Lewis-Parker Lecture

Mathematicians in Industry: Experiences of One High-Tech Migrant Worker

By Jack B. Brown

1. Introduction

I[1]want to thank the AACTM for inviting me to give the 2001 Lewis-Parker Lecture. I did not have the pleasure of knowing Dr. Lewis, but this opportunity is especially meaningful to me because of the impact Dr. Parker had on my professional career. Dr. Parker created the mathematics program at Auburn University essentially single-handedly, and I have been a beneficiary of his efforts for almost thirty-four years. He was not Department Head when I went to Auburn in 1967, but he was Dean of the Graduate School during my first five years on the faculty. He and Mrs. Parker came to many departmental parties during that period and for years after he retired. I was very fond of him.

I enjoyed reading the biographies of both Dr. Lewis and Dr. Parker which were published in the Fall 1990 issue of the *Alabama Journal of Mathematics* when the establishment of the Lectureship was announced. I especially liked Ann Parker Battle's statement

[1]This is the text of the 2001 Lewis-Parker Lecture which the author presented to the 51[st] annual meeting of the Alabama Association of College Teachers of Mathematics. Because of the nature of the talk (a description of some of the author's personal experiences working in industry) the paper is written in first person narrative style, which is the style in which the lecture was given. The author thanks the Editors of the *Alabama Journal of Mathematics* and the referee for many helpful suggestions that greatly improved this article.

[1] about what Dr. Parker said soon after he had established the first Ph.D. program at Auburn. Concerning his determination not to lower his standards for entrance to Graduate School, she quoted him as saying "If we let someone enter that is unqualified, someone else may let him graduate."

Before beginning my remarks, I will warn you that this Lewis-Parker Lecture will be somewhat different from previous lectures. Almost without exception, they have been survey articles (similar to [2]), outlining an important area of pure or applied mathematical research in which the speaker had made numerous original contributions. However, I thought the AACTM might actually prefer to hear another kind of lecture. Thus, I will discuss some of the applied mathematical problems I have encountered during my 25 summer appointments (plus one sabbatical year) at various applied governmental and industrial research centers around the United States, including Sandia National Laboratory, Bell Labs, TRW, the National Security Agency, and the Centers for Communications Research and Computing Sciences of the Institute for Defense Analyses.

This talk is the current version of a discussion I presented to Pi Mu Epsilon at Auburn University[2] every three years or so over the last twenty years. My purpose in doing this was twofold. First, I wanted to give applied mathematics majors some idea of the kind of work they might end up doing after graduation. The second purpose was directed toward providing some motivation (and hopefully generating some enthusiasm) concerning the courses in analysis, probability, statistics, differential equations, linear and abstract algebra, and other courses required in their applied mathematics degree programs. The degree requirements in our undergraduate Applied Mathematics degree plan were rather extensive, bordering on oppressive, in my opinion. During their first two years, the Applied Mathematics majors were required to take four quarters of calculus, courses in linear algebra and ordinary differential equations, plus a programming course (FORTRAN or C). Then they were required to take year-long courses in each of (1) Abstract Algebra, (2) Real Analysis, (3) Numerical Analysis, and (4) Probability and Mathematical Statistics. On top of this, they had to take at least four additional mathematics courses (≥ 20 quarter hours). I usually recommended they choose from among the following: more probability and statistics, partial differential equations, more ordinary differential equations, vector calculus, complex variables, more linear algebra, and Fourier analysis. They had to take

[2]In recent years, this talk has also been given to student-faculty groups at other universities (e.g. Washington and Lee University, University of Louisville, and Mississippi State University).

a minor consisting of at least five courses ($\geq$ 25 quarter hours) in some applied discipline. I always recommended they minor in physics because that was my undergraduate minor, but they seldom followed my recommendation. Many of the students who were pursuing the possibility of a career in Actuarial Science chose to take their minor in Business Administration, which was appropriate.

I have to admit that I certainly didn't undertake such an extensive plan of study for my own undergraduate degree. I took no abstract or linear algebra courses, although I learned a good bit of linear algebra in the year-long course in ordinary differential equations I took after completing the beginning ordinary differential equations course. I took a year-long course in probability followed by a year-long course in mathematical statistics, as well as a semester course on these topics taught in the Actuarial Science Department at the University of Texas.

As I proceed through a discussion of some specific problems I have had the opportunity to work on during my (primarily) summer appointments during the past 40 years, I will also describe how the computing capabilities available to me have evolved over the years.

