Statistical Papers 42, 143–169 (2001)

Statistical Papers

The quincunx: history and mathematics

Joachim Kunert, Astrid Montag, and Sigrid Pöhlmann

Department of Statistics, University of Dortmund, 44221 Dortmund, Germany

Received: December 30, 1999; revised version: July 19, 2000

The quincunx was invented by Sir Francis Galton in 1873 to demonstrate binomial distributions. During the last 125 years it has been used to illustrate the laws of the binomial and the normal distribution. In the first part of this paper we describe the historical background of Galtons invention and take a look at the discoveries he got from it. The second part of this paper discusses the mathematical background of the quincunx. We discuss the various limit theorems which explain the phenomena observable by this apparatus.[1]

1 History

1.1 Meteorologist, Geographer and Eugenicist

Sir Francis Galton was born on February 16, 1822 in Sparkbrook, England, as one of nine children. Following his fathers order he first studied medicine, later on he studied mathematics. Because of his fathers death and the resulting inheritance he was able to give up his studies. Instead, he turned to scientific exploration tours. For his exploration of

[1]This work was supported by DFG.

Africa he was awarded the gold medal of the Royal Geographical Society. His interests then turned towards statistical aspects of meteorology where he tried to develop methods to analyse and forecast the weather. This includes first steps for the use of graphical methods to analyse multivariate data.

The publication of *The Origin of Species* (Darwin 1859) by his cousin Charles Darwin (1809-1882) influenced his interests. He lost his interest in geography and meteorology and turned to heredity. Additionally, he worked on anthropology, sociology and various other fields. Notably he invented the identification of individuals through fingerprints. In heredity, his main interest was the improvement of mankind by selection of reproduction. For this, he invented the expression *eugenics*. Later his friend Karl Pearson (1857-1936) wrote:

> *"We see that his researches in heredity, in anthropometry, in psychometry and statistics, were not independent studies; they were all auxiliary to his main object, the improvement in the race of man."* (quoted according to David (1987, p. 360))

In 1904 he founded the *Galton Laboratory* at the *University College*, London, where he had been appointed as professor for applied mathematics. The *Galton Laboratory* developed from Galtons *Eugenic Record Office* and Karl Pearsons *Biometric Laboratory*.

In 1909 Galton was knighted. He died on January 17, 1911 in Surrey, England.

Even if some of his views were refuted by the work of Gregor Mendel (1822-1884), which became known only in the 20th century, and even if eugenic effort can easily be abused and perverted, Galton yet was an outstanding scientific pioneer. Characteristical for Galtons studies of heredity was his statistical approach, following the example of the Belgian statistician Adolphe Quetelet (1796-1874), who first took anthropological measures and fitted normal curves to this data.

Like Quetelet, Galton tried to measure talent. Later he tried to show that talent is inherited. So he developed the idea of regression, which he published in *Hereditary Genius* (Galton 1869) and pursued for 20 years. Galton was fascinated by the *very curious theoretical law of deviation from an average* (Galton 1908) in biological and intelligence measurement data. Following Quetelet he suggested that the possibility to fit a curve to data could be considered as a test: do the data (e. g. height of different races) belong to one group or have they to be regarded as different (for each race)? His main theme was not to discover things in common but to find the differences and the heredity of those differences.

Galton worked out methods for scaling: If data of the same species can be represented by a normal distribution and if the unit of one species can be shown by measurable quantities (as stature or examination results) which follow such a curve - if therefore on the basis of measurable quantities a species can be identified as related - then the process can be reversed relating to quantities which are difficult to understand. A qualitative variable as e. g. intelligence, which at best is ordinal scaled, will follow a normal distribution if the data come from a single population. Based on that Galton worked out the *Statistics by intercomparison* (Galton 1875): The data are sorted by quantity and are fitted to the inverse of a normal distribution function (quantile function), so that e. g. the median is by 0 and the upper quartile is by 1 (see Fig. 1).

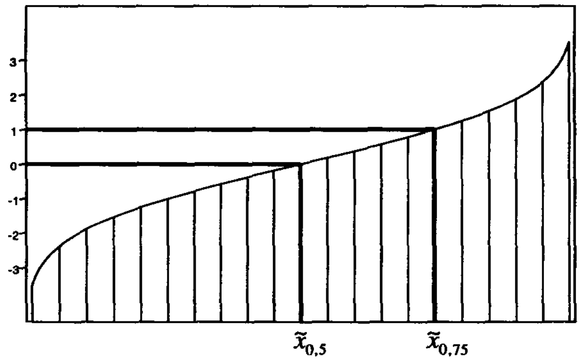

Figure 1:
21 observations are sorted by quantity and assigned to the inverse of a suitable normal distribution function in equidistant steps, so that the median $\tilde{x}_{0,5}$ is by 0 and the upper quantile $\tilde{x}_{0,75}$ is by 1.

This method grew to the most used (and most abused) method of scaling of psychological tests, even though this analogical conclusion - the scale is suited to the measurement of talent, because it is suited to the measurement of height - is shaky. Galton used it to give general statements about the abilities of different races, which corresponded to the prevailing prejudice of that time.

1.2 The invention of the quincunx

Galton's passion was the normal distribution and so he developed an instrument to illustrate it - a device, which generates a histogram which is similar to the normal distribution. It was most likely produced in 1873 by a company called Tisley & Spiller and was used by Galton as teaching aids, e. g. for a lecture at the Royal Society on February 27, 1874.

In *Natural Inheritance* (Galton 1889), which contains many of his scientific results from 1874 to 1889 and deals with the compatibility of heredity and the law of errors, Galton gave a detailed description of his device and called it "quincunx". This expression is borrowed from agriculture, where it is used to describe the cultivation of fruit trees in equidistant lines alternately "at gap". Also any arrangement of five objects like spots on a die with one in each corner and one in the middle is called quincunx (Posten 1986). This term became a synonym for the Galton board.

