Available online at www.sciencedirect.com

ScienceDirect

Historia Mathematica 34 (2007) 206–220

HISTORIA MATHEMATICA

www.elsevier.com/locate/yhmat

Richard von Mises and the "problem of two races": A statistical satire in 1934

Reinhard Siegmund-Schultze, Sandy Zabell *

Available online 11 October 2006

Abstract

The article "Problem of two races," which appeared in 1934 in French and Russian in the *Matematichesky Sbornik* (Moscow), authored by the noted applied mathematician and statistician Richard von Mises (1883–1953), is interpreted as having a twofold agenda. On its surface it is merely a detailed mathematical analysis of the statistical problem of comparing the distributions of a quantitative characteristic in two different classes. But on closer examination it serves two auxiliary purposes. On the one hand, von Mises, who as a Jew fled Nazi Germany in 1933, satirically attacked Nazi racial doctrines by resorting to statistical parody. On the other hand, von Mises employed his theoretical results to further his own distinctive program of objective Bayesian statistics.

Zusammenfassung

Der Artikel "Problem der zwei Rassen," den Richard von Mises (1883–1953), 1934 in Französisch und Russisch im "Matematichesky Sbornik" (Moskau) veröffentlichte, wird dahingehend interpretiert, dass er eine zweifache Absicht verfolgte. Oberflächlich betrachtet ist er eine detaillierte mathematische Analyse des statistischen Problems des Vergleichs der Verteilungen eines quantitativen Merkmals in zwei verschiedenen Klassen. Aber eine genauere Untersuchung zeigt, dass die Arbeit zwei Nebenzwecke verfolgte. Einerseits kritisierte der 1933 als Jude aus Deutschland geflohene von Mises die Nazi-Rassendoktrin, indem er statistische Argumente parodistisch benutzte. Andererseits verwendete von Mises seine theoretischen Resultate, um sein spezielles Programm einer objektiven Bayesschen Statistik zu propagieren.

Keywords: Richard von Mises (1883–1953); Bayesian statistical inference; Frequency theory of probability; Racist policies of Nazi Germany; Ludwig Bieberbach; George Pólya; Otto Neurath; Ludwig Marcuse

1. Introduction

In 1938 the leadership of the *Deutsche Mathematiker-Vereinigung* (DMV, the German Association of Mathematicians) discussed the expulsion of its remaining Jewish members in Germany and abroad.[1] This was done in close collaboration with the Nazi ministry of education. In these discussions, the president of the DMV, Dr. Wilhelm Süss (1895–1958), noted that Jews and other emigrants who left Germany because of the Nazis might later engage in propaganda attacks against the Third Reich. On March 9, 1938, Süss wrote a letter to Helmut Hasse, Conrad Müller, and Emanuel Sperner, three of his colleagues in the DMV, noting that

* Corresponding author. Fax: (847) 491 8906.
E-mail address: zabell@math.northwestern.edu (S. Zabell).

1 [Remmert, 2004], in particular "Die 'Judenfrage' in der DMV," pp. 223–228. The quotation below is given there in German on page 223.

doi:10.1016/j.hm.2006.05.004

> My attention has been drawn to v. Mises, who allegedly published a statistical proof in Moscow, according to which Jews are intellectually superior to Nordic people.[2]

This remark refers to a 1934 paper by the Austrian-German and Jewish mathematician Richard von Mises (1883–1953), who resigned his professorship at the University of Berlin in 1933 and left Germany for Istanbul (Turkey).[3]

The 1934 paper appeared in the *Matematichesky Sbornik* (in French with an accompanying Russian translation), entitled "Problème de deux races" [Mises, 1934]. The last section of this paper [pp. 207–209] was the source of Dr. Süss's comment, but, far from being a serious argument for racial superiority, the section had (apart from its mathematical content) a purely satirical intent apparently beyond the intellectual horizon of anti-Semitic German mathematicians such as Dr. Süss.

Von Mises's paper is unusual in a number of respects. It consists, for the most part, of a detailed mathematical analysis of an interesting statistical problem, one of greater applicability than its unusual title might suggest.[4] It contains as one illustration of its results a Bayesian application based on von Mises's own rather distinctive blend of objective, frequency-based probability and Bayesian inference, but its actual use in the example von Mises discusses is curious. The paper was published in a distinguished Russian journal, but one unlikely to be read by most statisticians outside the Soviet Union. In later years von Mises himself rarely if ever referred to the paper; and although reprinted in his *Selecta* [Mises, 1963/1964], it has almost entirely escaped notice.

Anyone familiar with von Mises's philosophy and writings will no doubt be rather surprised to learn that a paper of his would contain a result of the kind alleged by Süss. In this article, after briefly describing the content of the section Süss found so provocative, we present a detailed analysis of the mathematical content of von Mises's paper and assess its statistical contribution. We then discuss in detail the reasons for the unexpected satirical twist von Mises gave it, placing it in its political context.

2. The statistical "problem of two races"

In the introduction to his paper, von Mises poses a problem, which he describes as well known in both statistics and actuarial science:

> A set containing n individuals is composed of two classes, αn individuals belonging to a class A and $\beta n = n - \alpha n$ to a class B. Let z be any distinctive characteristic whatever, say for example the height of an individual in the set. We assume the existence, for each of the two classes, of probability distributions $A(z)$ and $B(z)$ respectively. In particular $A(z)$ denotes the probability that the distinctive characteristic of an individual in class A does not exceed the value z. We ask: what is the probability $P(x)$ that, among the m individuals displaying the largest values of z in the set, one finds x individuals in class A and $y = m - x$ in class B? [Mises, 1934, p. 192]

In addition to this question, of a very classical kind in mathematical probability, von Mises also considered the "inverse" problem: "given the values of x and y, what is the resulting conditional probability of a specified relationship holding between $A(z)$ and $B(z)$?" [Ibid., p. 193]. These results, comprising the first five sections of the paper, are then applied in the final section to two examples, the second of which von Mises suggests, "will, because of its particular subject matter, perhaps be of special interest." In this example von Mises considers:

> In a European country, the inhabitants of which number about $65 \cdot 10^6$, the population consists of two races in the relative proportions $\alpha = 0.009$, $\beta = 0.991$. [Ibid., p. 208]

[2] Freiburger Universitätsarchiv, Bestand DMV, E4, No. 36 (Schriftführer Müller), Teil 5 (Judenfrage).

[3] Von Mises did so in the wake of an initial round of dismissals of Jews from academic positions throughout Germany. Technically speaking, von Mises had not been forced to leave his position: having served as a state official prior to the first World War and in the Austrian military during it, he was exempt from this initial round of expulsions. But he realized the ultimately untenable nature of his situation. (And rightly so: he would have certainly been dismissed by 1935.) On such dismissals and ensuing emigration from Nazi Germany in general, see [Siegmund-Schultze, 1998]. For further biographical information regarding von Mises, see [Siegmund-Schultze, 2004].

[4] This discrepancy between the title of the paper and its contents already hints at its partly satirical intent.

