Table 7 (variance test for homogeneity for distribution truncated below one). Statistic $= \sum x_i^2$. $n = 3(1)9, n\bar{x} = 8(1)12$. An earlier condensed version of these tables is given by Rao and Chakravarti (1956).

Kathirgamatamby (1953) gives some tables for testing the sample index of dispersion.

Kiefer and Wolfowitz (1956) give tables for evaluation of the operating characteristic function and average sample time function for sequential probability ratio testing of λ.

Ractliffe (1964) gives experimentally obtained values of the cumulative frequency of (6.4-1) for $\lambda = 5(5)15, 30, 55, 80, 100, 130$ and $u = 0(0.25)4.25$, a statistic useful for testing the significance of the difference between two Poisson variables.

Chapter 9

HISTORICAL REMARKS

9.1 SIMEON DENIS POISSON (1781–1840)

The eminent scientist S. D. Poisson contributed significantly to many branches of mathematics and physics, and held important posts in the scientific and academic establishment of France. During the later years of his life, he became interested in the application of mathematical probability to the administration of justice. His major work in this area, the famous *Recherches sur la Probabilité des Jugements* [Poisson (1837*)], although designed as a concrete contribution to juridical practice, contains so much preliminary material of a purely mathematical and probabilistic nature that it must be regarded as a textbook on probability with illustrations from the courts of law. The German edition [Poisson (1841)], which differs but slightly from the original (and not at all in the section containing the Poisson distribution) bears a much more accurate title: *Lehrbuch der Wahrscheinlichkeitsrechnung und deren wichtigsten Anwendungen.*

In the course of his mathematical development, Poisson (Section 81; pp. 205–207 of the French edition) obtains $p_x(\lambda)$ as a binomial limit in a style very like that used in modern textbooks. The value p_0 is singled out for special mention, and the cumulative form is observed to converge to unity. Poisson then turned to other matters, and does not refer again to his discovery, either in the remainder of the book or in his subsequent publications.

Newbold (1927), Jensen (1954), and David (1962) prefer to give the credit for the discovery to de Moivre (1718), and David quotes the relevant passage. Be that as it may, it would certainly appear that the first person to grasp the statistical significance of the Poisson formula was Bortkiewicz, whose contributions are mentioned in Section 9.3.

* Certain authors give the date 1832 for Poisson's *Recherches*; the *Catalogue Generale des Livres Imprimes de la Bibliotheque Nationale* (Vol. 139, p. 982) gives only 1837, and the volume dated 1837 does not refer to any earlier edition.

9.2 FROM POISSON TO BORTKIEWICZ

There are almost no references to the Poisson distribution between 1841 and 1898. Poisson's name occurs, to be sure, in the statistical works of the nineteenth century, but in the context of "the law of large numbers" a name which Poisson used in connection with the Gauss-Laplace normal distribution. The splendor of normal theory dazzled scientific writers of the period and led many of them to believe that it was a "universal" law, governing every possible kind of variation.

Two very surprising exceptions appeared near the middle of the period. Seidel (1876) gives an explicit statement of the Poisson probabilities, obtained as a binomial limit (in the context of counting events) but without mentioning Poisson. Abbe (1879) also mentions the Poisson probabilities (in a paper dealing with blood corpuscle counting), but passes quickly on to a conventional application of the normal distribution, and also does not cite Poisson.

In addition to these unquestioned references to the Poisson probabilities, we might mention the theorem of Boltzmann (1868) to the effect that the probability of no (randomly placed) points in an interval of length t is $\exp(-\lambda t)$ as constituting a discovery of p_0. Whitworth (1886) gives this result as his Proposition LI.

In spite of these very meager references, there is some evidence that the Poisson limit of the binomial was known to a few mathematicians of the nineteenth century. Bortkiewicz (1915b) states that Poisson's work was called to his attention by Tschuprow in the mid-nineties. Kruskal* writes: "I am quite sure that the French probabilists, notably Bienaymé, working at the same time and after Poisson, were well aware of the limit. In Cournot's *Exposition de la Theorie des Chances* I notice a footnote on p. 331 that clearly ascribes to Poisson an approximation very closely related to the usual Poisson limit." On the other hand, Czuber's long monograph (1899), treating all aspects of probability theory, never mentions the Poisson distribution, although the name of Poisson occurs 28 times and that of Bortkiewicz five times.