> *"I shall now illustrate the origin of the Curve of Frequency, by means of an apparatus shown in Fig. 7* [see Fig. 2], *that mimics in a very pretty way the conditions on which Deviation depends. It is a frame glazed in front, leaving a depth of about a quarter of an inch behind the glass. Stripes are placed in the upper part to act as a funnel. Below the outlet of the funnel stand a succession of rows of pins struck squarely into the backboard, and below these again are series of vertical compartments. A charge of small shots is enclosed. When the frame is held topsey-turvy, all of the shots runs to the upper end; then, when it is turned back into its working position, the desired action commences[...] The shot passes through the funnel and issuing from its narrow end, scampers deviously down through the pins in a curious and interesting way; each of them darting a step to the right or left, as the case may be, every time it strikes a pin. The pins are disposed in a quincunx fashion, so that every descending shot strikes again a pin in each successive row. The cascade issuing from the funnel broadens as it descends, and, at length, every shot finds itself caught in a compartment immediately after freeing itself from the last row of pins. The outline of the columns of shot that accumulate in the successive compartments approximates to the Curve of Frequency[...] and is closely of the same shape however often the experiment is repeated."*
> (Galton (1889, p. 63ff))

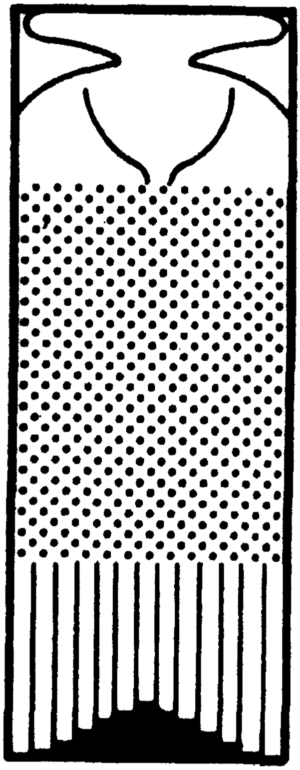

Figure 2:
Galtons illustration of the quincunx. (Fig. 7 in Galton (1889, p. 63))

The original quincunx (see Fig. 3) can be found today at the *Galton Laboratories*, London. The inscription by Galton reads as follows:

"Instrument to illustrate
The principle of the
Law of Error or Dispersion
by
Francis Galton F.R.S.

Charge the instrument by reversing it, to send all the shots into the pocket. Then sharply re-reverse and immediately set it upright on a level table. The shot will all drop into the funnel, and running thence through its mouth, will pursue devious courses through the harrow and will accumulate in the vertical compartments at the bottom, there affording a representation of the law of dispersion."
(quoted according to Stigler (1986, p. 277))

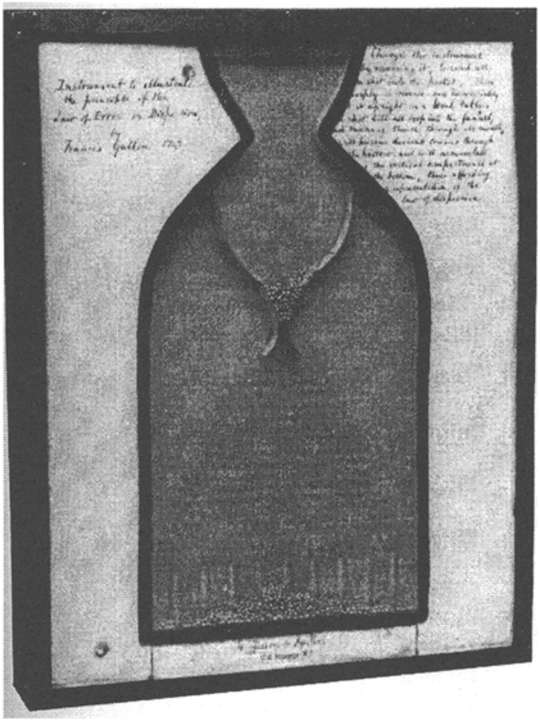

Figure 3:
The original quincunx. (Fig. from Stigler (1986, p. 277))

Galton was confronted with a problem: He could recognize the appearance of the normal distribution in his data, but he could not connect the curve with the heredity of characteristics from generation to generation. In some sense the classical theory of errors hindered him in his search for a connection. If the human characteristics follow a normal distribution, since they are the result of a large number of small influences, none of which is of outstanding importance, all of them together follow a normal distribution as well, how can a single factor - the parents - have a measurable influence? And why does the variation of the characteristics of the population not increase from generation to generation? Galton was able to find a way out of this dilemma with his formulation of regression and its connection with the bivariate normal distribution. By experimenting with peas he discovered the "tendency to the middle": Children of small parents will on the average be taller than their parents, children with very tall parents will on the average be smaller than their parents. In *Natural Inheritance* (Galton 1889) he gave a detailed description of this law of regression (at first called "law of reversion"). Above all, he examined the conditions which cause the law of errors. The classical conditions of Laplace, the independent and identical distribution of the random variables, do imply that the data follow a normal distribution, but this conditions were too restrictive for Galtons use. He wanted to show that the sufficient conditions of Laplace are not necessary for the normal distribution. Galton wrote:

> *"Considering the importance of the results which admit of being derived whenever the law of frequency of error can be shown to apply, I will give some reasons why its applicability is more general than might have been expected from the highly artificial hypotheses upon which the law is based. It will be remembered that these are to the effect that individual errors of observation, or individual differences in objects belonging to the same generic group, are entirely due to the aggregate action of variable influences in different combinations, and that these influences must be*
>
> - *all independent in their effects*
> - *all equal*
> - *all admitting of being treated as simple alternative ‚above average‘ or ‚below average‘ and*
> - *the usual Tables are calculated on the further supposition that the variable influences are infinitely numerous."*
>
> (Galton (1875, p. 38))

and continued:

> *"[The first three of the conditions] assuredly do not occur in vital or social phenomena, nevertheless it has been found in numerous instances, where measurement was possible, that the latter conform very fairly, within the limits of ordinary statistical inquiry, to calculations based on the (exponential) law of frequency of errors. It is a curious fact, which I shall endeavour to explain, that in this case a false hypothesis, which is undoubtedly a very convenient one to work upon, yields true results."*
> (Galton (1875, p. 39ff))

This argumentation was not wrong, but it did not really apply to the heart of the matter. It was decisive for Galton that the number of influences did not really have to be infinitely numerous. The article from 1875 contains a passage which is the key to the solution of the problem. Later Galton recognized that: E. g. the intensity of light is a decisive factor for the growth of fruit. If the same sort of fruit grows at different locations (e.g. at a southern, western and northern slope) on the average the fruit will be of different sizes - big, moderate and small. So why should the whole crop follow a normal distribution?