One immediately recognizes Germany as being the "European country" in question,[5] and the 0.9% is a reference to its Jewish inhabitants (including those converted to Christianity); this percentage was indeed a realistic estimate for the country as a whole (although not for large cities such as Berlin, where the percentage was substantially higher). Never actually explicitly referring to either Germany or the Jews throughout, von Mises went on to argue:

> Only a very small number of these inhabitants perform scientific research in physics and chemistry. An absolute scale for measuring the scientific ability (the value z) of a physicist or a chemist does not exist. However, one might grant that the recipients of the Nobel Prize form a group having the largest values of z. The list of recipients for the years 1901 to 1933 contains the names of 27 individuals coming from the said country, among which 5 belong to class A, so that one has $m = 27$, $x = 5$, $y = 22$. [Ibid., p. 208]

The figures given by von Mises are accurate: between 1901 and 1933, 27 individuals working in Germany received the Nobel Prize in physics or chemistry.[6] The five Jewish Nobel Prize laureates von Mises had in mind may have been Albert Einstein (physics, 1921), James Franck and Gustav Hertz (both physics, 1925), Fritz Haber (chemistry, 1918), and Richard Willstätter (chemistry, 1915).[7]

Von Mises contrasted the percentage of Jewish recipients of the Nobel Prize with the much lower percentage of Jewish inhabitants of Germany, using three different statistical arguments. The first of these was a classical test of significance, von Mises concluding that if scientific talent were equally distributed between the A and B classes, the probability of finding five or more Nobel Prize winners in the smaller A class was less than 1 in 100,000.

For this reason, rejecting the hypothesis of equal distributions of talent, von Mises went on to estimate how much more likely it was that a person in the A class had outstanding ability in physics or chemistry than one in B. Using the actual observed odds of 5 : 22, von Mises concluded that the probability for A was approximately 25 times greater than that for B.

Finally, using a third, Bayesian argument, von Mises concluded that there was a nearly 85% chance that the probability of having such ability among A individuals was at least 20 times and at most 42 times greater than the corresponding probability for B individuals.

How did von Mises arrive at these conclusions? Did he really believe such arguments? What was his motive in making them? The remainder of this paper is an attempt to answer these questions.

3. Von Mises's statistical analysis

3.1. Von Mises's model

There are several standard sampling models for the generation of the entries in a two-by-two table:

x	$a - x$	a
y	$b - y$	b
m	$n - m$	n

The most common sampling models include n not fixed (entries Poisson), n fixed (test of independence, entries multinomial), one set of margins fixed (comparison of two binomial proportions), and both sets of margins fixed (hypergeometric sampling). Von Mises considers an apparently different model: the individuals fall into two classes, A and B, containing a and b members, respectively. He postulates the existence of a numerical measure Z of some characteristic of the members of the population. These numerical values are not observed, however, only how many of the top m values fall into each of the two groups.

Let $A(z)$ and $B(z)$ denote the distributions of this numerical quantity in two populations (*Kollektivs*, in von Mises's terminology). That is, it is assumed that the Z-values of the A-individuals are a random sample from an $A(z)$

[5] And was recognized as such by George Pólya: see Section 5.3 below.

[6] Some of the 27 recipients were not born in Germany (Wilhelm Ostwald, for example, who received the prize in 1909 for chemistry, was born in Riga, then part of Russia).

[7] Adolf von Baeyer (chemistry, 1905) and Otto Wallach (chemistry, 1910) are also sometimes included on lists of Jewish Nobel prize winners. Both had died by the time von Mises wrote, and he may not have been aware of their background.

distribution, and the Z-values of the B-individuals are a random (and independent) sample from a $B(z)$ distribution. Thus we have

$$Z_i^A \sim A(z), \quad 1 \leqslant i \leqslant a; \qquad Z_j^B \sim B(z), \quad 1 \leqslant j \leqslant b.$$

Suppose that of the m top-ranking individuals, x belong to class A and $y = m - x$ to class B. Then, assuming $A(z)$ and $B(z)$ satisfy natural regularity conditions, von Mises shows that the probability of observing x of the top m falling into the A class is (suppressing the dependence of the $A(z)$ and $B(z)$ functions on z in the integrals)

$$P(x) = \binom{a}{x}\binom{b}{y}\left[x\int_0^1 A^{a-x}(1-A)^{x-1}B^{y-b}(1-B)^y\,dA + y\int_0^1 A^{a-x}(1-A)^x B^{y-b}(1-B)^{y-1}\,dB\right].$$

The remainder of the paper consists of an analysis of the consequences of this expression.

3.2. Case 1: the null hypothesis $A(z) = B(z)$

In Section 2 von Mises considers the special case when the distributions $A(z)$ and $B(z)$ coincide. In this case $P(x)$, the probability that x individuals in class A rank among the top m values, is

$$P(x) = \binom{a}{x}\binom{b}{y} m \int_0^1 A^{n-m}(1-A)^{m-1}\,dA = \frac{\binom{a}{x}\binom{b}{y}}{\binom{n}{m}}.$$

That is, von Mises's model reduces in the null case to the hypergeometric distribution.

Note that in the case of the null hypothesis $A = B$, all dependence on $A(z)$ disappears. There is an instructive parallel here with Bayes's celebrated argument in his paper of 1764 [Bayes, 1764]. In the case considered by Bayes (but described from a modern perspective), a sequence of points $X_0, X_1, X_2, \ldots, X_n$ are chosen at random (that is to say, uniformly) from an interval $[a, b]$, and the X_j, $j \geqslant 1$, are scored by whether they are greater than or less than X_0. Bayes showed that if S_n denotes the number of $X_j \leqslant X_0$, then

$$P(S_n = k) = \frac{1}{n+1}, \quad 0 \leqslant k \leqslant n.$$

It is an interesting observation, dating back to Karl Pearson [1921], that the distribution of S_n does not depend on the specific choice of distribution for the X_j, provided that distribution is continuous (that is, the cumulative distribution function $F_X(x) = P(X_0 \leqslant x)$ is continuous in x). (The parallel goes even further. In von Mises's derivation, the total probability is computed by conditioning on the value of the smallest value of z among the top m, and then the remaining $n - 1$ values are graded on whether they are greater than or less than z.)

The appearance of the hypergeometric distribution in this context is *not* an independent derivation of the Fisher exact test, published the following year [Fisher, 1935, pp. 48–51]. For von Mises, the sampling model under the assumed null *is* the hypergeometric; for Fisher, the exact test is applicable whatever the underlying sampling model generating the table. (Fisher argued that the analysis of such tables should be conditional on the observed values of the margins, which he regarded as ancillary, and that the conditional distribution of $P(x)$ was hypergeometric.)

In the limit as $n \to \infty$ but $m, x, y, \alpha = a/n, \beta = b/n$ remain fixed, von Mises notes that

$$\lim_{n\to\infty} P(x) = \binom{m}{x}\alpha^x\beta^{m-x};$$

this is the standard observation that the limit of the hypergeometric (under the appropriate limiting regime) is the binomial. (It is curious that, with the single exception of a paper by Romanovsky written in 1931, von Mises makes no reference throughout to any earlier work except his own, and makes no attempt to identify which aspects of his paper are novel.)