2. Sandia and Bell Laboratories Appointments (1960s and 70s)

My first technical appointment came during the Summer of 1961 when I was a beginning masters student. I reported to Sandia National Laboratory in Albuquerque, NM, for the first of three summer appointments in the Quality Assurance Department there. I always tell the students that I didn't get this job on the "first try." I had applied at a number of laboratories and aerospace companies the year before without success. I persevered and hit paydirt on my second annual campaign. In the intervening year, I had taken the year-long course in mathematical statistics, the probability-statistics course in the Actuarial Science Department at the University of Texas, and had made a fairly high mark on the probability-statistics exam administered by the Society of Actuaries. These credentials are probably what helped me "make the cut" on my second try. When I reported to Sandia, the security clearances of the summer hires were not completed yet, and they set us all up with desks in a barracks-like building outside the security perimeter until the clearances were finished. My supervisors brought me the mathematical statistics book by Hoel [3] and suggested I start working my way through it in preparation for reporting to the Quality Assurance Department. I told them that I had just taken a course out of that book and, indeed, had the book with me[3]. After ascertaining that I had not covered every section in the book, they suggested I work through the rest of the book. My security clearance was finished pretty quickly, and I reported to work. I will not go into detail about specific problems I worked on during those first three summers. They involved the standard statistical inference techniques (hypothesis testing, construction of confidence intervals, regression analysis, etc.) we teach in the undergraduate courses in probability and mathematical statistics. I had to do my own computing of the statistics associated with weapons component test data collected in the field. My "computer" was a Monroe desk calculator. Hey, at least it was electric.

After the third summer in the Sandia Quality Assurance Department, I stayed at the University of Texas for several summers in order to take certain graduate courses that were offered only in the summer and to begin work on my Ph.D. dissertation. I took the measure theory course taught by Prof. R. L. Moore during one of the summers and was lucky enough to be in the movie [4] the MAA filmed in that class. After my dissertation was "in progress," I returned to Sandia for my last summer as a graduate student and for the summer after receiving my Ph.D. This time I was assigned to the Statistics and Computing Division, which was one of the Lab's applied consulting units. This was a "dream job" for someone like me because of the wide variety of interesting problems that were brought to the Division by engineers and scientists from divisions all over the Labs. Most of the staff of my Division were Masters or Ph.D. statisticians. Many of the problems assigned to these people involved more theoretical mathematics than they preferred, and so many of these problems were passed on to me. The computational facilities at Sandia were much more modern by this time, but I was "insulated" from them. This was in the days when analysts were not necessarily expected to be able to program a computer, and I had a programmer who was assigned to assist me in transforming any ideas I had about solving some problem into computational form.

One of the first problems I was assigned was that of finding "parametric tolerance limits" for a "mixed" probability distribution. In undergraduate probability and mathematical statistics courses, we concentrate on the two main types of distributions, those which we label as "discrete-type" and those we label as "continuous-type." Indeed, most modern texts (including the text by Hogg and Tanis [5] which I used in the probability course I taught last semester) are constructed so that when you open the

[3]The point I am making to the students here is that they did not necessarily expect me to already know how to do the work I was going to be assigned.

book, you see on the inside front cover a list of the most important discrete distributions and their parameters, and on the facing page to the right you see a list of the most important continuous-type distributions and their parameters. Most of these texts have **one** section devoted to "mixed distributions," where the cumulative distribution function (or "c.d.f."), $F(x)$, is a convex combination, $F(x) = \varepsilon F_d(x) + (1-\varepsilon)F_c(x)$, (where $0 < \varepsilon < 1$) of a discrete c.d.f., $F_d(x)$, and a "continuous-type" (more accurately, absolutely continuous) c.d.f., $F_c(x)$. In undergraduate mathematical statistics courses, we routinely teach the students how to obtain what is called an "upper confidence bound" on the mean of an underlying distribution, given observations, $x_1, x_{2,} \ldots, x_n$, of a random sample from a normal distribution having unknown mean, μ, and variance, σ^2. A "95% upper confidence limit on μ" is a number $b = g(x_1, x_{2,} \ldots, x_n)$ for which you can say, "I'm 95% sure that $\mu \leq b$" (this statement has a precise technical interpretation). This is a standard statistical technique based on using the so-called t-distribution. Note that μ is the same thing as the 50^{th} percentile, $\pi_{.50}$, when one is dealing with normal distributions. One example of a "tolerance limit" would be a similar confidence bound on some other percentile. For example, a 95% upper tolerance limit on $\pi_{.90}$ is a number $B = h(x_1, x_{2,} \ldots, x_n)$ for which you can say, "I'm 95% sure that $\pi_{.90} \leq B$." This procedure is based upon a not-so-standard statistical technique using the so-called non-central t-distribution. The theory was developed by Wald and Wolfowitz [6] in 1946. The problem posed to me was to find out how you find an upper 95% tolerance limit for a distribution which is a "perturbed" normal, i.e. having a "mixed" c.d.f. satisfying the formula given above, where $\varepsilon > 0$ is relatively small, F_s is ε times the so-called "Heaviside" distribution having a jump discontinuity of size 1 at $x = 0$ and F_c is a normal c.d.f., $\Phi\left[(x-\mu)/\sigma\right]$. I might note that the problem was not described to me in these terms, but was instead described to me in terms of probability densities, where the discrete part was assumed to have the so-called Dirac delta function, $\delta(x)$, as its "density."

My main reason for even including this problem in my discussion is to provide the students with some motivation for taking the section on mixed distributions in their text books seriously and to suggest to the teachers of such mathematical statistics courses that maybe we should not skip that section. I did not provide the customer with a complete analytic solution to his problem, but I did provide him with a technique that I could prove was "asymptotically correct" (as $n \to +\infty$). I should also mention that in preparing for this talk, I logged onto the American Mathematical Society's Mathematical Reviews data base, *MathSciNet*, and fed it the phrase "tolerance limit." It came back with a list of 111 references, including the 1946 Wald-Wolfowitz paper. It is conceivable that someone has worked out a technique for finding tolerance limits for such mixed distributions which are exact for every sample size n.