9.3 LADISLAUS VON BORTKIEWICZ (1868–1931)

One of the important scientific discoveries of the nineteenth century was the relationship between probability and statistics. This connection, which today appears so obvious that teachers often find difficulty in explaining the distinction to students, was quite obscure one hundred years ago. Probability was thought of more or less as we now think of

* Private communication (paraphrased).

combinatorial mathematics; and statistics was regarded as the layman regards the word—substantial columns of figures in national yearbooks. In fact, the spate of papers purporting to show a connection between probability and statistics continued well into the twentieth century, as exemplified by Edgeworth's (1913) paper and his article on probability in the Eleventh Edition of the *Encyclopedia Britannica.* The mathematical reviewing journal *Jahrbuch über die Fortschritte der Mathematik* changed from the section heading *Kombinationslehre und Wahrscheinlichkeitsrechnung* to the more modern *Wahrscheinlichkeitsrechnung nebst Gleichungsrechnung. Statistik. Versicherungswesen* as late as 1916.

The discovery of this connection between a theoretical mechanism and its numerical manifestation was first observed in the normal distribution, an achievement associated with several of the most illustrious names in mathematics: Gauss, Laplace, Poisson, Lexis, Quetelet, Czuber. It was Bortkiewicz who first recognized that a similar connection exists between Poisson's formula and certain kinds of discrete data. Furthermore, he worked out in detail many of the theoretical and practical consequences of his discovery. It can be said, therefore, that although Poisson (or de Moivre) discovered the mathematical expression (1.1-1), Bortkiewicz discovered the probability distribution (1.1-1). It would also be fair to claim that the Poisson distribution is second in importance to the normal, whether regarded from the point of view of abstract theory or judged by its breadth of application. Furthermore, the contribution of Bortkiewicz is enhanced in importance by the fact that it came about quite abruptly as the work of a single individual.

Ladislaus von Bortkiewicz, best known today for his example of the frequency of death by horse-kick in the Prussian Army, was born in St. Petersburg of Polish ancestry. After studying law and public administration in Russia, he was sent by the government to Germany for graduate work. Bortkiewicz received his Ph.D. in 1892 from the University of Göttingen, where he was a student of Lexis. His dissertation [Bortkiewicz (1893)] contains an application of differential equations to stochastic quantities, but does not mention Poisson or the Poisson distribution. In an encyclopedia article published the following year [Bortkiewicz (1894)] Poisson appears fleetingly, but not the Poisson distribution; the principal emphasis is on a characteristically nineteenth century approach to normality.

After spending a few years in Strassburg Bortkiewicz returned to Russia as a civil servant. In 1901, however, he settled permanently at the University of Berlin, remaining a professor of *Staatswissenschaft* until his death in 1931.

It was during his stay in Strassburg as *privatdozent* that Bortkiewicz wrote his first, most important, and most famous contribution to the Poisson distribution—a monograph with the title *Das Gesetz der Kleinen Zahlen* [Bortkiewicz (1898)]. Coming after a half century of nearly perfect vacuum, this small pamphlet contains the binomial Poisson limit, with credit to Poisson; the mean, variance, mean deviation, and other moments; differential and difference relations for the probabilities; the normal limit of the Poisson; a form of the Poisson-gamma relationship; comparison with suicide and accident data (including the notorious horse kicks); and four-place tables of $p_x(\lambda)$ for $\lambda = 0.1(0.1)10.0$. Although [as pointed out by Kruskal (1956)] the main purpose of Bortkiewicz in *Das Gesetz der Kleinen Zahlen* was somewhat different—it involved the then popular Lexis ratio—the monograph revived and in a sense introduced the Poisson distribution as a statistical concept.