3.3. *Case 2: the alternative hypothesis $A(z) \neq B(z)$*

The next two sections of von Mises's paper consider the general case, under the two regimes $n \to \infty$, m fixed, and $n \to \infty$, m/n fixed. This is the technical heart of the paper, and its results are likely original.

3.3.1. *Regime $n \to \infty$, m/n fixed*

The contents of this subsection are not needed in what follows, and are stated here merely for completeness. The necessary technical assumptions are omitted. Let $\mu = m/n$, $\xi = x/n$, z_0 such that $\mu = 1 - \alpha A(z_0) - \beta B(z_0)$, $A_0 = A(z_0)$, $B_0 = B(z_0)$, $\xi_0 = \alpha(1 - A_0)$, $\eta_0 = \beta(1 - B_0)$, $B_0' = (dB/dA)(z_0)$, and

$$h^2 = \frac{\alpha\beta(\alpha + \beta B_0')^2}{\alpha^3\eta_0(\beta + \eta_0) + B_0^2\beta^3(\alpha - \eta_0)}.$$

Then

$$P(x) \approx \frac{h}{\sqrt{2\pi n}} e^{-\frac{1}{2}\frac{h^2(x-n\xi_0)^2}{n}}.$$

3.3.2. *Regime $n \to \infty$, m fixed*

This is the key result von Mises needs for his application. He introduces the assumption that

$$\lim_{A \to 1} \frac{1 - B(z)}{1 - A(z)} = \lambda.$$

In this case he derives the asymptotic formula

$$\lim_{n\to\infty} P(x) = \binom{m}{x} \frac{\alpha^x (\lambda\beta)^y}{(\alpha + \lambda\beta)^m}.$$

That is, the limiting distribution of $P(x)$ is binomial with parameters m and $\alpha/(\alpha + \lambda\beta)$.

3.4. *The inverse problem*

Von Mises now turns to the inverse problem in Section 5. He assumes for simplicity that the "lambda-relation" holds exactly:

$$\frac{1 - B}{1 - A} = \lambda, \quad A + B \geqslant 1; \qquad \frac{A}{B} = \lambda, \quad A + B \leqslant 1.$$

In this case the general expression for $P(x)$ reduces to a function von Mises denotes as $Q(\lambda)$. (This is quite similar to Fisher's [1935, pp. 50–51] choice of alternatives to the null in his analysis of the two-by-two table.)

If $q(\lambda)$ is the (as yet unidentified) prior density of λ, then the posterior density of λ (given x) is

$$p(\lambda) = C \cdot q(\lambda) Q(\lambda),$$

C being a normalizing constant that ensures that the expression integrates to one.

In order to apply this analysis to an actual example (as he intends later on), von Mises must confront head on the choice of the prior density $q(\lambda)$. He begins by noting that the common choice of taking $q(\lambda)$ to be constant in some range is arbitrary, and that some other choice, such as taking $q(\lambda)$ to be proportional to $(1+\lambda^2)^{-1}$, might well seem more natural, but readily admits the ultimate arbitrariness of the undertaking:

> [O]ne cannot attach too much importance to a solution of an inverse problem the outcome of which depends crucially on the choice, arbitrary in some measure, of the prior probability. [Ibid., p. 206]

Because of this von Mises resorts to a then already common argument that for large sample sizes (here m), the posterior density is largely independent of the prior density. In consequence, one can choose a prior that is mathematically tractable.

In order to evaluate the resulting expression, von Mises introduces a convenient change of variable, letting

$$z = \frac{\alpha}{\alpha + \lambda\beta}, \qquad 1 - z = \frac{\lambda\beta}{\alpha + \lambda\beta}.$$

(This seems an unfortunate notational choice, since it unnecessarily conflicts with the prior, very different use of the variable z.) Obviously $0 \leqslant z \leqslant 1$, and von Mises elects to use the uniform prior on z, obtaining

$$p(\lambda)\,d\lambda = (m+1)\binom{m}{x} z^x (1-z)^{m-x}\,dz.$$

This is an interesting result: if $\lambda = 1$, then the probability of an individual in the top m falling in the A group versus the B group is proportional to the respective sizes of the two groups; and other values of λ represent different weightings of the two groups (large values of λ favoring B and small values of λ favoring A).[8] Although no prior density for the initial parameter λ naturally suggests itself, transforming from the factor λ to the probability z permits one to use the mathematically more convenient uniform prior on z.

3.5. The application

To illustrate his results, von Mises postulates a country for which $\alpha = 0.009$, $\beta = 0.991$ and $m = 27$, $x = 5$.

3.5.1. The first statistical argument

Von Mises computes from his formulas that, under the null hypothesis $A = B$,

$$P(5) = \binom{27}{5}(0.009)^5 \cdot (0.991)^{22} = 0.00000391.$$

It is not entirely clear why von Mises views this number as relevant, since in computing a significance probability it is necessary to compute the probability of obtaining a value *as great or greater* than the one seen; here $P(x \geqslant 5)$.

It is in fact not entirely clear whether von Mises understood the difference. On the one hand, he notes that

$$P(x > 0) = (0.991)^{27} = 0.2166, \qquad P(x > 1) < 0.03, \qquad 1 - P(x < 5) \leqslant 10^{-5}.$$

The last of the three inequalities is equivalent to noting that the standard (one-sided) level of significance $P(x \geqslant 5)$ is less than 1 in 100,000. But von Mises then goes on to summarize the above by stating:

> It follows from all this that—under the assumption that A and B are equal—the result $x = 5$ which is observed would have an extremely small probability. [Ibid., p. 208]

This appears to (incorrectly) revert back to regarding to $P(5)$ as the quantity of primary interest.[9]

---

8 The λ factor corresponds either to Fisher's ψ [1935, pp. 50–51], or its reciprocal, depending on the labeling of the categories.

9 It is true that the major and preponderant contribution to the tail probability arises from the lead term: $P(x \geqslant 5) = 0.00000404$, so $P(x = 5)$ is 97% of $P(x \geqslant 5)$.

The satirical nature of the analysis becomes apparent when von Mises continues on in similar vein to consider the case $m = 1, x = 1$: “One might wonder whether or not the ‘greatest physicist’ would be found among the names belonging to the first of the two classes” [Ibid., p. 208]. The calculation is trivial in this case: $P(1) = 0.009$; presumably von Mises had Einstein in mind here.

3.5.2. The second statistical argument

Using the formulas he had derived based on the assumption of differing distributions of talent in the two classes A and B, the second argument von Mises employed was based on a point estimate of the λ parameter, setting $0.009/(0.991\lambda)$ equal to the observed ratio of $5/22$, so that $\lambda = 0.04$. Thus, von Mises argued [pp. 208/209]:

> On the other hand, in rejecting the hypothesis $A(x) = B(x)$ we arrive at probabilities which are much more reasonable for the observation that $x = 5$, $y = 22$. Let us grant that λ satisfies the equation
>
> $$\alpha : \lambda\beta = x : y = 5 : 22, \qquad \lambda = 0.04 = \frac{1}{25}.$$
>
> This corresponds to the hypothesis that the probability of great talent in physics or chemistry is judged to be 25 times higher for an individual in class A than in class B. In this case one has $P(5) \ldots = 0.194$.
>
> This value is about 5.5 times larger than the average probability (*probabilité moyenne*) $1 : 28$. The result $x = 5$ now constitutes the most probable case.