The next problem I am going to describe is one of the most interesting I have come across in my 40 years of (part-time) applied work. It involves the probability distribution of what might be called a "random geometric series." It is described as follows. Suppose we are given an infinite sequence $X_0, X_1, \ldots$ of independent random variables, each having a Bernoulli$\left(\frac{1}{2}\right)$ distribution. In other words, assume we can toss a coin infinitely many (independent) times and if on the i^{th} toss, we get a head, then $X_i = 1$, otherwise $X_i = 0$, so that $P(X_i = 1) = P(X_i = 0) = \frac{1}{2}$ for each i. We are also given a number $0 < r < 1$ and the random variable, Y_r is defined to be $Y_r = X_0 + rX_1 + r^2X_2 + \ldots$. It is clear that the range of the random variable Y_r is a subset of $\left[0, \frac{1}{1-r}\right]$, and that in the case where all of the $X_i = 1$, $Y_r = \frac{1}{1-r}$. Of course, this event has probability zero in the random case. What is actually going on is that there is a secret sequence of zeros and ones. Another person is trying to guess the terms of the sequence and every time he is correct, that $X_i = 1$ for him and he is building up his version of Y_r. I know what the secret sequence is and I can force all of the X_i's to equal 1 so that my version of $Y_r = \frac{1}{1-r}$.

The electrical engineer who brought the problem to us had designed an electronic gadget that essentially contained a component that behaved like Y_r and he wanted us to provide him with a computer tabulation of the c.d.f., F_r, for Y_r. He had already modelled the distribution with its normal approximation and some presumably more accurate approximations, which he started to tell me about. I asked him not to tell me any more so that I would not be influenced by knowledge of his approach as I tried to figure out how I would approach the problem myself. It is fairly easy to figure out what the mean and variance have to be and that, when $r = \frac{1}{2}$, Y_r has the continuous uniform distribution on the interval $[0, 2]$. Shortly after figuring that out, I was able to see that when $r < \frac{1}{2}$, F_r is a **continuous singular** distribution, indeed that when $r = \frac{1}{3}$, F_r is **the Cantor distribution!!!** I cannot tell you how excited I was when I figured this out. You see, when we teach the undergraduate probability-mathematical-statistics sequence, we tell the students about discrete distributions, distributions of continuous type (i.e. absolutely continuous distributions), and sometimes, about distributions which are mixtures of these two types. We do not tell them at that level about the third type of probability distribution, the continuous singular type. I rushed to my boss's office to tell him about this. After I told him what I had found, he said, "I know you mathematicians, you just don't like the Dirac delta function." I replied that while he was correct that I do not like the Dirac delta function, this distribution did not involve that because its c.d.f., F, was actually continuous but its derivative, F', was equal to zero almost everywhere, so that $\int_{-\infty}^{\infty} F'(t)\,dt = 0$. To this he replied, "There is no such thing as that." Let me point out at this stage that my boss was a Ph.D. statistician, this was the second summer I had worked for him, and we had become (and still are) very good friends. Anyway, I started going into detail about what continuous singular distributions were like and described the Cantor distribution in detail. For the x's in the middle third interval, $\left[\frac{1}{3}, \frac{2}{3}\right]$, that is removed at the beginning of the construction of the Cantor set, you set $F(x) = \frac{1}{2}$. You then connect with straight lines between $(0,0)$ and $\left(\frac{1}{3}, \frac{1}{2}\right)$ and again between $\left(\frac{2}{3}, \frac{1}{2}\right)$ and $(1,1)$, and you have the first approximation to the Cantor function. Then, you modify this construction by making $F(x)$ for the x's in the next two "take-out" segments in the construction of the Cantor set be equal to $\frac{1}{4}$ and $\frac{3}{4}$, respectively, connect the gaps so as to create a polygonal function and you have the second approximation to the Cantor function. Continuing this process yields a sequence of continuous c.d.f.'s which converge uniformly to the continuous Cantor distribution. We then computed the sum of the lengths of the "take-out" segments and saw that it equaled 1, so that the Cantor distribution is singular. At the end of this discussion, my boss said, "But there can't be such a thing." I said, "Wait a minute, I was watching you nod your head affirmatively all the time we were going though this. How can you still say there can't be such a thing?" He said, "Because if there were such a thing, I would have known about it!" That stumped me for a while. Then I told him that I would bet that he **did** know about continuous singular distributions, but that he had just forgotten about them because this was the first time he had come across one in a practical situation. I asked him what text he had used in the most advanced theoretical probability course he had taken in his graduate program. It was the classic, "Feller, Volume II" [7], which he actually had on the shelf in his office. I asked him to get it down and look in the Index under "Cantor," and that directed us to the page where Feller started discussing the continuous singular case. Indeed, there was a discussion of the fact that the $Y_{\frac{1}{3}}$ described above has the Cantor distribution (with the scale on the axis changed so that the c.d.f. achieves value 1 at $y = 1.5$ rather than $y = 1$). I have to say that that hour-long discussion with my boss at Sandia was one of the most enjoyable hours of my professional career.