> *"The question is, why a mixture of series radically different, should in numerous cases give results apparently identical with those of a single series [...]. Now if it so happens that the "moderate" phase occurs approximately twice as often as either of the extreme phases (which is an exceedingly reasonable supposition, taking into account the combined effects of azimuth, altitude, and the minor influences relating to shade from leaves etc.) then the effect of aspect will work in with the rest, just like a binomial of two elements. Generally the coefficients of $(a+b)^n$ are the same as those of $(a+b)^{n-r} \cdot (a+b)^r$."*
> (Galton (1875, p. 45))

So the growth of the fruit can be traced back on two things: one main influence factor and other factors. Given the main influence factor the other factors follow the classical theory of errors. Let r be the number of interference factors, r can be very large. If the main influence factor can be seen as the sum of $n-r$ interferences, $(n-r)+r=n$ inferences result, although $n-r$ is small. In the end the interference distribution is a normal distribution, even if the main influence factor is not normally distributed. At this time Galton was aware of the fact, that this conclusion may not be drawn without further ado, but he was vague about this aspect and did not go into details. Later he changed his mind not at least caused by the quincunx.

1.3 The two-stage quincunx

Galton developed a second version of the quincunx, the two-stage quincunx. It seems, however, that this version has never been really built. Nevertheless it was an important aid for further discoveries. A letter written by Galton on January 12th, 1877 (see Fig. 4) is still existing. There he explained the two-stage quincunx to his cousin George Darwin.

The question is: What happens if the runs are interrupted at some position on the board and the balls are caught there in compartments? The distribution arising should have a shape which is similar to the normal distribution, but the dispersion should be smaller than the one the curve would have reached at the bottom of the board.

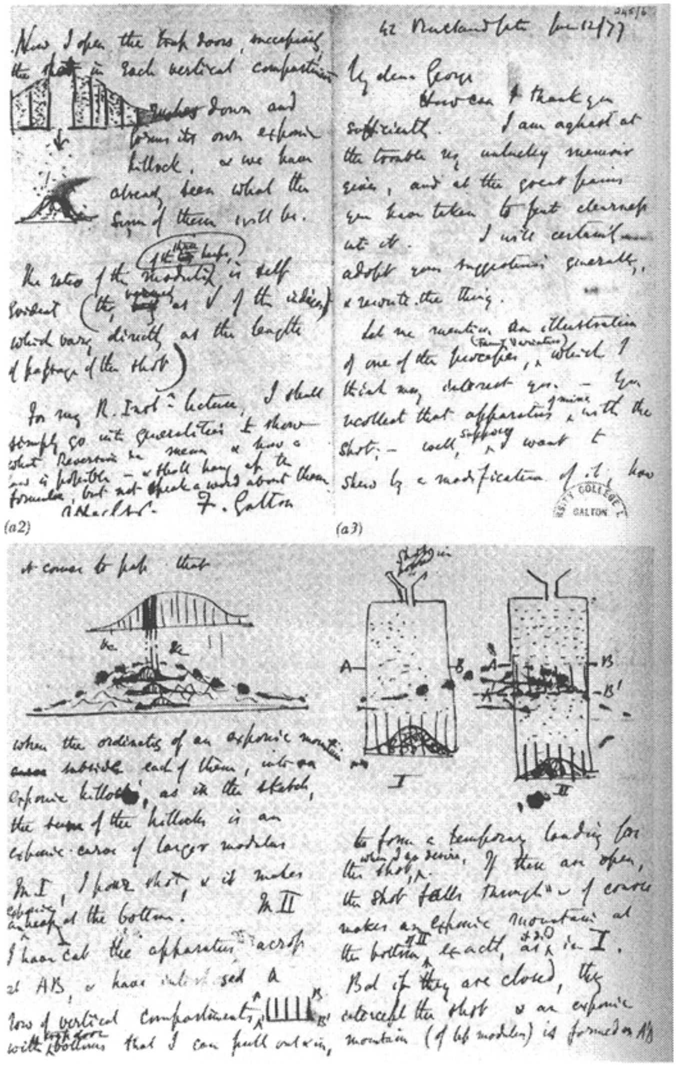

Figure 4:
Galton's letter, written on January 12th, 1877 to his cousin Georg Darwin. (Fig. from Stigler (1986, p. 278))

What happens then, if a single of the upper compartments is opened and the balls from this compartment run through the second part of the quincunx? In Galtons words a small "normal hillock" will be formed under the compartment concerned (see Fig. 5). And what, if all compartments are opened? Each one produces a small normal hillock, all together they build a hillock which can not be told apart from the hillock arising from a run without interruption.

Galton experimented with peas and understood, that a data set which is normally distributed (e. g. the weight of the peas) can be divided into smaller data sets, which are normally distributed as well. Each of the seven subgroups of his experiment was descended from parents having a certain weight; this can be understood as seven interim compartments on the two-stage quincunx (see Fig. 5, 6).

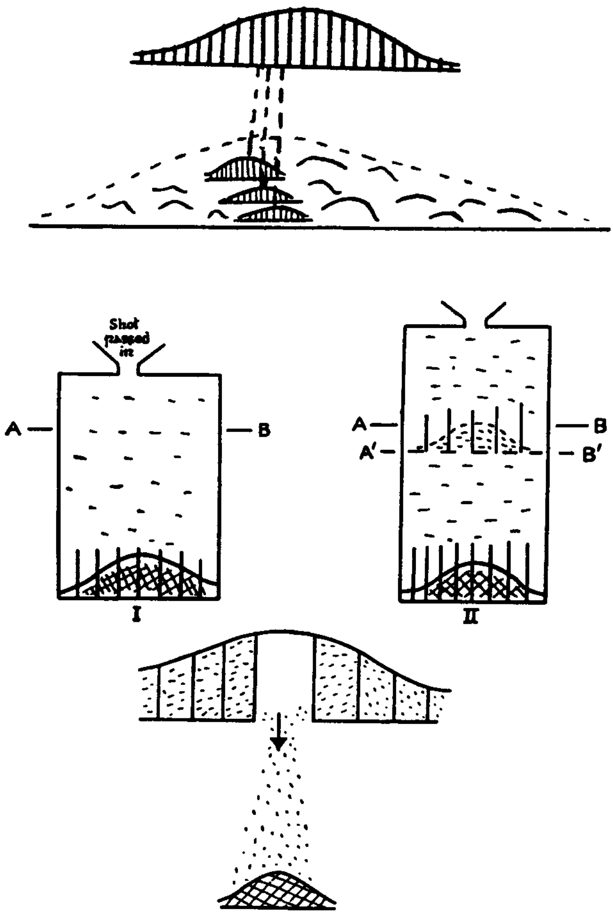

Figure 5:
Drawing by Pearson for his Galton biography (Pearson (1914-1930)), based on Galton's original letter. (Fig. from Stigler (1986, p. 279))

Galton asked why the width of the variation does not increase but remains constant. He came to realize that the reason is *regression*. The weights of the descendants are normally distributed, yet the mean is not the weight of the parents but the mean is closer to the mean of the total population.