Put another way, the observed odds of $5 : 22$ in favor of an A individual rather than a B individual having great talent in physics or chemistry is 25 times greater than the naïve estimate of the odds based simply on their respective population proportions of α and β.[10] The final comments, however, do not seem particularly incisive. The *probabilité moyenne* $1/28$ is irrelevant.[11] And the fact that $x = 5$ is the modal value of $P(x|\lambda)$ when $\lambda = 0.04$ corresponds to no recognized criterion of estimation. (The method of maximum likelihood would maximize $P(x|\lambda)$ as a function of λ.)

3.5.3. The third statistical argument

Finally [p. 209] von Mises applies the technical Bayesian results in his paper to the example and concludes:

> Thus there is a probability of around 85% for the fact that among the individuals of class A the probability of an eminent talent in physics or chemistry is at least 20 times and at most 42 times greater than in class B.[12]

On this provocative note, von Mises concluded his paper. What were his intentions?

4. Von Mises’s statistical philosophy

In order to understand the statistical analysis in von Mises’s paper, it is important to recognize that he had strong opinions regarding the nature of mathematical probability and statistical inference, views that blended in idiosyncratic fashion both frequentist and Bayesian perspectives.

4.1. Von Mises’s frequentism

Von Mises, not surprisingly given his background in applied mathematics, advanced an objective, frequentist theory of probability, one more complex and nuanced than had previously been advanced.[13] According to von Mises, probabilities are only properties of repetitive events or mass phenomena:

[10] If p is the probability of an event, then $p/(1 - p)$ is the odds in favor of the event.

[11] The 28 reflects that there are 28 possible values for x ($0 \leqslant x \leqslant 27$); see p. 208. (And of course the “1 : 28” is itself a mistake; it should be either 1 : 27 or 1/28.)

[12] These limits can also be interpreted, as above, as the limits on the multiple of α/β giving the odds favoring A over B.

[13] The primary element of novelty in von Mises’s theory is its introduction of invariance under place selection, discussed below. See generally [Von Plato, 1994, Ch. 6].

> The rational concept of probability, which is the only basis of probability calculus, applies only to problems in which either the same event repeats itself again and again, or a great number of uniform elements are involved at the same time. Using the language of physics, we may say that in order to apply the theory of probability we must have a practically unlimited sequence of uniform observations. [Mises, 1957, p. 11]

Central to this concept of probability is von Mises's concept of a *Kollektiv*: an infinite sequence of observations or events, each giving rise to one of several possible outcomes. It is assumed that the relative frequency of each of these outcomes has a limiting value, termed its *probability*. In order to rule out such obviously nonrandom sequences as 101010101010..., it is also assumed that the limiting frequency remains unchanged when one passes to a subsequence of the *Kollektiv* chosen according to some rule or algorithm termed a *place selection*.

Such an approach places von Mises in a clear line of descent from earlier frequentists as Cournot, Ellis, Venn, and Fechner, the notion of place selection giving his theory its own distinctive flavor.[14] Von Mises never gave a mathematically precise definition of the concept of place selection; one natural interpretation is to assume that place selections correspond to the class of computable functions (although this concept did not exist at the time when von Mises first proposed his definition). Because an outcome can be embedded in more than one possible *Kollektiv*, probabilities are attributes of the *Kollektiv* rather than the individual outcome: "Let me insist on the fact that in no case is a probability value attached to a single event as such, but only to the event inasmuch as it is the element of a well-defined sequence" [Mises, 1941, p. 341].

4.2. Von Mises's Bayesianism

Despite this empirical view of the nature of probability, von Mises also adhered, as many of his contemporaries still did in the 1920s, to an ultimately Bayesian view of the nature of statistical inference: parameters or hypotheses are assigned *prior* (or *initial*) probabilities, and then new information is incorporated via conditioning using Bayes's theorem, resulting in *posterior* (or *final*) probabilities. It is this unexpected blend of an objective notion of probability within the ordinarily subjective Bayesian framework that gives von Mises's view of statistical inference its own distinctive flavor.

The use of Bayes's theorem of course is not controversial if the prior distribution is in fact known.[15] For von Mises this prior distribution must itself arise from some *Kollektiv* (see, e.g., Mises, 1941, p. 349; 1942, p. 361). This *Kollektiv* represents an objective but in general unknown probability distribution, one that cannot be supplied by a nonempirical appeal to some form of the principle of indifference. The statistician is therefore faced with the quandary of what to do. Von Mises's resolution of this dilemma is to invoke what he termed the "second law of large numbers": under very general conditions, if the sample is large, then the posterior probability distribution for a parameter is largely independent of the choice of prior distribution, and is sharply peaked about the sample estimate; see [Mises, 1941, 1942, 1957]. Thus any mathematically convenient prior can be used.

This result is sometimes termed the *Laplace–Bienaymé theorem* [Heyde and Seneta, 1977, Sect. 5.2; Hald, 1998, p. 501], but is better known today as the *Bernstein–von Mises theorem* (see Diaconis and Freedman, 1990, p. 1318, who briefly discuss its history). Von Mises proved a rigorous version of it in his first paper on probability theory after World War I [Mises, 1919]. The argument itself, as a justification for the use of flat priors, can already be found in earlier authors such as Cournot [1843, Sect. 95].

4.3. The problem of two races redux

It is interesting to view the example that von Mises gives at the end of his paper in the light of the above. Von Mises's Bayesian calculations are curious from several points of view. First, as noted earlier, for von Mises probabilities are only meaningful in the context of a *Kollektiv*. But more than one *Kollektiv* may be relevant. Take, for example, the Bayesian analysis of tossing a coin. There is a parameter p representing the probability of heads, a prior distribution

[14] For Ellis and Venn, see [Zabell, 1991].

[15] See, for example, R.A. Fisher's discussion of this point in [Fisher, 1973, pp. 18–20].

$d\mu(p)$, and the probability, *given* p, of getting k heads out of n, the binomial probability

$$b(k; n.p) = \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \leqslant k \leqslant n.$$

Given a specific coin, we may imagine an infinite sequence of tosses of that coin, constituting one *Kollektiv*. The limiting frequency of heads is p, the probability of heads. Or we may imagine tossing the coin k times, over and over again, giving rise to a second *Kollektiv*; the probability of seeing k heads in n tosses then being $b(k; n, p)$. The two *Kollektivs* are related, and it is possible to derive the second probability from the first.

But suppose instead that the coin is regarded as having been chosen from some large population of coins, or is generated by some process that produces a sequence of coins, one after the other. Each coin has a probability p of heads, and the totality of such attributes, one per coin, can itself be regarded as a *Kollektiv*. This is von Mises's interpretation of the prior or initial probability $d\mu(p)$. (Strictly speaking, in later years von Mises did not require the prior probabilities to satisfy the randomness requirement of insensitivity to place selection; see, e.g., Mises, 1942.)

