpublished manuscript.

Meijer, O. G. (1983). The essence of Mendel's discovery. In Orel, V. and Matalová, A., editors, *Gregor Mendel and the Foundation of Genetics*, pages 123–178. The Mendelianum of the Moravian Museum, Brno.

Mendel, G. (1933). *Versuche über Pflanzenhybriden*. Ostwald's Klassiker der Exakten Wissenschaften. Akademische Verlagsgesellschaft, Leipzig.

Pearl, J. (2000). *Causality. Models, Reasoning, and Inference*. Cambridge University Press, Cambridge.

Pearson, K. (1900). *The Grammar of Science*. Adam and Charles Black, London, 2nd edition.

Pearson, K. (1957 (1911)). *The Grammar of Science*. Meridian Books, New York, reissue of the 3rd edition.

Spirtes, P., Glymour, C., and Scheines, R. (2000). *Causation, Prediction, and Search*. MIT Press, Cambridge, Massachusetts.

Stern, C. and Sherwood, E. R., editors (1966). *The Origin of Genetics: A Mendel Source Book*. Freeman, San Francisco.

Stigler, S. M. (1986). *The History of Statistics. The Measurement of Uncertainty before 1900*. The Belknap Press of Harvard University Press, Cambridge (Mass.).

Stigler, S. M. (2002). *Statistics on the Table. The History of Statistical Concepts and Methods*. Harvard University Press, Cambridge (Mass.).

Venn, J. (1889). Critical notices. *Mind*, XIV(55):414–420.

Weldon, W. (1902). Mendels laws of alternative inheritance in peas. *Biometrika*, 1:228–254.

Bert Leuridan
Research Assistant of the Research Foundation - Flanders (FWO)
Centre for Logic and Philosophy of Science, Ghent University, Belgium.
Bert.Leuridan@Ugent.be