I should point out that I do not talk about continuous singular distributions when I teach undergraduate probability-mathematical-statistics courses, but I do talk about the Cantor function when I teach the undergraduate course in analysis. Of course, the engineer who brought us the problem didn't know (or care) anything about singular distributions or the Cantor function, he just wanted to know what the numbers were. In particular, he wanted to know what some of the percentiles were at the extreme right tail of the c.d.f. F_r for some values of r close to 1. My solution for the engineer was obtained as follows. First, I showed that F_r satisfies the functional equation

$$F_r\left[t\right] = \frac{1}{2}F_r\left[\frac{t}{r}\right] + \frac{1}{2}F_r\left[\frac{t-1}{r}\right] \quad .$$

Then, I saw that if $r = \left(\frac{1}{2}\right)^{\frac{1}{n}}$ for some positive integer n, F_r has closed form

$$F_r\left[t\right] = \frac{t^n}{n!2^{\frac{n+1}{2}}} \quad \text{for} \quad 0 \le t \le \frac{1}{r} \quad .$$

F_r has symmetry about the median, $\frac{1}{2(1-r)}$, so finding an early percentile, say $\pi_{.001}$, will determine the late one, say $\pi_{.999}$, for example. The closed form expression gets you started, and then you can use the functional equation to extend computationally. With my programmer's help, I was able to provide the engineer with the kind of information he needed, using the exact distribution, F_r, rather than some kind of approximating distribution.

After leaving Sandia that summer to report to my new job on the faculty of Auburn University, I investigated a very interesting open problem concerning the nature of the c.d.f., F_r. The problem was to distinguish between cases where the distribution is absolutely continuous and cases where it is continuous-singular, in particular, to characterize the $\frac{1}{2} < r < 1$ for which F_r is singular. It seemed at first like F_r should be absolutely continuous for all $\frac{1}{2} < r < 1$, but that turned out not to be the case. It is a well-known theorem of Lebesgue (see [7]) that **every** c.d.f. is a convex combination, $F = \alpha F_d + \beta F_c + \gamma F_s$, of a discrete c.d.f., F_d, an absolutely continuous c.d.f., F_c, and a continuous singular c.d.f., F_s, where $\alpha \ge 0$, $\beta \ge 0$, $\gamma \ge 0$, and $\alpha + \beta + \gamma = 1$. It follows from a theorem of Jessen and Wintner [8] that the c.d.f., F_r is either purely absolutely continuous (i.e. $\beta = 1$) or else purely continuous singular (i.e. $\gamma = 1$). Erdös showed in 1939 [9] that F_r is singular if $r = 1/\lambda$, where λ is a "PV-number." In particular, F_r is singular for $r =$ "the golden ratio," $\frac{\sqrt{5}-1}{2}$, which is a root of the polynomial equation, $x^2+x-1=0$. Erdös showed in 1940 [10] that F_r is absolutely continuous for almost every $r \in [t,1]$ for some $\frac{1}{2} < t < 1$ and conjectured that the same was true for $t = \frac{1}{2}$. Garsia conjectured in 1962 [11] that F_r is singular only when r is algebraic and satisfies a polynomial equation with all coefficients equal to ± 1 or 0. Solomyak showed in 1995 [12] that F_r is absolutely continuous for almost every $r \in \left[\frac{1}{2}, 1\right]$, thus confirming Erdös's 1940 conjecture. The problem of characterizing the r for which F_r is singular is still open (despite the fact that I tried to solve it).

During my first six years at Auburn, we could count on full-time teaching during the summer, so I did not seek summer employment in industry. But that changed in 1973, when we started getting half-time teaching during the summer. The change did not affect me that year because I was on sabbatical at the University of California Irvine studying probability under the direction of Professor Howard Tucker, and I had already planned to stay at UCI through the summer. However, I decided it was time to return to my "applied career" after that. My best friend at Auburn University, Coke Reed, was going to the IDA Communications Research Division in Princeton for the summer of 1974, so I decided to apply to go to Bell Labs that summer because it is also in New Jersey. I felt there was a chance I could get on there because there was a connection between Sandia Labs and Bell Labs in that Bell used to manage Sandia Labs. I was successful in arranging an appointment for the summer. The work there involved performing nonparametric statistical analysis of a very large set of data concerning customer calling characteristics in California. The problem was fairly interesting and I had to learn quite a bit of nonparametric statistics that came in handy in later years. I actually took another sabbatical to UCI a few years later, during which I had the opportunity to participate in a nonparametric statistics seminar led by Professor Madan Puri (one of the world's experts), who was also on sabbatical there. It is too bad I did not have this seminar experience **before** I went to Bell Labs because I would have been able to accomplish more while I was there. I knew before I went to Bell Labs that I would be expected to do my own programming, and I was prepared to do that. We did our numerical work on an IBM 360, but that was during the days when programs were stamped into punch cards and the card stacks were submitted at the computer center window to be run as batch jobs. There was always some turn-around time before you got the results back and found out there were bugs in your program, and you had to find and correct the bugs and then resubmit. There were usually several cycles like this until the program would finally run. It seems primitive by today's standards but it was standard operating procedure in those times. This was not my favorite summer appointment, but it was not because of the work I was doing. The problem was that Bell Labs Holmdel is in one of the most beautiful areas on the entire eastern seaboard, and as a result, only very rich people are actually able to live there. Consequently, I (and all the other summer hires) had to live in some apartment complex almost 20 miles from the Lab and faced an hour round trip drive every day. This is normal for people who live in metropolitan areas, but not for someone who had been living in a small town in eastern Alabama for seven years. Actually, the best thing I can say about that summer is that I was in a car-pool with three single college girls.

