On a statistical approach to Bertrand's problem

Verbinden wir (4.3) und (4.4) mit den Einsätzen (2.1) und (2.2), so resultiert die Behauptung (1.10).
Im verbleibenden Fall $V \leq \omega_k a^k$ lässt sich in analoger Weise auf (1.10) schliessen, womit der Beweis für das Hauptergebnis erbracht ist.

Marcel Iseli, Oberwangen[1])

LITERATURVERZEICHNIS

1 H. Hadwiger: Über die Flächeninhalte ebener Schnitte konvexer Körper. El. Math. *30*, 97–102 (1975).
2 H. Hadwiger: Vorlesungen über Inhalt, Oberfläche und Isoperimetrie. Berlin 1957.
3 T. Bonnesen und W. Fenchel: Theorie der konvexen Körper. Berlin 1934.

On a statistical approach to Bertrand's problem

Dedicated to Prof. H. Hadwiger on the occasion of his 70th birthday

Probability theory is the branch of mathematics which constructs and analyzes mathematical models for random phenomena. In the early stages of the examination of such a phenomenon there are usually different mathematical models conceivable, each one of them representing a possible explanation of our empirical observations. Statistical methods may then be applied to evaluate the appropriateness of these models and to help decide which one of them should be adopted. It is the purpose of this note to demonstrate on a particular example the usefulness of statistical methods for the treatment of geometrical stochastics, i.e. of their applications to the analysis of randomly generated stets.
The example to be discussed is the following: Denote by E^2 a Euclidean plane with an origin Z and a system of Cartesian $x-y$-coordinates, by $G(q, \theta)$ $[0 \leq q < \infty, 0 \leq \theta < 2\pi]$ the straight line in E^2 whose distance from Z is q and whose angle, formed by the perpendicular on G through Z and the x-axis, is θ, and by C_1 and C_2 the two circle-lines in E^2 with the same centre Z and radius 1 and $1/2$, respectively. Let $\binom{Q}{\Theta}$ be a random vector whose components take values in the above specified ranges according to a density function h. It is of interest to calculate

$$_hP := \Pr\left(G(Q,\Theta) \cap C_2 \neq \emptyset \mid G(Q,\Theta) \cap C_1 \neq \emptyset : h\right)$$

the probability that the random straight line $G(Q,\Theta)$ intersects C_2 conditionnally on it intersecting C_1. The result is given by the formula

$$_hP = \left[\int_0^{1/2}\int_0^{2\pi} h(q,\theta)\,d\theta\,dq\right] \cdot \left[\int_0^{1}\int_0^{2\pi} h(q,\theta)\,d\theta\,dq\right]^{-1} \tag{*}$$

1) Bemerkung der Redaktion: Der Verfasser dieser Note, ein hochbegabter Schüler von Prof. Hadwiger, ist am 12.5.78, etwas mehr als ein Jahr nach seiner Diplomierung, nach schwerer Krankheit gestorben.

and depends on the choice of h. In the literature three stochastic models are proposed which have some intuitive appeal and which amount to the following choices of h:

Model A: $h_A(q,\theta)=\pi^{-1}q$,
Model B: $h_B(q,\theta)=\pi^{-2}(1-q^2)^{-1/2}$ $[0\leq q\leq 1, 0\leq\theta<2\pi]$,
Model C: $h_C(q,\theta)=(2\pi)^{-1}$.

It is well known and easy to derive from (*) that

$$_{h_A}P=\frac{1}{4}, \quad _{h_B}P=\frac{1}{3}, \quad \text{and} \quad _{h_C}P=\frac{1}{2}.$$

Failure to understand that these results are the solutions of three different problems produces the so-called 'paradoxon of Bertrand'. In particular adoption of the model C is equivalent to making the assumption that the probability law of $G(Q,\Theta)$ is induced by the integral geometric kinematic density of straight lines in E^2. For further indications on the history of the problem see [2, 5] and on integral geometry see [4, 7].

Although the probabilistic aspects of calculating $_hP$ are nowdays well understood, there remains the problem of deciding which model actually to use in each concrete situation. In the following we present a comprehensive solution of the problem of pairwise comparison of the models A, B and C by optimal statistical tests extending earlier remarks made in [9].