Because of the regression the variance of the total population remains constant. In Fig. 6 Galton makes this visible: With the help of the narrowed channels the means of the groups are brought together just so far that the variance of the total population stays the same.

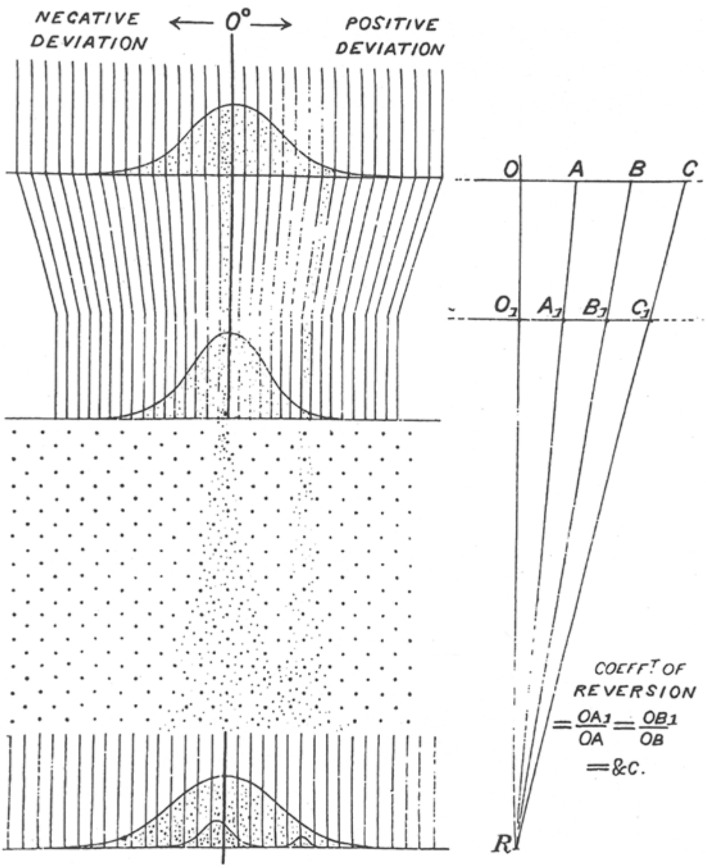

Figure 6:
Drawing by Galton: Two-stage quincunx and law of regression (Fig. from Galton (1877)).

However, from this point of view there is no obvious reason why this regression - which is necessary for the stability - should arise.

In 1885 Galton solved this problem. He collected data about the height of parents and adult children. Then he developed a formula to combine the height of both parents: He multiplied the mother's height with 1.08, added the father's height and divided by two. After that he examined the coherence between the height of the "midparents" and the children.

In contrast to the peas, which were a stratified sample, now an usual random sample is analysed. The height is divided into intervals, so the data given in the table (Fig. 7) can be regarded as grouped normally distributed data. The children's height is normally distributed, the sums of the columns result in the bottom of the quincunx. Each row contains the frequency distribution of a small normally distributed hillock, and the columns (height of the midparents) indicate the respective interim compartment.

Height of the mid-parent in inches	Height of the adult child														Total no. of adult children	Total no. of mid-parents	Medians
	<61.7	62.2	63.2	64.2	65.2	66.2	67.2	68.2	69.2	70.2	71.2	72.2	73.2	>73.7			
>73.0	—	—	—	—	—	—	—	—	—	—	—	1	3	—	4	5	—
72.5	—	—	—	—	—	—	—	1	2	1	2	7	2	4	19	6	72.2
71.5	—	—	—	—	1	3	4	3	5	10	4	9	2	2	43	11	69.9
70.5	1	—	1	—	1	1	3	12	18	14	7	4	3	3	68	22	69.5
69.5	—	—	1	16	4	17	27	20	33	25	20	11	4	5	183	41	68.9
68.5	1	—	7	11	16	25	31	34	48	21	18	4	3	—	219	49	68.2
67.5	—	3	5	14	15	36	38	28	38	19	11	4	—	—	211	33	67.6
66.5	—	3	3	5	2	17	17	14	13	4	—	—	—	—	78	20	67.2
65.5	1	—	9	5	7	11	11	7	7	5	2	1	—	—	66	12	66.7
64.5	1	1	4	4	1	5	5	—	2	—	—	—	—	—	23	5	65.8
<64.0	1	—	2	4	1	2	2	1	1	—	—	—	—	—	14	1	—
Totals	5	7	32	59	48	117	138	120	167	99	64	41	17	14	928	205	—
Medians	—	—	66.3	67.8	67.9	67.7	67.9	68.3	68.5	69.0	69.0	70.0	—	—	—	—	—

Source: Galton (1886a).

Figure 7:
Galton's table from 1885: the height of 928 adult children and the "middle height" of the 205 parents belonging to it. When publishing it again in 1889 Galton wrote that the first line is wrong, because four children can not have five parents but the last line (14 children from one family), which looks suspicious, is correct. (Fig. taken from Stigler (1986, p. 259))

Galton described the table from Fig. 7 as bivariate normal distribution. This freed him from the strict directional relation which is given by the two-stage quincunx. Now he saw a linear coherence between the height of the generations, which is disturbed by random influences on parents *and* descendants. Two related questions helped Galton to take this step:

• An anthropological question: If a single ancient bone is found, what does this tell about the height of the person it comes from?

• A question in forensic medicine concerning identification: What can be told about the relation between different measures of one person?

Galton connected this questions with the problems of heredity which he had already solved and recognized that these questions where special cases of a much more general question: the correlation. He developed a detailed statistical concept of correlation, which probably is the most famous of his legacies. As he said himself:

> *"Few intellectual pleasures are more keen than those enjoyed by a person, who, while he is occupied in some special inquire, suddenly perceives that it admits of a wide generalization and that his results hold good in previous unsuspected directions."* (quoted after Stigler (1989, p. 75))

Our description essentially follows Stigler (1986, 1989).

1.4 The Galton-Pearson-Board

In 1895 Karl Pearson published a modification of the quincunx (see Fig. 8). Here each movable row can produce a different probability for deviation. With this apparatus it should be possible to produce skewed distributions, which do not have to be binomial distributions (see Theorem 2.2).