A couple of years later, I decided to try to go back to Sandia Labs partly because my friend Coke Reed had arranged a summer job for himself at Los Alamos Labs. Sandia was originally sort of a "sister lab" to Los Alamos. The physics of nuclear weaponry was developed at Los Alamos and the engineering was done at Sandia. Sandia is now more of a Department of Energy lab. In any case, I was fortunate to be able to arrange an appointment in the Sandia Labs Fuel Cycle Risk Analysis Division that lasted two summers. By this date the computing setup at Sandia was more modern. They had a CDC 6600 which I could access interactively using a fairly fast TI terminal which had some kind of heat sensitive paper that fed through it. I was able to use line editing of programs I was writing, submit a compile command, get diagnostics back immediately, and correct errors in the program on-line. Of course, it was not a CRT connection that allowed page editing like we now have, but it was certainly an improvement over sitting at a punch-card machine and carrying stacks of punch cards to the Computer Center window. I was involved in a Nuclear Regulatory Commission project to study the surface flow of radioactive waste under the assumption that radionuclides started leaking from their underground storage site after a thousand years or so. The surface flow was assumed to be a large compartmental model (called the "Environmental Transport Model") described by a linear vector differential equation,

$$X'(t) \quad = \quad A\,X(t) \quad + \quad R \quad . \tag{1}$$

Each component, $X_i(t)$, of the vector function, $X(t,)$ represents the amount of some particular radioactive isotope we are tracking which is present in some subcompartment (in the stream flow, in the subsurface aquifer, in the soil adjacent to the stream, captured in the sediment in the bottom of the stream) of one of the compartments (a uniform stretch of the river, a lake, etc.) in the model. The A in equation (1) is the matrix that represents the flow rates between compartments and the R in the equation is the vector of input rates as the leaking radionuclides begin to reach compartments in the environmental model. As radioactive isotopes decay, they may change into other isotopes which are being tracked and this is considered to be a transfer to a new compartment. The flow rate for this kind of transfer would be constant, of course, but the flow rates between actual environmental compartments depend upon the weather, are seasonally periodic at best, and are more accurately modelled by stochastic functions. These two more complicated cases would be modelled by the following two equations,

$$X'(t) = A(t)\,X(t) + R\ , \tag{2}$$

$$X'(t,\omega) = A(t,\omega)\,X(t,\omega) + R\ . \tag{3}$$

The matrix $A(t)$ of equation (2) would represent the assumed seasonally periodic flow rates between compartments, and the matrix $A(t,\omega)$ of equation (3) would represent the stochastic flow rates that would be driving the flows, and the solution to the problem would be the stochastic process $X(t,\omega)$ that is the solution to the stochastic differential equation (3). The constant matrix, A, which drives the flow in equation (1) can be thought of as the average of the flow-rate matrices in the more accurate equations (2) and (3), i.e.

$$A = \int_0^1 A(t)\,dt = \int_0^1 E\left[A(t,\omega)\right]dt\ .$$

The problem I was assigned was to investigate the question of how large the errors were in assuming model (1) rather than the more accurate models (2) and (3). I obtained some analytical results concerning bounds on the differences between solutions to models (1) and (2) and some empirical results based upon a large Monte Carlo simulation run concerning the differences between the solutions to models (1) and (3).

One unexpected benefit that came out of these two summer appointments in the Fuel Cycle Risk Analysis Division was that I actually got some "outside publications" on this project. This did not happen in very many of my summer appointments because almost every project I worked on was classified, or else the results I obtained were considered "proprietary," and I was unable to attempt to publish them. Of these outside publications [13]-[16], the first two would not have counted for much as far as promotion and tenure at Auburn are concerned because they look too much like "technical reports." However, these were different from the many other technical reports I wrote over the years in that they were NUREG Reports of the Nuclear Regulatory Commission and were publicized and available to researchers in the field all over the world. These days it is easy for researchers to log onto the laboratory web sites and download such NUREG reports. The publication [15] might have carried a little more weight at Auburn because the Auburn University Library at least has volumes of the serial, *Scientific Basis for Nuclear Waste Management*, on the shelves. However, the Auburn University Promotion and Tenure Committee would probably have dismissed this as a "non-refereed" publication. But the publication [16] would have been accepted by the Promotion and Tenure Committee because the journal, *Ecological Modelling*, is a well-respected, refereed scientific journal. However, that journal probably would not have been considered to be a mathematics journal by the faculty in the Department of Mathematics. Thankfully, none of this mattered because I was already a tenured Full Professor at that time. One of my co-authors on those papers was Jon Helton, who was a "migrant worker" like me, making his home base at Arizona State University. Helton graduated from the University of Texas under the same supervising professor (H. S. Wall) under whom I studied. He is younger than I am and I did not know him at Texas, but I took Freshman trigonometry under his father (who was the "Mr. H." mentioned in the Moore movie [4]). My other co-author, Ron Iman, was a permanent staff member in the Statistics and Computing Division at Sandia, where I had worked during my previous appointment at Sandia. Iman recently served a term as President of the American Statistical Association.