In addition to *Das Gesetz der Kleinen Zahlen*, Bortkiewicz wrote a number of other works dealing with the Poisson distribution, which will be reviewed in their historical context in subsequent sections. Bortkiewicz' other mathematical works contain contributions to actuarial theory (1899a, 1906), to the Pearson Type III distribution (1922), to the foundations of probability (1899b, 1903), to Tchebycheff's inequality (1927), another encyclopedia article (1901) (which, by the way, coming three years after *Das Gesetz der Kleinen Zahlen*, barely mentions it in a footnote and does not discuss the Poisson distribution), an idea presaging nonparametric analysis (1917), and other statistical questions (1918, 1920). In his later years Bortkiewicz turned to problems of national economy but still published occasional statistical papers. In fact, his last published work [Bortkiewicz (1931)] appeared in the *Annals of Mathematical Statistics*.

A biographical note on Bortkiewicz, with a selected list of publications and reference to obituaries, is given in J. C. Poggendorff's *Handworterbuch*, Vol. VI (1923–1931), Verlag Chemie, Berlin, 1963. A similar notice appears in Vol. 2 (1955) of *Neue Deutsche Biographie*, Dunker and Humlot, Berlin. A much longer biographical article written by Thor Andersson and a complete publication list appear in Vol. 10 (1931) of *Nordisk Statistisk Tidskrift*, pp. 1–16. The English translation of Andersson's appreciation is contained in the English counterpart journal, *Nordic Statistical Journal*, Vol. 3 (1931), pp. 9–26. Bortkiewicz was closely associated with the establishment of these journals, and Vol. 1 of the *Nordic* (i.e., Vol. 8 of the *Nordisk*) is dedicated to him, and contains his portrait. Neyman (1933) provides an account of Bortkiewicz' contributions to the development of the Poisson distribution. Gumbel (1931), in an obituary and appreciation, also discusses the importance of Bortkiewicz' statistical work and gives a complete publication list.

Those who knew him agreed that Bortkiewicz was an academician of the old school, authoritative and, if necessary, disputatious, but objective, dedicated to scholarship, and unceasingly industrious.

9.4 FROM BORTKIEWICZ TO CHARLIER (1898–1906)

The Poisson distribution was evidently of some interest to Lexis, for it is the subject of a dissertation by his student Schmidt (1900). Most of Schmidt's work consists in applying the Poisson probabilities to specific demographic data; there is very little analysis. At about the same time Loria (1900) introduced the Poisson distribution to Italian scholars, the consequences of which will be discussed in Section 9.5. Loria himself does not appear to have been very interested in the matter; he contents himself with a brief remark on *Das Gesetz* and the law it proposes.

There were during this period two independent* discoveries of the Poisson distribution, one by Smoluchowski and the other by "Student." Smoluchowski (1904), dealing with the distribution of gas molecules, obtains the Poisson probabilities as the limit of the binomial, and then passes on to the normal approximation. "Student's" work arose from problems of particle counting in the plane, and is fairly systematic. He begins ["Student" (1909)] with the binomial limit, obtains the Poisson probabilities, computes the mean and variance, together with the first six moments and the values of the Pearson betas. He then fits the distribution to several samples of data. It is quite clear that "Student" understood the significance of (1.1-1) in the modern sense of a discrete probability distribution.

The most important consequence of Bortkiewicz' work during the first decade of this century was represented in three papers and a monograph of Charlier. In the first paper Charlier (1905a) refers to his earlier work on the normal distribution, and emphasizes that the "Gaussian law of error" is not always applicable. The Poisson distribution (with reference to Bortkiewicz) is introduced as an alternative. In his second paper Charlier (1905b) deals at great length with representation of functions by normal and Poisson probabilities, and obtains results (Section 4.5) now associated with his name. The third paper (in English) [Charlier (1906)] covers more or less the same ground as the second with the addition of several examples. The expressions "Type A" and "Type B" for normal and

* Here, and in what follows, "independent" discoveries means those with no reference to previous works on the Poisson distribution. It is worth noting, however, that without exception all of these authors give the probabilities exactly in the Poisson-Bortkiewicz format—for example, $e^{-\lambda}$ in the numerator rather than e^{λ} in the denominator.