Suppose a realization $G(q_1,\theta_1),\ldots,G(q_n,\theta_n)$ of a random sample of n straight lines has been observed and that we choose one of the models as null hypothesis and another as alternative hypothesis. The table (see p. 136, 137) characterizes most powerful tests of size a $[0<a<1]$ (i.e. tests with a probability of committing an error of the first kind smaller than a and with minimal probability of committing an error of the second kind among all tests satisfying this restriction). The hypotheses are listed in the columns (1a) and (1b). According to the fundamental theorem of Neyman and Pearson [6, 10] these tests depend on the random sample only in terms of the likelihood-ratio statistics, which are given in column (2) for the six different possible cases. Instead of using these statistics we may use equivalent formulations of the decision rules, which are based on the random quantities

$$U_n:=-\sum_{i=1}^{n}\ln(Q_i), \qquad V_n:=-\sum_{i=1}^{n}\ln(1-Q_i^2), \qquad W_n:=U_n+\frac{1}{2}V_n.$$

In column (3) is specified which one of these statistics is relevant in any particular case and in column (4) the corresponding critical regions are characterized, i.e. it is indicated under which condition the null hypothesis is rejected. In order to apply the tests, knowledge of the critical values is required. In this context we note that U_n, V_n and W_n are sums of independently identically distributed random variables. It thus follows that for $n\rightarrow\infty$ these quantities are asymptotically normally distributed provided that the corresponding mean-values and variances are finite [8]. More

Statistical analysis of Bertrand's problem.

Hypotheses Model under null hypothesis	Model under alternative hypothesis	Likelihood-ratio statistic	Equivalent statistic	Critical region
(1a)	(1b)	(2)	(3)	(4)
A	B	$\pi^{-n}\left\{\prod_{i=1}^{n}(Q_i[1-Q_i^2]^{1/2})\right\}^{-1}$	W_n	$w_n > W_n(1-a)$
A	C	$2^{-n}\left\{\prod_{i=1}^{n}Q_i\right\}^{-1}$	U_n	$u_n > U_n(1-a)$
B	A	$\pi^{n}\left\{\prod_{i=1}^{n}(Q_i[1-Q_i^2]^{1/2})\right\}$	W_n	$w_n < W_n(a)$
B	C	$\left(\frac{\pi}{2}\right)^{n}\left\{\prod_{i=1}^{n}([1-Q_i^2]^{1/2})\right\}$	V_n	$v_n < V_n(a)$
C	A	$2^{n}\left\{\prod_{i=1}^{n}Q_i\right\}$	U_n	$u_n < U_n(a)$
C	B	$\left(\frac{2}{\pi}\right)^{n}\left\{\prod_{i=1}^{n}([1-Q_i^2]^{1/2})\right\}^{-1}$	V_n	$v_n > V_n(1-a)$

Generally $T_n(a)$ (resp. $T_n(1-a)$) denotes the left-hand (resp. right-hand) critical value of the statistic T_n of size a under the corresponding null hypothesis. $z(t)$ is defined by the relation

$$(2\pi)^{-1/2}\int_{-\infty}^{z(t)}\exp\left[-\frac{1}{2}u^2\right]du=t \quad [0<t<1].$$

or less involved calculations [for an example see the appendix of this paper] yield the moments shown in the columns (5) and (6) of the table. The variance of W_n is equal to $n\,\sigma_{BA}^2$, if model B is valid. σ_{BA}^2 is defined by the expression

$$\sigma_{BA}^2:=\pi^2[6]^{-1}-2[\ln(2)]^2+(2\pi^{1/2})^{-1}\sum_{\nu=1}^{\infty}\left(\nu^{-1}\cdot\Gamma\left(\nu+\frac{1}{2}\right)\cdot[\Gamma(\nu+1)]^{-1}\right.$$
$$\left.\cdot\left[\psi(\nu+1)-\psi\left(\nu+\frac{1}{2}\right)\right]\right)$$

with

$$\psi(t):=\frac{d\ln(\Gamma(t))}{dt}$$

and Γ denoting the gamma function; its numerical value is 0.819. The optimal procedure for sufficiently large n is thus easy to implement. It consists in comparing the standardized statistics with the critical value $z(a)$ resp. $z(1-a)$ of the

Expected value of (3) under model (1a)	Variance of (3) under model (1a)	Critical region for large n
(5)	(6)	(7)
n	$n\left(1-\frac{\pi^2}{12}\right)$	$\left[n\left(1-\frac{\pi^2}{12}\right)\right]^{-1/2}[w_n-n]>z(1-a)$
$\frac{n}{2}$	$\frac{n}{4}$	$n^{-1/2}[2u_n-n]>z(1-a)$
$2n\ln(2)$	$n\sigma^2_{BA}$	$[n\sigma^2_{BA}]^{-1/2}[w_n-2n\ln(2)]<z(a)$
$2n\ln(2)$	$n\frac{\pi^2}{3}$	$\left[n\frac{\pi^2}{3}\right]^{-1/2}[v_n-2n\ln(2)]<z(a)$
n	n	$n^{-1/2}[u_n-n]<z(a)$
$2n(1-\ln(2))$	$n\left(4-\frac{\pi^2}{3}\right)$	$\left[n\left(4-\frac{\pi^2}{3}\right)\right]^{-1/2}[v_n-2n+2n\ln(2)]>z(1-a)$

normal distribution with mean 0 and variance 1. Tables of such values are given in most books on statistics.

