MOSCOW MATHEMATICAL JOURNAL
Volume 10, Number 4, October–December 2010, Pages 807–810

PROBABILITY PROBLEMS OF DOBRUSHIN TYPE

LARRY SHEPP

To my friend Roland

Abstract. I review five probability problem areas which I believe are in the spirit of my friend, Roland. Were he still around, he would shed light on them in his inimitable, insightful way. The novelty of this note is that it ties together a lot of my special papers in a way that perhaps only Roland would grasp; even though I cannot see the commonality. Anyone who wants to comment is more than welcome to do so: shepp@stat.rutgers.edu

2000 Math. Subj. Class. 60G60.

Key words and phrases. Probability, math.

I. Connectivity of Random Graphs

This topic is the one most closely connected to Dobrushin's work in phase-transition-physics. Make a random graph on the integers $\{0, 1, 2, \ldots\}$ by regarding the integers as the vertices and placing an edge between i and j with probability $p_{i,j} = \frac{c}{i+j}$ and no edge between i and j with the complimentary probability, $1-p_{i,j}$, independently for each pair, i, j. The constant c might represent a reciprocal temperature of a one-dimensional crystal where there are bonds between atoms. At high temperatures there would be fewer bonds; at lower temperatures more bonds. Thanks to work of Kalikow and Weiss [9], of Shepp [12], and of Durrett and Kesten [3], it is known that if $p_{i,j} = p(i, j)$, where $p(\lambda x, \lambda y) = \lambda^{-1} p(x, y)$, is homogeneous of degree minus one, then there is a sharp transition in the probability that the graph is *connected*, or has only one component. In the special case posed above, the transition is at $c_0 = \frac{1}{\pi}$, so that for $c < c_0$ the probability that the graph is connected is zero, while for $c > c_0$ the graph is connected with probability 1. The problem has been solved for only one dimensional graphs, and physicists do not consider the model "realistic" because of the one-dimensionality of it, and even more so because $p_{i,j}$ is not *homogeneous* (they would like it to be only a function of $(i - j)$. Alas, if $p_{i,j} = p_{i-j}$ in the one-dimensional case, then Kalikow and Weiss showed there is no "multiplicative" phase transition; they showed that the graph

Received January 15, 2010.

is connected if and only if

$$\sum_{n=-\infty}^{\infty} p_n = \infty,$$

and so if $p_{i,j}$ is replaced by the multiple $cp_{i,j}$, there will be no critical value of c. The problem I would like to get Roland to think about is what happens when one considers higher dimensions. Perhaps then there would be a way to get the desirable *multiplicative* phase transition and somehow also manage to retain homogeneity. Perhaps also this is too tall an order; some Roland-like insight is needed here.

II. Random Growth Models

Let seeds be placed at the points of a Poisson random set on $\mathcal{R} = (-\infty, \infty)$. Each seed grows at unit rate until it runs into another seed and then both seeds stop growing. What fraction of the line is covered when growth finally stops? By ergodicity, this is the same as $1 - p$, where p is the probability that when growth stops, the origin is *not* inside a seed. The probability, p, turns out to be $\frac{1}{e}$ by a theorem of Daley, Mallows and Shepp, [1]. Huffer [6] and also Winkler showed that the same answer, $\frac{1}{e}$, holds even if the seeds grow at random rates independent of each other. Again, it would be nice to be able to generalize to higher dimensions [2] but the "curse of dimensionality", would seem to defeat us. But, then again, read the next problem area.

III. Random Covering Models

A. Dvoretzky [4] realized that if one centers an arc of length l_n, $n \geqslant 1$, independently for each n, at a random uniform point on a circle of length one, then the condition that $\sum_{n=1}^{\infty} l_n = \infty$ may not guarantee that the entire circle is covered with probability 1, although it is true that each fixed point of the circle is covered with probability 1. Of course the fact that the circle is an uncountable set is the explanation of the apparent paradox. Eventually, Kahane [8], and Shepp [11], showed that the divergence of another series, namely,

$$\sum_{n=1}^{\infty} \frac{1}{n^2}\, e^{l_1+\cdots+l_n} = \infty$$

is necessary and sufficient for a.s. covering assuming only that the l_n decrease monotonically. Thus if $\{l_n,\, n \geqslant 1\}$ is the sequence $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \ldots\}$, then the series converges, so there is positive probability that the union of the arcs fails to cover some point, and this is also true if $\{l_n,\, n \geqslant 1\}$ is the sequence, $\{\frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \ldots\}$, but if $\{l_n,\, n \geqslant 1\}$ is the union of these two sequences, $\{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\}$, then coverage takes place with probability 1.

For an analogous version of this problem in higher dimensions, astonishingly enough, it is possible to generalize this result! If the circle is replaced by a torus of volume one, and if the arcs are replaced by *simplices*, scaled to have volume l_n, Fan and Kahane [5] gave the exact condition for a.s. covering as the divergence of an infinite series involving the sequence, $\{l_n,\, n \geqslant 1\}$. They showed that the series is different for each dimension. The problem remains open for general convex sets in place of simplices. It seems likely that the exact condition will depend on the shape of the convex set. The question for Roland is why is this problem different from the other problems in the sense that it is not limited by the "curse of dimensionality"?

In higher dimensions one can also ask whether it is possible to find a convex shape and a scaling sequence, l_n, for which every hyperplane of dimension k is covered with probability 1, but this does not hold for hyperplanes of dimension $k-1$. This problem is also open; I would conjecture the answer is yes, without much conviction.