Poisson series respectively are introduced here; the main point of the paper appears to be the reconciliation of the Poisson theory and practice with the Pearson system of frequency curves. Charlier's approach is again put forward in the monograph [Charlier (1910)], which has been reprinted (undated) fairly recently, from the second (1920) edition.

9.5 THE GINI CONTROVERSY (1907–1910)

One of the earliest contributions of Corrado Gini to mathematical statistics was an unjustified attack on Bortkiewicz [Gini (1907, 1908a, 1908b)] and on the so-called "law of small numbers." The matter apparently came to Gini's attention by means of a small note by Broggi (1907), quoting *Das Gesetz* in connection with the Lexian ratio.

It is difficult at a distance of half a century, without a fluent knowledge of Italian and German, to be sure exactly what issues were involved. It would appear, however, that Gini believed that the Poisson distribution, as based on the Lexis ratio by Bortkiewicz, was incorrect. In his first paper he concludes with the dogma: "The law of small numbers does not exist!" He furthermore believed either that constant binomial probabilities were necessary for the binomial approach to Poisson to be justified, or else possibly that no one had proved the contrary. This matter was settled by Bresciani (1908) who showed that fixed binomial probability is by no means necessary for the Poisson limit to be valid.

Another difficulty arose in connection with Bortkiewicz' use of the expression "law of small numbers." Nowadays this phrase is generally considered to be synonymous with the Poisson distribution, but several authors [Newbold (1927), Crathorne (1928), Winsor (1947), Lorenz (1951), and Kruskal (1956)] make it perfectly clear that Bortkiewicz had something different in mind. According to Kruskal (1956), the law means that "in binomial sampling with large sample size and small values of p_i, the data tend to look as if the p_i's were equal, and, as a secondary point, Poisson." In other words, the law of small numbers was intended to describe not the Poisson distribution itself, but the tendency of certain data to be Poisson distributed.

The following passage from *Das Gesetz* [quoted by Winsor (1947)] illustrates partly Bortkiewicz' views and partly the modern reader's difficulty in deciding what his views were:

"Under the condition of a restricted field of observation one obtains, as we know, a nearly normal dispersion, that is, almost complete agreement between the standard errors calculated by the direct and indirect methods. The smaller the field of observation, and the more rarely the phenomenon in question, as for instance suicide or accident, occurs in a given community, the better do the statistical results fit the appropriate mathematical formula. The hypothesis of a variable probability or variable expectation helps us to recognize this behavior as lawful, and in this sense we may call the fact, that small numbers of events (out of very large populations) are subject to or tend toward a definite norm of variation, the *Law of Small Numbers*."

Although, in view of these circumstances it may not be correct to say that the law of small numbers *is* the Poisson distribution, it would be right to say that the law of small numbers is *about* the Poisson distribution, just as Poisson's law of large numbers is about the normal distribution. It would surely appear that Bortkiewicz' choice of the name was influenced by Poisson's expression law of large numbers. Furthermore, the sense in which normal and Poisson are the fundamental continuous and discrete distributions (illustrated by the Lévy-Khintchine Theorem of Section 3.2 and Charlier's Types A and B series) does, in a way, correspond to the idea of large and small numbers.

The Lexis ratio is explained by von Mises (1964); its relationship to the Poisson distribution and Bortkiewicz' law of small numbers, by Kruskal (1956) and Crathorne (1928).

Gini's criticisms were answered by Bresciani (1908) and by Bortkiewicz (1908, 1909, 1910). In his usual style Bortkiewicz discusses the problem at great length: historically, numerically, mathematically, and philosophically, from the point of view of the personalities involved, not omitting some polemics. He was correct in all essential aspects of the controversy, but conducted the discussion on the somewhat irrelevant grounds of the Lexis ratio and its implications. It should be remembered, however, that good criteria for accepting models as adequate to explain observations were largely lacking, and consequently heavy emphasis was often placed on certain favorite statistics. Bortkiewicz' point of view in the Gini controversy is explained by Edgeworth (1909).