3. Defense-Aerospace Appointments (early 1980s)

A summer or two later, I decided to try to go back to Bell Labs for a summer, partly because my friend, Coke Reed, was going back to IDA/CRD in Princeton. I applied and was contacted by a technical group that wanted to hire me. I should explain that when the technical people have decided they would like for you to work for them, the appointment is still not guaranteed. The Personnel Office has to do a lot of paperwork, the Business and Finance Office has to do a lot of paperwork, and at installations where a security clearance is required, the Security Office has to arrange for the clearance (this is often the biggest hold up). Things dragged on for another month or so with no firm offer. I got a call out of the blue from a friend I knew in graduate school asking me if I would like to come work with his group at a defense aerospace consulting company called the Center For Analysis (CFA), in Newport Beach, California. I promise that I did not initiate this alternative job possibility — I have never applied for two different jobs for a summer because I know how much trouble it is for someone to set up a job for a summer hire. However, it was getting to be late April, and I had a growing family at that time, so I decided to take the job in Newport Beach. The fact that I was going to have a view of the Pacific Ocean out the window in my office and was going to earn the highest salary I had ever earned surely had nothing to do with it. I called the technical group at Bell Labs and told them that because of the delay in finalizing the Bell Labs appointment, I felt I had to accept this **unsolicited** offer I had received from a company in Southern California. The people in the technical group at Bell Labs were furious, not at me but at their own Personnel and Finance Offices that had been dragging their feet on my appointment. So, I packed up my family and headed west.

I worked at the Center For Analysis for three summers and then worked at TRW Strategic Systems Group for another summer when several of my CFA colleagues moved to TRW as project managers. The computing facilities at both of these locations were modern. At both locations, I **finally** had access to CRT terminals from which you could do interactive page-editing of programs, compile programs on the main computer, and receive diagnostics immediately. So, it was possible to produce bug-free code fairly quickly and then to obtain quality computational results with the ease we have come to expect today. The problems I worked on at both installations were similar, involving determining the reliability of infrared discrimination techniques and improving performance, if possible. One of the problems was the following. An "incoming" missile is dropping off objects as it streaks across the sky, and we are to train a telescope with an infrared sensor attached to it on these falling objects. Then we will have a computer analyze certain features of the optical signature that is being observed and make a decision as to whether the particular object we are looking at is a Warhead or a Decoy. We will know that the observations $x_1, x_2, \ldots, x_n$ are of a random vector which has one of two known multivariate normal distributions $N(M_w, \Sigma_w)$ or $N(M_d, \Sigma_d)$, where the M's and Σ's are the known mean vectors and covariance matrices for the two distributions. What was done was essentially what we teach students to do in undergraduate mathematical statistics courses, test the hypotheses

$$H_0 : \text{the underlying distribution is } N(M_d, \Sigma_d)$$

vs

$$H_1 : \text{the underlying distribution is } N(M_w, \Sigma_w).$$

The only difference between this problem and the hypothesis testing we teach at the undergraduate level is that the distributions are multivariate. We still had to consider the possibility of making errors:

Type I Error: reject H_0 when H_0 is true

and

Type II Error: accept H_0 when H_1 is true.

The Type I Error was called "False Alarm." The Type II Error was called "Leakage," and **you really hate it when it happens!** My problem was to determine what variables were best to use to keep both errors small and then write a program so the computer could automatically make decisions very quickly (in "real time"). I was able to make some improvements in the choice of features they had previously been extracting from the optical signature that decreased both error types. I also was able to improve the efficiency of the program they had been using so that the time to decision was decreased.

I would have returned to TRW for another appointment after that fourth summer in Southern California, but I took on the job of Head of the Auburn University Department of Mathematics. My colleagues at TRW arranged for the USAF Ballistic Missiles Office to keep me involved through a contract with Auburn University so that I could continue my work at Auburn the following summer. But of course, being Department Head, I did not really have time to do the job justice.

4. U. S. Governmental Cryptologic Work (mid-1980s to present)

Since the mid-1980's, I have done my applied work at three locations, the Center for Communications Research of the Institute for Defense Analyses (IDA/CCR) (seven summers), the IDA Center for Computing Sciences (one summer), and the National Security Agency (one year, five summers, consulting). This section of this paper will be somewhat short because most of the work I did at these places was highly classified and I cannot talk about it. There are publications that give some information about the nature of the work, the oldest and most extensive being the well-known book, *The Code Breakers* [17], by David Kahn ("CCR" is called "CRD" in this book). Another source is the article, "The Agency That Came in From the Cold," which appeared in the 1992 *A.M.S. Notices* [18]. It is a text of a lecture that Dr. Richard Shaker (who was Chief of Mathematical Research at NSA) gave at the national AMS/MAA meeting, held in Baltimore in 1992. I was at that lecture, and it marked a significant change in NSA policy concerning secrecy of its activities. Before that time, I would not have even been able to say that I worked at NSA.