During our investigation we found several older and newer drawings of Galton-Pearson-Boards, but - in contrast to many real quincunxes in all variations - we know of only one Galton-Pearson-Board existing in the Department of Statistics at Berkeley.

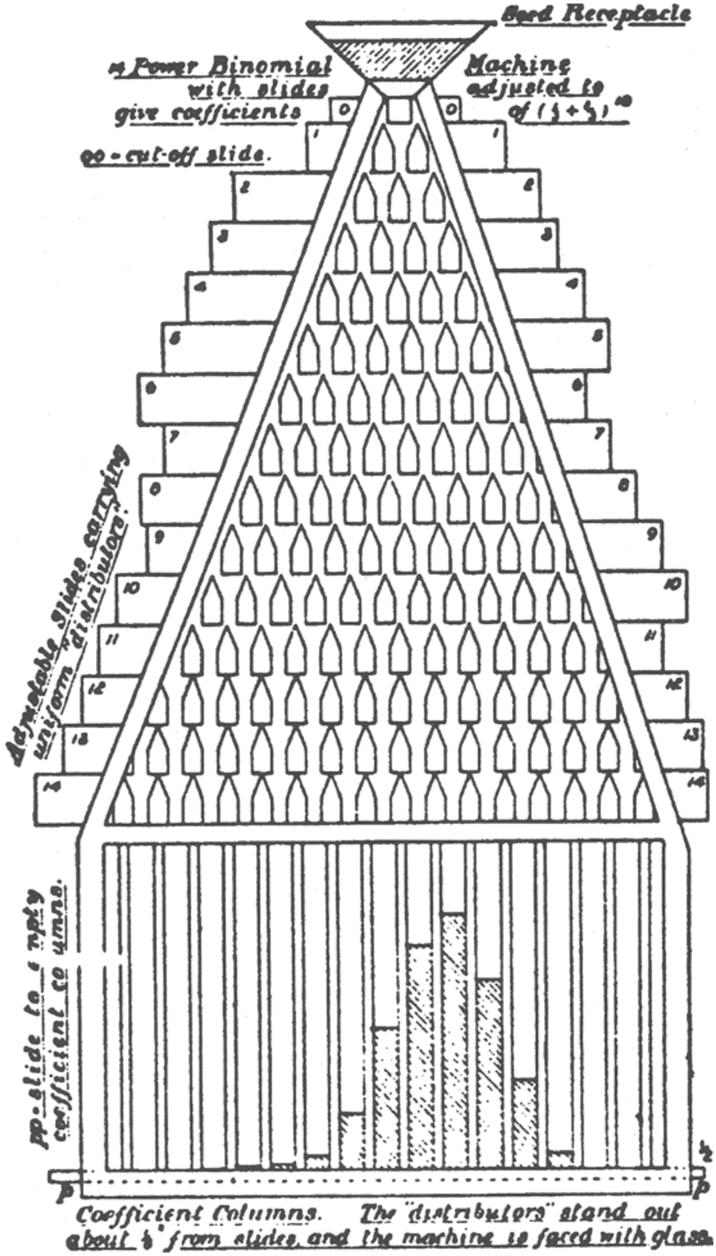

Figure 8:
Illustration by Pearson: Generation of skew distributions. (Fig. from Pearson (1895, p. 415))

2 Mathematical background

2.1 General overview

The quincunx is a nice means to show different statistical laws. The normal distribution, however, is not as evident as is often implied in the literature. There are some considerations necessary to see why an approximate normal shape appears in a large quincunx. We list the single steps in this section, which is followed by a more detailed mathematical derivation in section 2.2.

Binomial distribution. For one ball the probability to fall into a certain compartment is according to the binomial distribution.

Strong law of large numbers. If there are very many balls to roll down the quincunx, it follows from the law of large numbers that the proportion of balls in a given compartment converges to a constant, which depends on the compartment (see Fig. 9, see also Theorem 2.1).

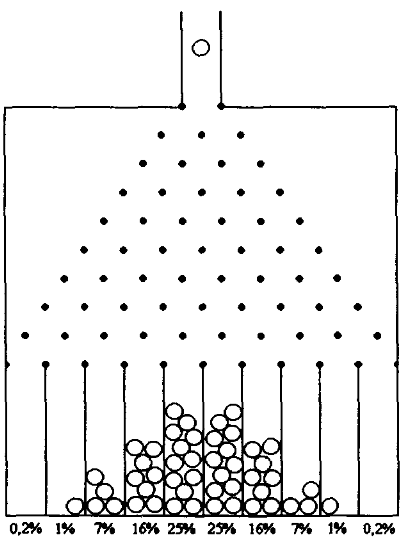

Figure 9:
Binomially distributed balls.

A series of boards: Imagine a series of quincunx boards getting larger and larger by adding new rows of nails. Let the number of balls be large but constant. Then the balls are distributed into more and more compartments, the number of balls in each compartment converges to zero.

Combination of compartments in a series of boards: Let us group the compartments of the above series of boards, such that the number f of compartments remains constant and all compartments have the same width. Then, as the number of rows increases, the proportion of balls to roll in the middle compartment increases (see the proportions in Fig. 10 in comparison with Fig. 9). Finally, with ever increasing number of rows, all balls go into the middle compartment.

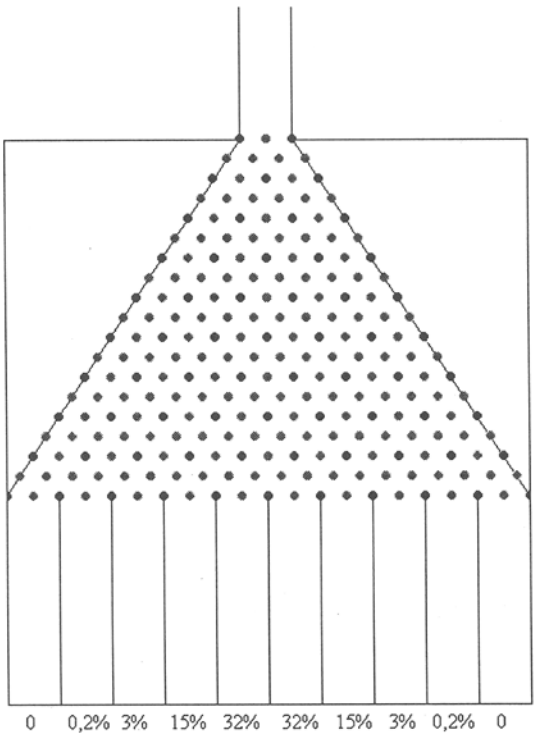

Figure 10:
Increasing number of nail rows.