9.6 PARTICLE COUNTING; DIFFERENTIAL-DIFFERENCE EQUATIONS (1910–1913)

While Bortkiewicz was wrangling with Gini about the Poisson as a binomial limit, a new justification for the Poisson distribution appeared. Following the discovery of radioactivity, some attempts had been made to fit the normal distribution to the empirical tables of frequency of intervals between consecutive emissions. These efforts failed, most conspicuously because the frequencies near zero were found to increase (negative exponential gaps) whereas normal theory (over the positive domain) required them to decrease. In an appendix to a paper of Rutherford and Geiger, Bateman (1910) shows that the numbers of particles emitted in fixed time periods satisfy a simple set of differential equations, and that the solutions to these equations are the Poisson probabilities.

This independent discovery of the Poisson distribution was not the first time that (1.1-1) had been obtained from (2.4-7)—Charlier (1905a) contains this result—but it was the first to be derived in the context of a practical application and as such attracted some attention.

The immediate consequences of Bateman's derivation illustrate the primitive state of ideas regarding probability which were current at the time. Marsden and Barratt (1911), who advocate the Poisson distribution ("Bateman's theoretical formula"), seem amazed, and perhaps even anguished, to find that the normal distribution does not fit all continuous data. Snow (1911) tries to make the Poisson into a Pearson distribution ("ideal frequency curves") obtaining a primitive form of the normal limit, but obscuring rather than illuminating Bateman's result.

It was precisely Bortkiewicz who recognized the significance of the differential-difference equations as an approach to a known distribution, and therefore as an additional justification for it. In a characteristically prolix monograph [Bortkiewicz (1913)] he expounds this fact, incidently chiding Bateman for stating that the mean deviation is "much more difficult to calculate" when it had already appeared in 1898 in *Das Gesetz*.

Some writers of the period continued to regard the Poisson when applied to demographic data as justified (if at all) as a binomial limit, and the Poisson when applied to particle counting as justified by means of the differential-difference equations. Bortkiewicz was able to see the connection between the two. Beginning with gap distributions, he developed both Poisson counting and Poisson differential-difference equations.

9.7 APPLICATIONS (1909–1915)

By 1909 the Poisson distribution was beginning to be known in scientific circles, and appears with the standard name "Poisson distribution" in textbooks.

Erlang (1909) found the Poisson distribution helpful in his now classic studies of telephone traffic, using a procedure similar to the differential-difference equation method and citing Poisson. Svedberg (1910), on the other hand, in applying the Poisson distribution to the distribution of particles in solution, refers only to Smoluchowski (1904). This work is continued by Svedberg and Inouye (1911).

Edgeworth's (1911) article on "Probability" in the Eleventh Edition of the *Encyclopedia Britannica* mentions Bortkiewicz and *Das Gesetz*, although the normal distribution continues to prevail with over fifty references to Laplace.

Mortara (1912) continues with the application of the Poisson distribution to demographic data, citing *Das Gesetz* and giving a large number of examples of good Poisson fit to Italian experience.

Timerding (1915) in a monograph on accident causation, uses the Poisson distribution as the basis for his statistical analysis.

A curious episode is the independent discovery at this time of the Poisson distribution by McKendrick [occasionally spelled M'Kendrick], who was serving as a surgeon in the Indian Army. His first two papers [McKendrick (1914a, 1914b)] contain the derivation by means of (2.4-7), and application to blood corpuscle and gas molecule counting. In subsequent publications McKendrick uses the Poisson distribution for the number of deaths from an epidemic (1915a), the number of cases of enteric fever in a house (1915b) and the number of malarial attacks suffered by an individual (1915c). None of these rather obscure publications refers to any other statistical work.

9.8 THE WHITAKER CONTROVERSY (1914–1916)

The last serious writer to misunderstand the meaning of the Poisson distribution was Whitaker (1914). This paper, based on the work of Bortkiewicz (*Das Gesetz* only), "Student" (1907), and Mortara (1912), might never have been published if the author or editor had been aware of the Gini controversy and its outcome. The criticisms put forward by Whitaker cover approximately the same ground as those of Gini (omitting anxiety about the Lexis ratio): that the Poisson distribution is justified only as an approximation to the binomial with fixed probability, that the law of small numbers is an inappropriate name [What is the small number, λ or r of (2.1-1)?], and that the examples cited by Bortkiewicz and Mortara are mostly better fitted by a binomial or negative binomial.