When I spent the 1993-94 academic year at NSA under the auspices of the NSA Mathematical Sabbatical Program (see [19] or log onto `http://www.nsa.gov/programs/msp/sabbat.html` for information about that program), I was invited to give a version of this talk at Washington and Lee University. I had never referred to NSA in earlier versions of this talk, so I went to Dr. Shaker's office and asked him if I could say anything. He said I could say anything that had been published in the open literature. In particular, I could quote freely from his article. A more recent source of information is a documentary film which was shown on the History Channel's "History's Mysteries" Series in January of 2001. An article about that documentary which appeared in my local newspaper at the time quoted the NSA Director, Lt. Gen. Michael V. Hayden, as saying, "We intercept communications of adversaries of the United States and attempt to turn that into wisdom for American policy-makers and commanders. By the same token we attempt to prevent other nations from doing that to the United States of America."

It probably comes as no surprise to anyone that the computing facilities at IDA/CCR, IDA/CCS, and (especially) NSA were without equal. I had a SPARC workstation on my desk at NSA from which I could connect to more powerful computers when I needed them. I don't know how many CRAY computers I had access to — I would not be able to tell you, even if I did know — but I had accounts on three different types of CRAYs because of the nature of the project I was working on at the time. It is well-known that NSA is the cryptologic arm of the US Department of Defense. The work is highly classified and I cannot say anything specific about it. What I **can** say is that the single most personally satisfying professional experience I have had in my entire career (pure or applied) occurred while I was working on that classified project at NSA in 1993-94.

Occasionally, an unclassified problem comes up which is of interest to NSA and I did spend two summers working part time on such a problem. It concerned "factor convergence" and "order independence" in a numerical algorithm used to "balance" nonnegative matrices. The algorithm was used by the famous statistician Deming, et. al. in 1940 [20], who raised the question of convergence of the algorithm. Statisticians call the algorithm "iterative proportional fitting." An engineer named D. T. Brown [21] gave an incorrect argument for the convergence in 1959. Engineers and information theorists sometimes call the algorithm "Brown's algorithm." The mathematician, R. Sinkhorn, gave an incorrect proof of convergence in 1964 [22] and then Sinkhorn and Knopp [23] gave a correct proof in 1967. They were actually preceded in doing this by an economist named Bacharach [24] who also gave a correct proof of convergence in 1965, but not many researchers knew about this paper because it was not reviewed by the *AMS Reviews*. The algorithm is called "Sinkhorn Balancing" or "Iterative Scaling" in the Linear Algebra community. The work that two of my NSA colleagues and I did was published [25] — this time in a refereed mathematics journal!

5. The "Bottom Line" — Keep the Employers Happy

I save the next problem for last because it provides a special lesson for students. It was a problem I solved during my appointment in the Statistics and Computing Division at Sandia Labs in the late 1960s. It was definitely not the hardest, nor the most interesting problem I worked on there, but it is probably what I am remembered for at Sandia. It is a problem of determining the minimum sample size that needs to be tested in order to make a predetermined confidence statement about the reliability of the items in a finite population, under the assumption that no defective items are found in the sample. We actually teach students in undergraduate mathematical statistics courses how to do this using normal approximations to the Binomial distribution. However, it is possible to do this exactly without using the normal approximation and I usually teach students how to do this. For example, it can be determined that if we test a random sample of size $n = 45$ from a Bernoulli(p) distribution, where $p = P\,(\text{defective})\,,$ $q = 1 - p = P\,(\text{functional})\,,$ and we observe zero defectives in the sample, then we can say "We are 90% sure the population is 95% functional," which is equivalent to saying that we have determined that .95 is a 90% lower confidence bound on the true proportion, q, of functional items in the population being tested. The Reliability and Quality Assurance Departments at Sandia Labs were making these sample size selections **constantly** because they were required by Department of Defense regulations to do sampling and run tests on a multitude of weapons systems and components of those systems on a periodic basis in order to verify readiness. They did this so often that one of the applied mathematics divisions at Sandia designed and produced a "slide-rule" that could be used to quickly determine the choice of sample size that would yield the desired percentage confidence in the desired percentage functionality in the case that zero defectives were observed in the sample being tested (which was almost always the case). However, these choices are based upon the assumption that the items tested are chosen independently from the populations so that the binomial distribution applies. That is not the true situation. The "population" is a **finite** set of N systems or components in storage somewhere and some number, K (hopefully zero), of them are defective. When the testing is done, we would choose the items to be tested randomly, but we would not be so foolish as to test the same item twice. It is really the hypergeometric distribution which should be used in determining the sample size to be tested rather than the binomial distribution. I thought that this should surely yield some advantages. After carefully studying exactly what a lower confidence bound on the number of functional items actually is, I determined that there would be great advantage in using the hypergeometric distribution. For example, if the finite population is of size $N = 80$, K of which are defective, and we test a sample of size $n = 35$ and observe zero defectives in the sample, we can correctly say that "We are $\geq$90% sure the population is $\geq$95% functional." You might ask, "What's the big deal about a sample size of $n = 35$ versus a sample size of $n = 45$?" They were running these tests on multiple systems and systems components each year. It cost thousands of dollars to run some of these tests, so the smaller the sample size required to achieve $\geq$90% confidence in functionality $\geq$95%, the less money the tests cost. Or, to put it another way, they were paying me a relatively minuscule fee of less than \$5,000 for the whole summer, thereby saving a great deal of money.