Here, for example, the value of the parameter $z = \alpha/(\alpha + \lambda\beta)$, relating the scientific abilities of the current citizens of a country (in the A and B classes), must also be regarded as a member of some *Kollektiv*. But although we can certainly imagine a coin as drawn from a large collection of coins, there is no serious sense in which the z-parameter, an attribute of a country, could be so regarded. (To begin with, the number of such z values could not exceed the number of countries on the planet.)

Von Mises insisted that his notion of probability was "objective," and derided adherents of the subjective school, sometimes going so far as to use epithets as "absurd." But any attempt to force the analysis of the problem of two races into the Procrustean bed of the *Kollektiv* forces one to conclude that either that concept is too narrow, or the example von Mises considers does not admit of rigorous statistical analysis.

But there are equally serious technical problems when von Mises makes two successive appeals to asymptotics. One, taking the limit as n becomes large, is certainly reasonable in the context of his example: $n = 65{,}000{,}000$. But in order to apply his own asymptotic result ("the second law of large numbers"), von Mises also requires m to be large, so large that the effect of any particular prior will be negligible. But this (a) depends in part on the particular prior being used (particularly "opinionated" priors requiring a larger sample size before their initial impact "washes out"); and in any case (b) the appeal to the asymptotics of the second law of large numbers is patently absurd for $m = 27$. For all his insistence on *Kollektivs* and the second law of large numbers, von Mises's Bayesian application of his results turns out to be in a situation where no *Kollektiv* is evident, the sample size is small, and the much-maligned Bayes–Laplace prior is used!

What is particularly curious is that von Mises was quite sceptical about inferences based on small samples. In his posthumously published textbook [Mises, 1964, pp. 347–348], he states (and the italics are his):

> It remains an invariable fact, pervading all problems in mathematical statistics, that *no substantial inference can be drawn from a small number of observations if nothing is known prior to the experiments about the object of experimentation*.

One can only conclude that the example von Mises chose to illustrate his new methodology was singularly ill-suited, given his own statistical *Weltanschauung*. This suggests that he had other motives, ones which led him to use the example, either blinding him to its defects, or inducing him to use it despite (or perhaps even because of) them. To such motives we now turn.

5. Von Mises's political agenda

The statistical results in von Mises's *Problème de deux races* are clearly both mathematically substantial and serious in intent. The application of those results to the distribution of talent among races, however, cannot be read as anything other than ironic. Given his broad interests in the social sciences, von Mises was certainly aware of the social, economic, and cultural reasons for the relatively high percentage of Jews among German intellectuals.[16] Nor could he

[16] See von Mises's textbook on philosophical positivism [Mises, 1951], first published in German in 1939, as well as his until recently unpublished manuscript, written about 1934, "Die Mathematik und das Dritte Reich" [Sheynin, 2003], which will be discussed below.

have seriously believed that the German Jews of his time were on average 25 times more scientifically talented than the German population as a whole, merely on the basis of a naive percentage comparison.[17]

5.1. Ludwig Bieberbach and Deutsche Mathematik

It seems clear that von Mises's example was intended instead as a satirical reply to Nazi racism, which had, particularly since 1933, advanced pseudo-scientific theories concerning alleged race-related "styles" of research and teaching, and implemented concrete legal measures to exclude Jewish students and professors from German universities.[18] Von Mises's former colleague in Berlin, Ludwig Bieberbach (1886–1982), for example, had developed a racist theory of "stylistic kinds of mathematical creativity" (*Stilarten mathematischen Schaffens*) claiming that "blood and race affect the kind (*Art*) of mathematical creativity" [Bieberbach, 1934, p. 358]. In a lecture given in early April 1934 to the *Förderverein* (the German Mathematical Education Association), Bieberbach defended a Nazi-organized student boycott of lectures by the Göttingen number theorist Edmund Landau, arguing that "representatives of overly different races do not mix as students and teacher" and vigorously supported a series of measures put in place by the Nazis to expel German Jews from German universities and bar those not yet enrolled from gaining entrance.[19] (Bieberbach's theories, as stated, were to a certain extent ambiguous: they could either be interpreted as just asserting that German and Jewish mathematical intellect were incompatible, or that the mathematics produced by foreign races was in fact actually inferior to *Deutsche Mathematik* according to some unspecified, ostensibly objective standard. Bieberbach's followers used whichever interpretation best suited the purposes of the moment; see Segal, 2003, pp. 375ff.)

More generally, the Nazis objected to the disparity between the percentage of Jews in German society and their representation in the professional and intellectual classes. On April 25, 1933, shortly after they came to power, the Nazis passed a law limiting the percentage of Jewish students in German universities to a maximum of 1.5%.[20] Here—as in Bieberbach's pseudo-philosophy—there were differing rationales. On the one hand, the measures reflected a belief on the part of some Germans that German Jews had unfairly achieved an unjustified overrepresentation in the German university system (rather than at least in part the reality that German society had effectively limited Jews to certain professions and excluded them from others). But these measures also reflected the view of some that intellect (*Intellekt*) per se was not the most important quality to foster in the schools and universities of the new Germany, character (*Charakter*) being more important.[21]

Von Mises himself (who had by then left Germany) had little patience with Bieberbach's racial theories. In a 1934 manuscript, "Die Mathematik und das Dritte Reich" (which remained unpublished under his name during his lifetime), von Mises dismissed contemptuously Bieberbach's "subordination of mathematics to the racial theory of the new German type" (*Rassenkunde neudeutscher Prägung*) [Sheynin, 2003, p. 138], and preferred to discuss instead the more general cultural roots of German anti-Semitism (what he described as "the basic German phenomenon of lack of restraint" [*deutsches Urphänomen der Masslosigkeit*]) and its underlying social causes.

Bieberbach's April lecture before the *Förderverein*, noted earlier, was the subject of a "detailed and ironic newspaper report" in the *Deutsche Zukunft* ("German Future"); as a result "news of Bieberbach's ideas spread quickly and prompted sharp reactions from abroad" [Mehrtens, 1987, p. 195]. (For example, an article by the Danish mathematician Harald Bohr appeared in the May 1, 1934 issue of the Danish newspaper *Berlingske Aften* [Ibid., p. 221].) It is thus possible that the ironic concluding application to von Mises's paper (which was received by the *Sbornik* on May 20) was in fact added in direct response to Bieberbach's lecture. (If so, this might explain the poor fit between the applied and theoretical portions of the paper, noted earlier.)

[17] Simple percentage comparisons of the kind performed by von Mises are problematic for a number of reasons, including (but not limited to) the choice of an appropriate benchmark population, the need to adjust for relevant explanatory covariates, and phenomena such as Simpson's paradox. In addition, comparisons based on observational data can only establish correlations between traits of members of a population under study, not direct causative links; see generally [Freedman et al., 1997, Ch. 2]. Von Mises would certainly have been aware of such problems.

[18] See generally Siegmund-Schultze [1998].