Integral Limit Theorem Once more, let the number m of rows of nails increase. This time we let the number f of compartments also increase, but not as fast as m. We assume that f is given by the formula

$$f = \lambda\sqrt{m}, \quad \text{where } \lambda \text{ is a constant.}$$

Let x be a fixed integer. If the number n of balls is kept constant, then, with increasing m, the proportion of balls in compartment $[\frac{f+1}{2}] \pm x$ (that is the xth compartment seen from the middle) converges against a quantity given by the normal distribution. However, only if the number f of compartments is large (and, consequently, the number m of rows is very large) a bell curve results.

This is the Central Limit Theorem. In the case of the binomial distribution it is also called the Integral Limit Theorem (see Theorem 2.4).

Local Limit Theorem. Another possibility is not to group the compartments but to increase the number n of balls together with the number m of rows, following the formula

$$n = \sqrt{m}.$$

So we assume that the number of compartments is $m+1$. Let x be a given integer. Then the fraction of balls in compartment $[\frac{f+1}{2}] \pm x$ converges to a fixed number, which is proportional to $e^{-x^2/2}$. Here really a bell form arises. This is the Local Limit Theorem (see Theorem 2.3). If the number of balls is increasing at a faster rate than $\sqrt{m}$, then the percentage of balls in the xth compartment (seen from the middle) converges to infinity.

Conclusion: The form of a standard normal distribution is not really easy to see.

2.2 Mathematical models

A ball rolling down a quincunx as shown in Fig. 11 when passing through row i will either deviate to the right (with probability p_i) or to the left (with probability $1-p_i$). In all, n balls are used and are collected in $m+1$ compartments $F_k, k = 0, \ldots, m$.

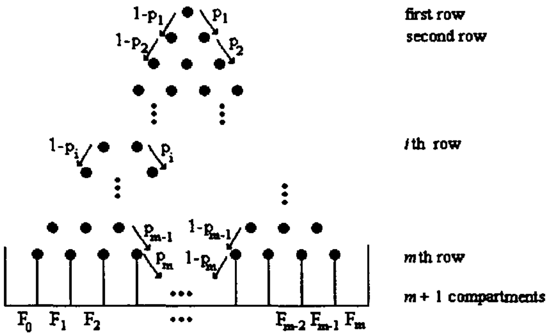

Figure 11:
Quincunx notations.

For ball j, $1 \leq j \leq n$, let the random variable Y_{ij} describe the behaviour of the ball at row i:

$$Y_{ij} = \begin{cases} 1 & , \text{ if the ball moves to the right} \\ 0 & , \text{ if the ball moves to the left.} \end{cases}$$

Then $P(Y_{ij} = 1) = p_i$ and $P(Y_{ij} = 0) = 1 - p_i$ and Y_{ij} corresponds to the outcome of a Bernoulli-experiment with probability of success of p_i. Note that the probability is the same for every ball j.

Now, if a ball j has completed a run, then the sum of the Y_{ij} determines the compartment in which it lands. More precisely, if $X_j^{(m)}$ = "number of the compartment in which ball j ends", then

$$X_j^{(m)} = \sum_{i=1}^{m} Y_{ij}.$$

If the ball always moves to the left, then it ends in compartment 0. If it moves to the right exactly k times, then it ends in compartment k.

This identity will be used later on to determine the probability distribution of $X_j^{(m)}$ for the standard quincunx (all $p_i = \frac{1}{2}$).

The strong law of large numbers. The first law to be seen from the quincunx is the strong law of large numbers. (Some considerations can be found in v. Mises (1931, pp. 144-146))

Theorem 2.1 (Strong law of large numbers)
For a sequence of i.i.d. random numbers X_i, $i = 1, 2, \ldots, n$ with expectation $E(X_i) = \mu$, the empirical mean converges a.s. to the expectation μ. More formally

$$P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mu\right) = 1.$$

Proof: see e. g. Gnedenko (1968, p. 253ff) □

To apply Theorem 2.1 let us introduce a third random variable $Z_j^{(m,k)}$ which indicates whether the ball j falls into compartment k, or not. That is

$$Z_j^{(m,k)} = \begin{cases} 1, & if\ X_j^{(m)} = k \\ 0, & else. \end{cases}$$

Then $\frac{1}{n}\sum_{j=1}^{n} Z_j^{(m,k)}$ is the proportion of balls to go into the kth compartment. Let $Z_j^{(m,k)}$ take the role of the X_i in Theorem 2.1.

Since the probabilities p_i, $1 \leq i \leq m$, are the same for every ball, we have for every ball the same probability $\pi_k^{(m)}$ to go into compartment k. Therefore, for every j we have that $P(Z_j^{(m,k)} = 1) = \pi_k^{(m)}$ is the same.

Then the strong law of large numbers implies that

$$\frac{1}{n}\sum_{j=1}^{n} Z_j^{(m,k)} \overset{\text{a.s.}}{\to} \pi_k^{(m)} \text{ as } n \to \infty$$

Therefore, by varying k, the (theoretical) distribution of the random variable $X_j^{(m)}$ can be approximated by the observed proportion of balls in the different compartments, provided a sufficiently large number of balls is used.

The standard quincunx. If all $p_i = p$ (equal probabilities at each row) then it follows from the remarks above, that

$$P(X_j^{(m)} = k) = P\left(\sum_{i=1}^{m} Y_{ij} = k\right) = \binom{m}{k} p^k (1-p)^{m-k},$$

since there are exactly $\binom{m}{k}$ different ways to fall through the quincunx, where the ball turns to the right exactly k times. The probability to turn to the right exactly k times (and to the left exactly $(m-k)$ times) is $p^k(1-p)^{m-k}$. Therefore $X_j^{(m)}$ is binomially distributed with parameters m and p,

$$X_j^{(m)} \sim B(m,p).$$

For the standard quincunx the probability for ball j to turn to the left or to the right is equal, $p = \frac{1}{2}$, at each row. Consequently, $X_j^{(m)} \sim B(m, \frac{1}{2})$.