This is the kind of thing that will make you an asset in the applied mathematics field. Saving money, or making money, is always a goal. In the routine course of summer employment, these types of challenges often present themselves, but sometimes they may not be obvious and thus they need some insight. Whereas one can hope to be challenged mathematically, and that has been most of the subject of this paper, I think that an additional reason why I was able to go back to Sandia quite a few summers is that I also had the good fortune to contribute to the Laboratory's financial health. As I look back over the years, I can say very definitely that the wide variety of interesting problems of my mathematics experience have given me a great deal of professional satisfaction on many levels.

References

[1] Anne Parker Battle, William Vann Parker, *Alabama Journal of Mathematics* (Fall 1990), 7–10.

[2] J. B. Brown, Restriction theorems in real analysis, *Real Analysis Exchange* 20 (1994), 520–526; MR 96f:46059.

[3] P. G. Hoel, *Introduction to Mathematical Statistics*, (1954), Wiley, New York.

[4] *The Moore Method, A Documentary on R. L. Moore*, Mathematical Association of America Video Classics, Number 3, Copyright 1966.

[5] R. V. Hogg and E. A. Tanis, *Probability and Statistical Inference*, 5th Ed., (1999), Prentice Hall, Upper Saddle River, NJ.

[6] A. Wald, A. and J. Wolfowitz, Tolerance limits for a normal distribution. *Ann. Math. Statistics* 17, (1946). 208–215; MR 8,478a.

[7] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. II, (1950), Wiley, New York.

[8] B. Jessen and A. Wintner, Distribution functions and the Riemann zeta function, *Trans. Amer. Math. Soc.* 38 (1935), 48–88; Zbl 014.15401.

[9] P. Erdös, On a family of symmetric Bernoulli convolutions, *Amer. J. Math.* 61 (1939), 974–976; MR 1, 52a.

[10] P. Erdös, On the smoothness properties of a family of Bernoulli convolutions., *Amer. J. Math* 62 (1940), 180–186; MR 1, 139e.

[11] A. M. Garsia, Arithmetic properties of Bernoulli convolutions, *Trans. Amer. Math. Soc.* 102 (1962), 409–432; MR 25 # 1409.

[12] B. M. Solomyak, On the random series $\Sigma \pm \lambda^n$ (an Erdös problem), *Ann. of Math.* (2) 142 (1995), 611–625; MR 97d:11125.

[13] J. B. Brown and J. Helton, Risk methodology for geologic disposal of radioactive waste: Effects of variable hydrologic patterns on the environmental transport model, NUREG/CR-1636, v. 4 (1981), 105 pgs.

[14] J. C. Helton, J. B. Brown, R. C. Iman, Risk methodology for geologic disposal of radioactive waste: Asymptotic properties of the environmental transport model, NUREG/CR-1636, v. 3 (1981), 173 pgs.

[15] J. Helton, R. Iman, J. Brown, S. Schreurs, Geologic waste disposal and a model for the surface movement of radionuclides, *Scientific Basis for Nuclear Waste Management*, v. II (1980), 795–802.

[16] J. Helton, R. Iman, and J. Brown, Sensitivity analysis of the asymptotic behavior of a model for the environmental movement of radionuclides, *Ecological Modelling* 28 (1985), 243–278.

[17] David Kahn, *The Codebreakers : The Story of Secret Writing*, Scribner, 1996, New York.

[18] R. J. Shaker, The agency that came in from the cold, *Notices Amer. Math. Soc.* 39 (1992), no. 5, 408–411.

[19] NSA Grant and Sabbatical Programs, *Notices Amer. Math. Soc.* 48 (September 2001), 836.

[20] W. E. Deming and F. F. Edwards, On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat.* 11 (1940) 427–444; MR 2,232g.

[21] D. T. Brown, A note on approximations to discrete probability distributions, *Information and Control* 2 (1959), 386–392; MR 22 #1474.

[22] R. Sinkhorn, A relationship between arbitrary positive matrices and doubly stochastic matrices, *Ann. Math. Stat.* 35 (1964), 876–879; MR 32 #7574.

[23] R. Sinkhorn and P. Knopp, Concerning nonnegative matrices and doubly stochastic matrices, *Pacific J. Math.* 21 (1967), 343–348; MR 35 #1617.

[24] M. Bacharach, Estimating non-negative matrices from marginal data, *Internat. Econ. Rev.* 6 (1965), 294–310.

[25] J. B. Brown, P. J. Chase, and A. O. Pittenger, Order independence and factor convergence in iterative scaling, *Lin. Alg. Appl.* 190 (1993), 1–38; MR 94m:60006.

Department of Mathematics
340 Parker Hall
Auburn University, AL 36849
brownj4@auburn.edu