[19] [Mehrtens, 1987, p. 227] (English translation from Bieberbach's German). Mehrtens notes that Bieberbach carefully avoided the explicit use of terms such as "Jewish" and "Aryan" mathematics, although several of his followers were more explicit in this respect.

[20] This was the infamous *Gesetz gegen die Überfüllung deutscher Schulen und Hochschulen* ("Law against the Over-crowding of German Schools and Universities").

[21] Regarding this view, see the Nazi poem quoted by George Pólya in a letter to von Mises, given in Section 5.3 below. The juxtaposition of intellect and character was common in Nazi writings of the period.

In his paper in the *Sbornik*, von Mises, who as a Jew living in Germany had been directly affected by Nazi racial policies, satirically countered the assertions of racial superiority explicit and implicit in the bureaucratic measures of the Nazi regime and racial theories such as Bieberbach's. If one was willing to uncritically credit simple percentage-difference comparisons, naively viewing them as a matter for serious discussion without taking into account the obvious social, economic, political, and cultural differences in contemporary German society, one could equally well deduce from similar comparisons a superiority of Jewish intellect. In the physical sciences (unlike mathematics) the degree of ability could even be arguably quantified, using the indicia of Nobel prizes. (The corollary for mathematics was obvious: if one arrived at paradoxical results in the case of the physical sciences as von Mises had shown, one was even less justified linking the quality of mathematical results with the racial background of their creators, where one did not even have such a clear and objective measure of quality.)

5.2. Why the Sbornik?

The satirical point of von Mises's paper was obviously lost on some people, such as Dr. Süss's informant (see the quotation at the beginning of this paper), who evidently took the analysis at face value as a serious argument for the racial superiority of the Jews. The point was presumably not lost, however, on most of von Mises's Russian readers, many of whom had experienced anti-Semitism at first hand (albeit not in as equally open, explicit, and state-regulated a way as in Nazi Germany). In any case it is natural to speculate about why von Mises decided to publish his paper in the Soviet Union. One obvious contributing factor might have been that von Mises had a number of professional ties with scientists in Russia, including at least one close friend, the physicist L. Mandel'shtam, and had visited the country at least once (in 1928, during the Congress of Russian Physicists).[22]

It is interesting to note that the *Matematichesky Sbornik* published von Mises's paper as the lead article in that issue of the journal, giving the complete text of the paper in both French and Russian (contrary to its usual practice of just appending a Russian abstract to papers written in languages other than Russian).[23] The decision to publish the paper not only as the lead article but in both languages might have been motivated by a desire to give it as wide a readership as possible.

It is important to note that at the time the article appeared contributions to Soviet journals by either non-Russians or by Russians in foreign languages, as well as by Russians to foreign journals, were under close political scrutiny in the Soviet Union. For example, an anonymous editorial published a few years earlier in the *Sbornik* ("Soviet mathematicians support your journal!") [vol. 38 (1931), part II, p. 1] appealed to its Russian readership not to publish abroad in "bourgeois" countries. In order to secure the international importance of Soviet mathematics, the editorial promised to continue to publish foreign abstracts of Russian articles. This practice was, in fact, maintained at least into the 1940s.

The decision to give von Mises's paper such prominence in an issue of the journal—despite such political scrutiny—suggests that this might have been to advance another agenda, pursued by some Russian mathematicians who supported von Mises's views against those of other Russian mathematicians and dogmatic Marxist philosophers who opposed those theories on scientific or philosophical grounds. These very interesting and multifaceted problems, however, cannot be discussed here and have to be reserved for future research.

5.3. Contemporary reactions

Von Mises's paper on the *Problème de deux races* appears to have gone largely unremarked in subsequent literature. The authors have not come across any later mention of it by von Mises in his papers or books, nor any comment on it

[22] Von Mises also had common mathematical interests with both S.N. Bernstein and A.Ya. Khinchin. Khinchin edited a Russian translation of von Mises's 1928 book *Probability, Truth, and Statistics* (the translation appeared in 1930, but—presumably for political reasons—left the word "truth" out of the Russian title!), was on the editorial board of the *Sbornik*, and was an admirer of von Mises's work in probability (while at the same time being a critic of its Machian positivist connotations); see Khinchin [1936/1944].

[23] It is perhaps no accident that one of the few exceptions to this practice occurs in the same issue of the *Sbornik*: a short paper by Jacques Hadamard "Un cas simple de diffusion des ondes" [vol. 41, pp. 402–404] is also followed by a complete Russian translation [pp. 404–407]. It is possible that this shorter piece was included in order to conceal the unusual nature of the translation of von Mises's much longer text.

by the editors of his work (in particular regarding its satirical intent).[24] It did not, however, go unnoticed at the time.[25] In fact, a very different reaction (compared to that of the DMV leadership mentioned in the introduction) came from someone on the other side of the political spectrum, George Pólya (1887–1985), who had corresponded with von Mises on probability theory since the publication of the latter's paper [Mises, 1919] on the *Fundamentalsätze*.[26]

On April 6, 1935, Pólya, who was born in Hungary and was Jewish by birth, wrote a letter to von Mises from his position in Switzerland.[27] "Thank you very much for your off-prints, in particular for 'Le problème de deux races'! Really very nice [*hübsch*], I was *very pleased* [underlined by Pólya in the letter] in every respect." Later in the letter, Pólya referred to von Mises's allusion to the "country, of which the number of inhabitants is about $65 \cdot 10^6$," remarking that he (Pólya) did not want to publish his planned "book on the solution of mathematical problems" in that country, although he would prefer to write it in German.[28] (In another letter written to von Mises a month later on May 4, 1935, Pólya ironically quoted a Nazi poem to illustrate the Nazi contempt for intellect noted earlier.)[29]

Even the novelist, philosopher, and theater critic Ludwig Marcuse (1894–1971), who as a Jew had fled Berlin in 1933, going first to France and then via the Soviet Union in 1939 to the United States, took note of von Mises's article.[30] In his German reminiscences of 1960 Marcuse wrote:

> Sometimes it was possible to simply laugh away [*wegzulachen*] the unbearable. Professor Mises, once ordinary professor for mathematics and applied mechanics at the University of Berlin, remained temporarily in his position as a former officer in the war. He then fled to Istanbul and published there in the *Recueil Mathématique* his paper "Problème de deux races," which studied the relative fraction of Jewish-German Nobel Prize recipients, culminating in the following formula:
>
> $$P > 1 - \frac{(x+1)(m-x+1)}{(m+2)^2(m+3)Z^2}.$$
>
> In German ["auf deutsch," that is to say, in simple or lay terms]: there is a probability of 85% that the talent of German Jews for physics and chemistry is at least 20 times, at most 42 times bigger than that of German non-Jews. [Marcuse, 1960, p. 165]

(The formula is the same as the one in von Mises's paper [1934, p. 207], except for a minor change in notation.)[31]

---

24 Notably none by von Mises's widow Hilda Geiringer, a capable statistician in her own right who edited von Mises's *Selecta* [1963/1964] and lectures on probability [Mises, 1964] and left detailed records on the process of editing in the von Mises Papers at Harvard.