Whitaker's mistaken ideas were corrected by Holwerda (1914) and Fisher, Thornton, and Mackenzie (1922), who remark that her criticism is "entirely vitiated by her neglect of the variation of random samples," and by Bortkiewicz (1915b). It is clear that Bortkiewicz was offended by what he considered an unjustified attack, and the implication of plagiarism (Whitaker dismissed the contribution of Bortkiewicz as "expanding by illustrations Poisson's work") drove him to provide, in a gigantic footnote, every reference to the Poisson distribution known to him. He also corrects Whitaker for supposing that he equated the law of small numbers with the Poisson distribution; this imputation he calls "direkt falsch." In bitter terms he directs Whitaker's attention to the Gini controversy and to his defense. He tends to associate (perhaps quite correctly) Karl Pearson with Whitaker's paper, and concludes:

"Those who believe, as I sincerely do, that Lexian dispersion theory is a firm basis for our science, will resist every attack on it, malicious or not. And so I hope that my 'Law of Small Numbers,' which is throughout oriented towards Lexian theory, will deserve its modest place near that theory and that it will still be remembered when Pearson's negative binomial has long fallen into well deserved obscurity."

The overwhelming style of Bortkiewicz' defense can be judged by the fact that he refers in detail to over fifty other publications. On the intrinsic merits of the Poisson distribution, he is without doubt correct. His presentation is spoiled in some respects by his insistence on the irrelevant Lexis ratio, and even more so by accepting the idea that the negative binomial distribution is somehow an antagonist of the Poisson. In this respect Bortkiewicz was guilty of exactly that error of those who supposed that the Poisson was a threat to the normal or the binomial.

Bortkiewicz' systematic recapitulation of the entire problem [Bortkiewicz (1915a)] was his last publication on the Poisson distribution. Karl Pearson is reported to have said, a propos this paper, that he "did not think Bortkiewicz had added anything to what Whitworth* had done" [Greenwood (1946)]. Pearson was willing to publish, nevertheless, as editor of *Biometrika*, not only Whitaker's (1914) paper, but in the same issue Soper's (1914) improved six-place tables of $p_x(\lambda)$, for $\lambda = 0.1(0.1)10.0$, and two years later K. Smith's (1916) paper which speaks familiarly and objectively of the Poisson distribution and problems of fitting it to data. Although K. Smith still refers only to Whitaker, it is apparent that by this time the Poisson distribution was no longer an unwelcome intruder in the Biometric Laboratory.

It is easy to regard scientific controversies of the past, conducted as they were with narrow and jealous intensity, as being trivial scuffles involving no more than inflated vanity. In the Gini and Whitaker controversies, however, looking beneath the contentious customs of the period, there is observable a genuine anxiety about the Poisson distribution, and especially about the propriety of fitting a distribution to data without a convincing theoretical model.

9.9 GENERALIZATIONS (1917–1920)

Many of the considerations mentioned in Chapter 3, leading to modifications of the Poisson distribution, were already beginning to be noticed

* Actually Boltzmann, see Section 9.2.

before 1920. Rather than argue that the deficiencies of the Poisson rendered it invalid, the most able writers proceeded quite properly to the development of generalizations lacking these deficiencies.

Erlang (1917) proposes the truncated Poisson, and obtains Erlang's loss formula (3.1-6); Erlang (1920b) suggests the Erlang process. Erlang (1920a) considers negative exponential gaps translated away from the origin (Type I Counter); this is treated systematically by Morant (1921).

"Student" (1919) attempts to modify the Poisson distribution to allow for unequal mean and variance in data. This is accomplished by parameter mixing (anticipating the Pólya process) by Greenwood and Yule (1920).

Cramér (1961) takes up the historical narrative at this point, emphasizing the development of many of the ideas mentioned in Chapter 3.