

A HISTORY OF & PROBABILITY & STATISTICS

And Their Applications before 1750



Anders Hald

Wiley Series in Probability and Mathematical Statistics
—Vic Barnett, Ralph A. Bradley, J. Stuart Hunter,
Joseph B. Kadane, David G. Kendall, Adrian F. M. Smith,
Stephen M. Stigler, and Geoffrey S. Watson: Advisory Editors

A History of Probability and Statistics and Their Applications before 1750

ANDERS HALD

Formerly Professor of Statistics
University of Copenhagen
Copenhagen, Denmark



A JOHN WILEY & SONS, INC., PUBLICATION

A NOTE TO THE READER

This book has been electronically reproduced from digital information stored at John Wiley & Sons, Inc. We are pleased that the use of this new technology will enable us to keep works of enduring scholarly value in print as long as there is a reasonable demand for them. The content of this book is identical to previous printings.

Copyright © 2003 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, e-mail: permreq@wiley.com.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representation or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services please contact our Customer Care Department within the U.S. at 877-762-2974, outside the U.S. at 317-572-3993 or fax 317-572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print, however, may not be available in electronic format.

Library of Congress Cataloging-in-Publication is available.

ISBN 0-471-47129-1

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

with being the first to have used this method correctly (see Rabinovitch, 1970b). Previously, first Bernoulli and then Pascal got the credit.

Stifel (1544, 1545) and Scheubelius (1545), both from Germany, used the figurate and binomial numbers in connection with the extraction of roots.

Tartaglia (1556) found that the number of ways n dice can fall is given by a figurate number. Cardano (1570) found the multiplicative form of the figurate numbers and pointed out the correspondence between the combinatorial and the figurate numbers.

Vieta (1591) found that the coefficients in his formula for $\sin nx$ in terms of $\sin x$ were figurate numbers.

Briggs (1633) used the figurate and binomial numbers in binomial expansions, in trigonometric calculations like Vieta, and in interpolation formulae.

In 1636 Fermat found the multiplicative form of the figurate numbers and used these numbers in his formula for the summation of power of integers; he communicated his results to Mersenne.

Mersenne (1636, 1637) extended Cardano's table of C_r^n up to $n = 36$ and discussed the three formulae given by Levi ben Gerson. He also gave the formula

$$\frac{n!}{a! b! c! \dots}$$

for the number of permutations of n things of which a things are of one kind, b of another, and so on. Mersenne used these results in connection with the arrangement of musical notes.

No doubt Pascal knew Mersenne's books; Pascal's contributions are thus essentially his proofs and the clear demonstration of the relations between the figurate, binomial, and combinatorial numbers.

5.3 THE CORRESPONDENCE OF PASCAL AND FERMAT AND PASCAL'S TREATISE ON THE PROBLEM OF POINTS

Besides the problem of points, Pascal and Fermat also discussed two rather simple dice games, which we shall discuss first.

In his letter of 29 July 1654, Pascal writes to Fermat on the simplest of de Méré's problems:

If one undertakes to throw a *six* with one die, the advantage of undertaking it in 4 throws is as 671 to 625. If one undertakes to throw a *double-six* with two dice,

there is a disadvantage of undertaking it in 24 throws. And nevertheless 24 is to 36 (which is the number of faces of two dice) as 4 to 6 (which is the number of faces of one die).

Pascal writes that he does not have time to give the solution and that Fermat will easily find it himself. The method for solving this problem had been given earlier by Cardano. It amounts to finding the smallest number of trials, n , say, such that

$$\frac{1 - q^n}{q^n} \geq 1 \quad \text{or} \quad q^n \leq \frac{1}{2}.$$

For $p = 1/6$, one finds $n = 4$ and the odds 671/625 as given by Pascal, and for $p = 1/36$, one gets $n = 25$ and the approximate odds 506/494, whereas $n = 24$ gives 491/509.

Obviously, de Méré's reasoning presupposes that n is proportional to 6^k , $k = 1, 2, \dots$. Since $p = 6^{-k}$ and $(1 - p)^n \cong e^{-np}$, the equation $(1 - p)^n = \frac{1}{2}$ has the approximate solution

$$n = (\ln 2)6^k = 0.6931 \times 6^k, \quad k = 1, 2, \dots,$$

which gives $n = 4.16$ for $k = 1$ and $n = 24.95$ for $k = 2$. Since n must be an integer, the strict proportionality is destroyed. The approximate solution is due to de Moivre (1712), see §14.4.