25 Most notably by Gumbel [1935, p. 130]. It is possible that von Mises's paper played a part in Gumbel's interest in the subject of extreme value theory. E.J. Gumbel (1891–1966) is perhaps best known for his work in statistical extreme value theory. He was one of the most outspoken leftist intellectuals in the Weimar Republic and for this reason had been dismissed from his position at the University of Heidelberg in 1932. Although likely not in full agreement with von Mises politically, he frequently published reviews and short notes in von Mises's journal, the *Zeitschrift für Angewandte Mathematik und Mechanik*.

26 Pólya's famous paper [Pólya, 1920] on the "Zentraler Grenzwertsatz" was originally planned as a critical reaction to von Mises's 1919 paper. One of us is about to publish this interesting exchange between these two pioneers of modern probability theory, as documented in von Mises's papers; see [Siegmund-Schultze, 2006].

27 Von Mises Papers, Harvard University Archives, HUG 4574.5, Box 3, Folder 1935. The letters from Pólya are in handwritten German.

28 This is the earliest reference we know of to Pólya's *How to Solve It* [Pólya, 1945] which only appeared a decade later, published in 1945 by the Princeton University Press (although not in German, as originally planned). In an interview in 1977 Pólya said the *How to* book had a special appeal for American readers and that the English version differed considerably from the German, and, as he said "I think to its advantage" [Pólya, 1977, p. 253].

29 The poem, whose authorship is unknown, gives a sense of the social context in which von Mises's satire appeared, and its likely impact on the mood of Jewish intellectuals such as von Mises and Pólya:

> Intellekt! Hinweg mit diesem Wort, dem bösen
> Mit seinem jüdisch grellen Schein!
> Nie kann ein Mann von deutschem Wesen
> Ein Intellektueller sein!

30 Not to be confused with the much more famous philosopher Herbert Marcuse (1898–1979). The two were not related.

31 A virtually identical statement (signed "m") also appears in a short note in "Das Neue Tage-Buch" (volume 3, 1935, p. 1150), a Parisian weekly published before the war by German exiles (and to which Marcuse was a frequent contributor). The note drily ends by suggesting that von Mises's result might result in probability theory being labelled as non-Aryan ("Dies Resultat dürfte der Wahrscheinlichkeitstheorie das Prädikat 'nicht-arisch' eintragen")!

6. Von Mises's views on race

Does von Mises's satirical article mean that he was free of the prejudices of his time? Does it mean that he did not believe in differences of any kind among human races? There is no evidence that von Mises ever believed in the superiority of some human races over others, a position which would have made him a "racist" in the contemporary and already then disputed meaning of the word, exemplified for instance by Nazi racism. There is, however, evidence that von Mises did share the then widely held belief in the differing capacities of human races. In a 1945 correspondence with Otto Neurath, an old acquaintance of von Mises from the days of the Vienna Circle, von Mises expressed opinions not dissimilar to those of, say, some contemporary members of the English eugenics movement, while Neurath's position seems closer to that of many modern readers. (Excerpts from this correspondence appear below in the Appendix.) The correspondence between Neurath and von Mises reminds us of the elusive and labile nature that the concept of "race" has exhibited over time, and the scientific and sociological debates regarding it that continue up to the present day (some modern scholars even rejecting the very notion of race itself).[32] It would appear that von Mises could not fully liberate himself from the particular prejudices of his own time, while Neurath perhaps did.[33]

Acknowledgments

The authors are grateful to Bernard Bru for many helpful comments and suggestions. Thanks go to Mrs. Magda Tisza for permitting the use of the von Mises Papers at the Harvard University Archives and to the Rijksarchief in Noord-Holland for permission to quote the correspondence between Neurath and von Mises. Persi Diaconis and the anonymous referees of the paper also made a number of helpful comments and suggestions.

Appendix. Von Mises and Neurath on race in their final correspondence

From the correspondence kept at Rijksarchief Noord Holland, Otto Neurath Papers, folder no. 268.

Note: Although German was the native tongue of both von Mises and Neurath, the two wrote in English. Neurath, who was writing from England, died December 22, 1945, 6 weeks after his last surviving letter to von Mises, quoted here. The use of italics reflects capitalization in the original letters.

Von Mises (Harvard University), July 16, 1945 to Neurath:

> Dear Dr. Neurath,
>
> ... I read with great interest all your Plato criticism... My own point of view is different... I think that not everything Plato says is good, but neither is everything evil that Nazism stands for. That human races exist and are clearly distinguishable with respect to their achievements is to me a plain fact. How to handle this situation in a really humane way is very difficult to say. To be sure, intermarriage is not the satisfactory answer. At least not in all cases. I also cannot see why one should not try to develop certain characteristics in human beings in the same way as growers of roses do. If you like planned economy why not planned breeding? Perhaps a thousand years from now, one will produce for each generation the required number of medics, of business men, of abstract thinkers etc. And do you know that Popper-Lynkeus who certainly had no leaning towards racism recommended the killing of new-born babies if there are too many of them? (Birth control was unknown to him)...
>
> Very cordially yours [signed] R. v. Mises

[32] For two very different views, see [Lewontin, 1974; Edwards, 2003].

[33] It is interesting to note that in 1956, Felix Bernstein, another emigrant Jewish mathematician, argued in a letter to the *New York Times* (June 24, 1956, p. E8) that there were three types of mathematical aptitude: geometrical, algebraic-combinational, and purely logical; and that these were not uniformly distributed geographically. Bernstein's thesis, however, was not that one type of aptitude was superior to the others, but that a student's particular type or types of aptitude should be taken into account in teaching mathematics.

Neurath (Oxford) to von Mises, 10 November 1945 [from a copy in the Neurath papers]:

... I do not think of Plato and the Nazis as collection of single ideas, plans, attitudes but of both as of certain patterns. And the pattern is the following: there are certain higher authorities, who decide the fate of human beings, they are thinking of some "artistic" product, which may be characterized by *nation, justice, race, beauty, fatherland, truth* or something like that...

There are various of your remarks I should like to see a little more elaborated. What has Nazism to do with *race* as such? Only with *race* as something for which one has to sacrifice something... If Plato says one has to sacrifice human beings for getting a state structure able to fight the Foreigners who are of another blood, then we are confronted with the same attitude, without the Race theory, of which Plato had no idea. He is sacrificing people for the Beauty, the Justice etc. without merci, friendship, brotherhood. *That is the point*...

That happiness of each single individual has to be put into account is just what I always praised in Popper-Lynkeus. And I think we agree in this point. But even Popper has his Platonic elements—unfortunately. E.g. where he speaks of killing the new-born children of mothers with a lot of them... [Neurath adds, Popper was informed about contraceptive medicine, contrary to von Mises's assumption]...

You are writing "I cannot see why one should not try to develop certain characteristics in human beings"... Who is this "one," the master planner or who?... I care [not] a damn for the production of highly efficient business men, abstract thinkers etc.... but for the production of good friends, kind lovers, nice parents, nice children... Do you think people happier, when specialists—that is a rather Platonic idea, as if occupation as such is something wonderful... Why should a weak man's happiness count less than a strong man's happiness. Please I do not ask you to have this attitude yourself. I do not ask Plato or the Nazis to change their attitudes, I only tell... I should not like to live in a Nazi state even when there were not just Belsen camps...