Galton-Pearson-Board. By the movable rows of the Galton-Pearson-Board different probabilities p_i can be derived for each row i. This will lead to a distribution of the $X_j^{(m)}$ which need not be a binomial distribution. In fact, we now show that the distribution of the $X_j^{(m)}$ is a binomial distribution if and only if all p_i are the same.

Theorem 2.2
It holds:

$$\exists\, p \text{ such that } X_j^{(m)} \sim B(m,p) \Leftrightarrow p_i = p \;\forall\, i = 1,\dots,m. \tag{1}$$

Proof: The direction " $\Leftarrow$" has been shown in the last paragraph. The other direction can be seen, if it is assumed that

$$X_j^{(m)} \sim B(m,p).$$

It follows that for every k

$$P\left(\sum_{i=1}^{m} Y_{ij} = k\right) = P(X_j^{(m)} = k) = \binom{m}{k} p^k (1-p)^{n-k} \tag{2}$$

and especially for $k = m$, that

$$p^m = P\left(\sum_{i=1}^{m} Y_{ij} = m\right).$$

Since

$$P\left(\sum_{i=1}^{m} Y_{ij} = m\right) = P(Y_{ij} = 1 \;\forall\, i = 1,\dots,m) = \prod_{i=1}^{m} p_i,$$

it follows that

$$p = \left(\prod_{i=1}^{m} p_i \right)^{1/m},$$

the geometric mean of the p_i. It is well-known that the geometric mean is less or equal to the arithmetic mean

$$p = \left(\prod_{i=1}^{m} p_i \right)^{1/m} \leq \frac{1}{m} \sum_{i=1}^{m} p_i$$

(see e. g. Kendall and Stuart (1969, p. 36)).

Therefore,

$$p \leq \frac{1}{m} \sum_{i=1}^{m} p_i. \tag{3}$$

On the other hand, if $k = 0$ in (2), it follows with the same arguments that

$$1 - p = \left(\prod_{i=1}^{m} (1 - p_i) \right)^{1/m}$$

and that

$$1 - p \leq \frac{1}{m} \sum_{i=1}^{m} (1 - p_i) = 1 - \frac{1}{m} \sum_{i=1}^{m} p_i. \tag{4}$$

Consequently, combining (3) and (4),

$$p = \frac{1}{m} \sum_{i=1}^{m} p_i,$$

and the geometric and arithmetic mean coincide. Therefore it follows that all p_i are equal. □

Drift. It appears plausible that a ball which has moved to the right in one level has a higher probability than $\frac{1}{2}$ to move to the right again in the next level. Such a phenomenon would destroy the assumption of independence and therefore the binomial distribution. (It is one of the main difficulties in constructing a quincunx, to try to avoid this dependence to happen.) Some theory to this can be found in v. Mises (1964, pp. 288-289).

2.3 Convergence in distribution

The approximation of the binomial distribution by the normal distribution can not easily be shown by the quincunx (see section refmathall). Mathematically it is based on the limit theorems of de Moivre-Laplace which describe the local and the global approximation of the binomial distribution by the normal distribution. Some similar considerations were done by v. Mises (1931, pp. 144-146).

Theorem 2.3 (Local Limit Theorem)
Consider a sequence $(X^{(m)})$ of random variables where $X^{(m)} \sim B(m,p)$ and let (k_m) be a sequence of numbers in $\mathbb{N}_0$ with

$$\lim_{m\to\infty} \frac{k_m - mp}{\sqrt{mp(1-p)}} = x.$$

Then we have for $m \to \infty$

$$P\left(X^{(m)} = k_m\right) \approx \frac{1}{\sqrt{(m+1)p(1-p)}}\,\varphi(x), \tag{5}$$

in the sense that the ratio of the l.h.s. and r.h.s. converges to 1. Here $\varphi(x)$ is the density function of the standard normal distribution.

Proof: see Gnedenko (1968, p. 94ff) □

Theorem 2.4 (Integral Limit Theorem)
Let $\Phi(x)$ be the distribution function of the standard normal distribution. If, additionally to the conditions of Theorem 2.3, the sequence of numbers (l_m) in $\mathbb{N}_0$ fulfills

$$\lim_{m\to\infty} \frac{l_m - mp}{\sqrt{mp(1-p)}} = y > x$$

then

$$\lim_{m\to\infty} P\left(k_m \le X^{(m)} < l_m\right) = \Phi(y) - \Phi(x).$$

Proof: see Gnedenko (1968, p. 104ff) □

For the quincunx this means that the probability for one *single* compartment

$$P\left(X_j^{(m)} = k\right)$$

as well as the probability for *several* compartments

$$P\left(X_j^{(m)} \in \{a, a+1, \dots, a+k\}\right),$$

$$\text{with} \quad \{a, a+1, \dots, a+k\} \subset \{0, \dots, m\}$$

can be approximated with the help of the normal distribution. It is decisive, however, that *the goodness of this approximation can only be improved, if the number m of rows increases.*

In other words: A given quincunx (m fixed) will yield an increasingly better fit to the binomial distribution, if the number n of balls increases, but will not take a "smoother" shape in sense of a bell curve.

Galton-Pearson-Board. If different p_i are allowed, then a more general theorem is needed: the Theorem of Lindeberg-Feller.

Theorem 2.5 (Lindeberg-Feller)
Let $X^{(m)}$ be as in Theorem 2.1. Then

$$X^{(m)} \overset{\text{asy}}{\sim} \mathcal{N}\left(\sum_{i=1}^{m} p_i, \sum_{i=1}^{m} p_i(1-p_i)\right)$$

$$\Leftrightarrow \sum_{i=1}^{\infty} p_i(1-p_i) \qquad \textit{diverges.} \tag{6}$$

Proof: see e.g. Gnedenko (1968, p. 318). □

The notation $X^{(m)} \overset{asy}{\sim} \mathcal{N}(\xi_m, \delta_m^2)$ means that the distribution of

$$\frac{X^{(m)} - \xi_m}{\delta_m}$$

converges weakly to the $\mathcal{N}(0,1)$, where ξ_m and δ_m are sequences of constants. It is not necessary that ξ_m and δ_m^2 are the mean and the variance of $X^{(m)}$, see e.g. Serfling (1980, p. 20).

So the asymptotic behaviour of the quincunx depends on the structure of the deviation probabilities p_i. In the special instance that all p_i are equal to a number p, we have

$$S_m = \sum_{i=1}^{m} p_i(1-p_i) = mp(1-p) \underset{m\to\infty}{\longrightarrow} \infty,$$

so that condition (6) is fulfilled. Therefore the Integral Limit Theorem 2.4 is a special instance of Theorem 2.5.