IV. Zeros of Random Polynomials

Let $P_n(t) = \sum_{k=0}^{n} \xi_k t^k$ denote a random polynomial with independent identically distributed coefficients with a given distribution F. Much is now known about the number, ν_n, of real roots, [7], [10]. If ξ's have zero mean and finite variance, then $E\nu_n \sim \frac{2}{\pi}\log n$. If the ξ's are in the domain of attraction of the stable law with parameters, α, β, then $E\nu_n \sim c(\alpha,\, \beta)\log n$. The maximum value of $c(\alpha,\, \beta)$ occurs at the point $\beta = 0$, $\alpha = 0^+$, where the limiting value of $c(\alpha,\, \beta) = 1$, among all values of α, β. The question now arises as to which distribution, F, would produce this *maximal expected number of real zeros*. Since $c(\alpha,\, 0)$ decreases from one at $\alpha = 0^+$ to $\frac{2}{\pi}$ at $\alpha = 2$, and since the smaller the value of the parameter α, the fatter are the tails of the symmetric stable law with parameter α, one might believe that the fatter the tails of the ξ's, the more real zeros would be obtained. The stables have tails, $1 - F(a) \sim x^{-\alpha}$, as $x \to \infty$. However, a recent paper, [13], gives evidence against this. It shows there is an F with *still fatter tails*, $1 - F(x) \sim (\log x)^{-a}$, as $x \to \infty$, yet, for $a > 1$,

$$E\nu_n \sim \frac{a-1}{a}\log n$$

as $n \to \infty$, and the mean number of real zeros decreases as the tails get fatter. This is very strange, because as the tails get fatter, starting from say a normal r.v., the mean number of real zeros, first increases to a maximum of $\log n$, asymptotically, and then decreases as the tails get fatter than the stables. This seems to shed light on Descartes's rule of signs. On the basis of this result, I conjecture strongly that the expected number of real zeros for an arbitrary distribution F is asymptotically at most $c\log n$, with $c = 1$. Much remains to be done to obtain this result, if it is actually true.

V. The Distribution of a Gaussian Functional

Some (inferior) statisticians have convinced themselves that one can claim that two processes, $W_1(t)$, $W_2(t)$, $0 \leqslant t \leqslant 1$, are not independent if their correlational statistic,

$$\theta = \frac{\int_0^1 W_1(t)W_2(t)\,dt - m_1 m_2}{S_1 S_2},$$

where

$$m_i = \int_0^1 W_i(t)dt,\ S_i = \sqrt{\int_0^1 W^2(t)dt - m_i^2}, \quad i = 1,\, 2,$$

is far away from zero. Note that $-1 \leqslant \theta \leqslant 1$ by Schwarz's inequality.

Of course this is false for the case of *independent* standard Browian motions, $W_i(t)$. It may be possible (see [14]) to find the exact distribution of θ in this case and it is far from zero. Finding it is apparently not easy. For other cases of independent, identically distributed processes, $W_i(t)$, the problem remains very open. Perhaps Roland would have had no trouble with this one.

References

[1] D. J. Daley, C. L. Mallows, and L. A. Shepp, *A one-dimensional Poisson growth model with non-overlapping intervals*, Stochastic Process. Appl. **90** (2000), no. 2, 223–241. MR 1794537

[2] D. J. Daley, H. Stoyan, and D. Stoyan, *The volume fraction of a Poisson germ model with maximally non-overlapping spherical grains*, Adv. in Appl. Probab. **31** (1999), no. 3, 610–624. MR 1742684

[3] R. Durrett and H. Kesten, *The critical parameter for connectedness of some random graphs*, A tribute to Paul Erdős, Cambridge Univ. Press, Cambridge, 1990, pp. 161–176. MR 1117011

[4] A. Dvoretzky, *On covering a circle by randomly placed arcs*, Proc. Nat. Acad. Sci. U.S.A. **42** (1956), 199–203. MR 0079365

[5] A. H. Fan and J.-P. Kahane, *Rareté des intervalles recouvrant un point dans un recouvrement aléatoire*, Ann. Inst. H. Poincaré Probab. Statist. **29** (1993), no. 3, 453–466. MR 1246642

[6] F. W. Huffer, *One-dimensional Poisson growth models with random and asymmetric growth*, Methodol. Comput. Appl. Probab. **4** (2002), no. 3, 257–278 (2003). MR 1965408

[7] M. Kac, *On the average number of real roots of a random algebraic equation. II*, Proc. London Math. Soc. (2) **50** (1949), 390–408. MR 0030713

[8] J.-P. Kahane, *Some random series of functions*, D. C. Heath and Co. Raytheon Education Co., Lexington, Mass., 1968. MR 0254888

[9] S. Kalikow and B. Weiss, *When are random graphs connected*, Israel J. Math. **62** (1988), no. 3, 257–268. MR 955131

[10] B. F. Logan and L. A. Shepp, *Real zeros of random polynomials. II*, Proc. London Math. Soc. (3) **18** (1968), 308–314. MR 0234513

[11] L. Shepp, *Covering the circle with random arcs*, Israel J. Math. **11** (1972), 328–345. MR 0295402

[12] L. Shepp, *Connectedness of certain random graphs*, Israel J. Math. **67** (1989), no. 1, 23–33. MR 1021358

[13] L. Shepp and K. Farahmand, *Expected number of real zeros of a random polynomial with independent identically distributed symmetric long-tailed coefficients*, to appear in Teor. Veroyatn. Primen.

[14] L. Shepp and A. Wyner, *The spurious empirical correlation of independent sum processes*, In preparation.

Statistics Department, Rutgers University, Piscataway, NJ
E-mail address: shepp@stat.rutgers.edu