I do not know of what you are thinking seriously, when speaking of the *race* problem. As far as I can see, we have a *political problem* of this kind—of course, as we once had a *witchcraft problem*... but I could not find any scientifically sound material which shows me that the race problem in human society has *biological* or *efficiency* importance. I searched for that. The Nazi literature which specialized in this field is simply dreadful, not because I dislike their habits and attitudes but also because their scientific procedure is under any decent level. I have books and articles on that. Mainly blab-blab.

Did you ever read *Race, Reason, and Rubbish* by Gunnar Dahlberg, London Allen and Unwin 1942? A useful book. That different nations have different success in history is not doubted because it is, as you say a "plain fact," but that this is connected with any hereditary qualities—we have not a grain of proper evidence... The detailed analysis of "tests" etc. show us e.g. that in the USA the Negroe-abilities are not sufficiently fixed up to now. Something depends apparently upon the selection of tests. If one uses white-children habits as a start the black will not have the best results and vice versa. Impressive that North Negroes often are "better" than South Whites, which indicates an environment element... But I do not say anything definite about that, I only do not know from where you got your definite opinion. Intermarriage as a *political* solution is open to discussion of course, as a *biological* measure—I do not know anything of value, which tells me what happens, when human beings of different skin pigment marry...

References

Bayes, T., 1764. An essay towards solving a problem in the doctrine of chances. Philosophical Transactions of the Royal Society of London for 1763 53, 370–418.

Bieberbach, L., 1934. Stilarten mathematischen Schaffens. Sitzungsberichte Preussische Akademie der Wissenschaften, Physikalisch-Mathematische Klasse 1934, 351–360.

Cournot, A.A., 1843. Exposition de la théorie des chances et des probabilités. Libraire de l'Hachette, Paris.

Diaconis, P., Freedman, D., 1990. On the uniform consistency of Bayes estimates for multinomial probabilities. Annals of Statistics 18, 1317–1327.

Edwards, A.W.F., 2003. Human genetic diversity: Lewontin's fallacy. Bioessays 25, 798–801.

Fisher, R.A., 1935. The logic of inductive inference. Journal of the Royal Statistical Society Series A 98, 3954.

Fisher, R.A., 1973. Statistical Methods and Scientific Inference, third ed. Hafner, New York.

Freedman, D., Pisani, R., Purves, R., 1997. Statistics, third ed. Norton, New York.

Gumbel, E.J., 1935. Les valeurs extrêmes des distributions statistiques. Annales de l'Institut Henri Poincaré 5, 115–158.

Hald, A., 1998. A History of Mathematical Statistics from 1750 to 1930. Wiley, New York.

Heyde, C.C., Seneta, E., 1977. I.J. Bienaymé: Statistical Theory Anticipated. Springer-Verlag, New York.

Khinchin, A.Ya., 1936/1944. R. Mises' Frequentist Theory and the Modern Ideas of the Theory of Probability. Published posthumously in the original Russian by B.V. Gnedenko in 1961, and in English translation in 2004 by O.B. Sheynin, with commentary by R. Siegmund-Schultze, in Science in Context 17, 391–422.

Lewontin, R.C., 1974. The Genetic Basis of Evolutionary Change. Columbia Univ. Press, New York.

Marcuse, L., 1960. Mein Zwanzigstes Jahrhundert. Auf dem Weg zu einer Autobiographie. Paul List Verlag, München.

Mehrtens, H., 1987. Ludwig Bieberbach and 'Deutsche Mathematik'. In: Philips, E.R. (Ed.), Studies in the History of Mathematics. Mathematical Association of America, Washington, DC, pp. 195–241.

Mises, R. von, 1919. Fundamentalsätze der Wahrscheinlichkeitsrechnung. Mathematische Zeitschrift 4, 1–97.

Mises, R. von, 1928. Wahrscheinlichkeit, Statistik und Wahrheit. Springer-Verlag, Wien.

Mises, R. von, 1934. Problème de deux races. Matematichesky Sbornik 41, 359–374. In the original, a Russian translation follows immediately on pages 374–389. We quote and translate here from the French version in von Mises's *Selecta* [1963/1964, vol. II, pp. 192–209].

Mises, R. von, 1941. On the foundations of probability and statistics. Annals of Mathematical Statistics 12, 191–205 [pp. 340–355 of *Selecta*, vol. II].

Mises, R. von, 1942. On the correct use of Bayes' formula. Annals of Mathematical Statistics 13, 156–165 [pp. 360–370 of *Selecta*, vol. II].

Mises, R. von, 1951. Positivism, a Study in Human Understanding. Harvard Univ. Press, Cambridge, MA.

Mises, R. von, 1957. Probability, Statistics, and Truth. Allan & Unwin, London (second English edition, following the third German edition (1951) of [Mises, 1928]).

Mises, R. von, 1963/1964. Selected Papers of Richard von Mises, 2 vols. American Mathematical Society, Providence, RI [= *Selecta*].

Mises, R. von, 1964. Mathematical Theory of Probability and Statistics. Edited and Complemented by Hilda Geiringer. Academic Press, New York/London.

Pearson, K., 1921. The fundamental problem of practical statistics. Biometrika 13, 1–16.

Pólya, G., 1920. Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem. Mathematische Zeitschrift 8, 171–181.

Pólya, G., 1945. How to Solve It. Princeton Univ. Press, Princeton, NJ.

Pólya, G., 1977. George Pólya. Interviewed by G.L. Alexanderson. In: Albers, Donald J., Alexanderson, G.L. (Eds.), Mathematical People. Profiles and Interviews. Birkhäuser, Boston, pp. 247–253 (1985).

Remmert, V., 2004. Die Deutsche Mathematiker-Vereinigung im "Dritten Reich": Fach- und Parteipolitik. Mitteilungen der Deutschen Mathematiker-Vereinigung 12, 223–245.

Segal, S., 2003. Mathematicians under the Nazis. Princeton Univ. Press, Princeton, NJ.

Sheynin, O., 2003. Mises on Mathematics in Nazi Germany. Historia Scientiarum 13, 134–146.

Siegmund-Schultze, R., 1998. Mathematiker auf der Flucht vor Hitler. Quellen und Studien zur Emigration einer Wissenschaft. Vieweg, Braunschweig/Wiesbaden. To appear in extended English translation, Princeton Univ. Press, 2007.

Siegmund-Schultze, R., 2004. A non-conformist longing for unity in the fractures of modernity: Towards a scientific biography of Richard von Mises (1883–1953). Science in Context 17, 333–370.

Siegmund-Schultze, R., 2006. Probability in 1919/20: The von Mises–Pólya-controversy. Archive for History of Exact Sciences 60, 431–515.

Von Plato, J., 1994. Creating Modern Probability. Cambridge Univ. Press, Cambridge, MA.

Zabell, S.L., 1991. Ramsey, truth, and probability. Theoria 57, 211–238.