Appendix

Determination of the expectation $E[W_n]$ and $Var[W_n]$ of W_n under validity of the model B:

$$E[W_n]=E[U_n]+\frac{1}{2}E[V_n]$$

$$=-2n\pi^{-1}\int_0^1[\ln(q)]\,[1-q^2]^{-1/2}dq-n\pi^{-1}\int_0^1[\ln(1-q^2)]\,[1-q^2]^{-1/2}dq$$

$$=n\ln(2)-n(2\pi)^{-1}\left[\Gamma\left(\frac{1}{2}\right)\right]^2\left[\psi\left(\frac{1}{2}\right)-\psi(1)\right]$$

$$=2n\ln(2).$$

For the transformations the formulae (6.3.2) and (6.3.3) of [1], p. 258, and the formulae [4.241 (7)] and [4.253 (1)] of [3], p. 535, 538, were used.

$$\operatorname{Var}[W_n]=\operatorname{Var}[U_n]+\frac{1}{4}\operatorname{Var}[V_n]+\mathrm{E}[U_n V_n]-\mathrm{E}[U_n]\mathrm{E}[V_n]$$

$$=2n\pi^{-1}\int_0^1[\ln(q)]^2[1-q^2]^{-1/2}dq-n[\ln(2)]^2$$

$$+n(2\pi)^{-1}\int_0^1[\ln(1-q^2)]^2[1-q^2]^{-1/2}dq-n[\ln(2)]^2$$

$$+2n(\pi)^{-1}\int_0^1[\ln(q)][\ln(1-q^2)][1-q^2]^{-1/2}dq-2n[\ln(2)]^2$$

$$=n\pi^2[6]^{-1}-2n[\ln(2)]^2$$

$$+n(2\pi^{1/2})^{-1}\sum_{\nu=1}^{\infty}\left(\nu^{-1}\Gamma\left(\nu+\frac{1}{2}\right)[\Gamma(\nu+1)]^{-1}\left[\psi(\nu+1)-\psi\left(\nu+\frac{1}{2}\right)\right]\right)$$

$$=n\sigma_{BA}^2.$$

For the transformations the formulae (6.3.2), (6.3.3), (6.4.2), (6.4.4) and (23.2.24) of [1], p. 258, 260, 807, and the formulae (4.256), [4.261(17)] and [4.261(21)] of [3], p. 539, 541, were used.

F. Streit, Genève

REFERENCES

1 M. Abramowitz and I.A. Stegun: Handbook of Mathematical Functions. Dover, New York 1970.
2 J. Bertrand: Calcul des probabilités, p. 4–5. Gauthier-Villars, Paris 1907.
3 I.S. Gradshteyn and I.W. Ryzhik: Table of Integrals, Series and Products. Academic Press, New York 1965.
4 H. Hadwiger: Vorlesungen über Inhalt, Oberfläche und Isoperimetrie, chap. 6. Springer, Berlin 1957.
5 M.G. Kendall and P.A.P. Moran: Geometrical Probability, p. 9, 10. Griffin, London 1963.
6 E.L. Lehmann: Testing Statistical Hypotheses, p. 63–68. Wiley, New York 1966.
7 L.A. Santaló: Integral Geometry and Geometric Probability. Addison-Wesley, Reading 1976.
8 L. Schmetterer: Einführung in die mathematische Statistik, p. 127. Springer, Wien 1966.
9 F. Streit: On Methods and Problems of Geometrical Stochastics. Bull. ISI *46* (2), 600–605 (1976).
10 B.L. van der Waerden: Mathematische Statistik, p. 251–258. Springer, Berlin 1965.

Kleine Mitteilungen

Eine Asymptotenkonstruktion der Hyperbel

Den allgemein bekannten affinen Eigenschaften der Hyperbel[1]) sei mit diesem Beitrag noch eine weitere hinzugefügt, die zur raschen Konstruktion der Asymptoten einer Hyperbel herangezogen werden kann. Diese Eigenschaft lautet:

1) Siehe Literaturangabe.