It is clear that, if there is an m_o such that $p_i \equiv 0$ for all $i \geq m_o$, then $\sum_{i=1}^{m} p_i(1-p_i) = \sum_{i=1}^{m_o} p_i(1-p_i)$ for all $m \geq m_o$. It then also holds that $X^{(m)}$ is not asymptotically normal as all $X_j^{(m)}$ for $m \geq m_o$ have the same distribution as $X_j^{(m_o)}$.

There are also situations when all $p_i \in (0,1)$, but $\sum_{i=1}^{\infty} p_i(1-p_i)$ converges and therefore $X_j^{(m)}$ is not asymptotically normal.

As an example for that consider $p_i = \frac{1}{2^i}, i \in \mathbb{N}$. Then for every m we have

$$EX_j^{(m)} = \sum_{i=1}^{m} p_i = \frac{1}{2}\sum_{i=0}^{m-1} \frac{1}{2^i} = \frac{1}{2} \cdot \frac{1-(\frac{1}{2})^m}{1-\frac{1}{2}} = 1-(\frac{1}{2})^m < 1$$

and

$$VarX_j^{(m)} = \sum_{i=1}^{m} p_i(1-p_i) = \sum_{i=1}^{m}(\frac{1}{2})^i - \sum_{i=1}^{m}(\frac{1}{4})^i$$

$$< \sum_{i=1}^{m}(\frac{1}{2})^i = 1-(\frac{1}{2})^m < 1.$$

Remember that $X_j^{(m)}$ is a random variable with nonnegative integer values. Therefore,

$$P(X_j^{(m)} \in \{0,1,2,3\}) = P((X_j^{(m)} \leq 3) =$$

$$P(\frac{X_j^{(m)} - EX_j^{(m)}}{\sqrt{VarX_j^{(m)}}} \leq \frac{3-EX_j^{(m)}}{\sqrt{VarX_j^{(m)}}})$$

$$\geq P(\frac{X_j^{(m)} - EX_j^{(m)}}{\sqrt{VarX_j^{(m)}}} \leq \frac{3-1}{1})$$

$$\geq P(|\frac{X_j^{(m)} - EX_j^{(m)}}{\sqrt{VarX_j^{(m)}}}| \leq 2).$$

By Tchebycheff's inequality (see e. g. Kendall and Stuart (1969, p. 88)), we conclude that $P(X_j^{(m)} \in \{0, 1, 2, 3\}) \geq \frac{3}{4}$.

Therefore, the fixed finite set {0,1,2,3} has a probability of more than $\frac{3}{4}$ for every m. This implies that $X_j^{(m)}$ is not asymptotically normal.

3 The quincunx in practice

For didactic reasons diagrams of quincunxes can be found in popular and scientific statistical books and articles, in textbooks of statistics for the humanities and economics, for engineers and natural scientists as well as in school books for middle and upper schools, sometimes even for primary schools. As audiovisual aids some schools and universities have a self- or professionally-made quincunx.

Not all of the existing quincunxes do work properly. On the contrary, it is quite difficult to produce a good quincunx. The nails have to be driven in exactly, an additional row of nails at the margins is necessary. The balls must be of equal size and weight, they must be smaller than the space between the nails, but they may not be "too" small. If the quincunx is closed, the funnel must be big enough to contain all balls.

As simple as a drawing of a quincunx seems to be, there are many inaccuracies and mistakes in illustrations of the quincunx published in the literature. Especially dubious in some publications is the emphasis of the normal distribution under neglect of the binomial distribution, even though a binomial distribution actually is produced. For more details about this and for a detailed bibliography see Kunert, Montag, Pöhlmann (2000).

References

Darwin, Charles (1859): *The Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life.* New York Library, New York.

David, F. N. (1987): Galton, Francis. In: Kruskal, William H. and Judith M. Tanur (eds.): *International Encyclopedia of Statistics* (vol. 1). The Free Press - Macmillan, New York.

Galton, Francis (1869): *Hereditary Genius: An Inquiry into its Laws and Consequences.* Macmillan, London.

Galton, Francis (1874): On a proposed statistical scale. *Nature*, 9, 342-343.

Galton, Francis (1875): Statistics by intercomparison, with remarks on the law of frequency of errors. *Philosophical Magazine, 4th series*, 49, 33-46.

Galton, Francis (1877): Typical laws of heredity. *Proceedings of the Royal Institution of Great Britain*, 8, 282-301.

Galton, Francis (1889): *Natural Inheritance.* Macmillan, London.

Galton, Francis (1908): *Memories of my life.* Macmillan, London.

Gnedenko, B. V. (1968): *The Theory of Probability*, 4. edn. Chelsea Publishing Company, New York.

Kendall, Maurice and Alan Stuart (1969): *The Advanced Theory of Statistics, vol. 1: Distribution Theory*, 3. edn. Charles Griffin, London.

Kunert, Joachim, Astrid Montag and Sigrid Pöhlmann (2000): Das Galtonbrett und die Glockenkurve. In: Link, Jürgen und Ute Gerhard (eds.): *Flexibler Normalismus und Orientierung mittels 'Kurvenlandschaften': Fälle und Modelle* (working title). Synchron Wissenschaftsverlag der Autoren, Heidelberg. (in preparation)

Mises, Richard von (1931): Vorlesungen aus dem Gebiet der angewandten Mathematik, I. Band: Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und Theoretischen Physik. Deuticke, Leipzig.

Mises, Richard von (1964): Mathematical Theory of Probability and Statistics (Ed. by H. Geiringer). Academic Press, New York.

Pearson, Karl (1895): Contributions to the Mathematical Theory of Evolution II. *Philosophical Transaction of the Royal Society*, A186. Royal Society, London, 343-414 and 7 plates.

Pearson, Karl (1914-1930): *The Life, Letters and Labours of Francis*

Galton (3 vols. in 4 parts). Cambridge University Press, Cambridge.

Posten, Harry O. (1986): Quincunx. In: Kotz, S. and N. Johnson (eds.): *Encyclopedia of Statistical Science* (vol. 7). Wiley, New York.

Serfling, Robert (1980): *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

Stigler, Stephen (1986): *The History of Statistics.* The Belknap Press of Harvard University Press, Cambridge.

Stigler, Stephen (1989): Francis Galtons Account of the Invention of Correlation. *Statistical Science*, 4/2. Institute of Mathematical Statistics, Hayward, California, 73-86.