In an undated letter from Fermat to Pascal, presumably the earliest one of those preserved, they discuss the value of a throw in a dice game in which the player undertakes to throw a 6 (i.e. at least one 6) in eight throws. Fermat says that after the stakes have been made, the player should have 1/6 of the total stake in compensation for not making the first throw. If, after this has been settled, they further agree that the player should not make the second throw, he should have 1/6 of the remaining stake as compensation, i.e. 5/36, of the total stake, and so on. Fermat calls this the *value* of the throw and remarks that this result is in agreement with the one communicated to him by Pascal. The value of the k th throw is thus $(1/6)(5/6)^{k-1}$, $k = 1, 2, \dots, 8$, times the total stake, and by summation the value of the eight throws becomes $1 - (5/6)^8 = 0.767$ times the total stake, which is the compensation the player should have for giving up the play altogether. (This last result is not given by Fermat.) It will be seen that Fermat's value equals the probability of winning times the total stake. Suppose now that the player has made three throws without success. What is then the value of the fourth throw? It seems that Pascal in a previous letter, not longer extant, has given this value as 125/1296, whereas Fermat states that according to his principle, the value is

1/6 of the remaining stake, i.e., 1/6 of the total stake. Fermat stresses that this principle holds no matter the number of unsuccessful throws. Presumably he means that the value of the remaining five throws should be found as for the eight original throws, i.e., as $(1/6)(5/6)^{k-1}$, $k = 1, 2, \dots, 5$, so that the value of the five remaining throws becomes $1 - (5/6)^5 = 0.598$ times the total stake. Hence if the play is interrupted after m unsuccessful throws, the player should have $1 - (5/6)^{8-m}$, $m = 0, 1, \dots, 7$, times the total stake in compensation for giving up the play. Expressed in modern terminology it will be seen that Fermat's reasoning is based on the geometric distribution $p(1-p)^{k-1}$, $k = 1, 2, \dots$, for $p = 1/6$.

Before discussing the correspondence further, we shall give an account of Pascal's *Usage du Triangle Arithmétique pour déterminer les partis qu'on doit faire entre deux joueurs qui jouent en plusieurs parties* in which he solves the division problem, first by recursion on the expectations and next by use of the arithmetic triangle. He always assumes that the two players have an equal chance of winning in a single game.

He begins by stating two obvious division principles (axioms): (1) If a player gets a certain sum, whatever the outcome of the game this sum should be allotted to him; (2) if the total stake falls to the winner with probability 1/2 and the game is not played, the stake should be divided equally between the two players. Formulated in another way we may say that the *value* of the game or the player's *expectation* equals $t/2$ if the total stake equals t .

From these principles he proves Corollary 1: If a player gets the amount s , when he loses, and the amount $s + t$ when he wins, and the game is not played, then he should have the amount $s + t/2$. In the proof he first uses principle (1), which says that the player should first have the amount s , which belongs to him independently of the outcome of the game, and furthermore he should have $t/2$ accordingly to principle (2). In Corollary 2 he simply remarks that the same result may be obtained by adding the two amounts and dividing by 2. It will be seen that Pascal in these corollaries derives the player's expectation for the case where the two outcomes of the game correspond to different prizes, which need not be limited to the total stake and zero as in principle (2).

Like Cardano (see §4.2), Pascal explains that the division rule should depend only on the number of games each player lacks in winning. We shall let $e(a, b)$ denote A's share of the total stake, or A's expectation, if the play is interrupted when A lacks a games and B lacks b games in winning. Pascal begins by discussing six examples for a total stake equal to 8; later he changes to the stake 1, which means that the expectation equals A's probability of winning. Since Pascal keeps strictly to the division problem, he always speaks of the *fraction* of the stake going to A; he does not use the term probability or chance in this context. For convenience we shall usually set the total stake

Table 5.3.1. Pascal's recursive calculation of $e(a, b)$

	a	b	$e(a-1, b)$ $e(a, b-1)$	$e(a, b)$
(1)	0	$n > 0$		1
(2)	n	n		$1/2$
(3)	1	2		
(3a) A wins	0	2	1	$3/4$
(3b) A loses	1	1	$1/2$	
(4)	1	3		
(4a) A wins	0	3	1	$7/8$
(4b) A loses	1	2	$3/4$	
(5)	1	4		
(5a) A wins	0	4	1	$15/16$
(5b) A loses	1	3	$7/8$	
(6)	2	3		
(6a) A wins	1	3	$7/8$	$11/16$
(6b) A loses	2	2	$1/2$	

equal to unity so that $e(a, b)$ represents both A's expectation and his probability of winning; the interpretation will be clear from the context. Pascal usually sets the stake equal to a power of 2 to avoid fractions in $e(a, b)$. Table 5.3.1 gives Pascal's results divided by 8, i.e., A's probability of winning.

The results stated in the first two examples are obvious. In the third and the following examples Pascal considers the two possible outcomes of the next game: Either A wins, or he loses. For each of these cases the expectation is known from the previous calculations, and the final expectation is then found as the average of the two intermediate ones according to Corollary 2. Pascal concludes that the method may be used in general. It will be seen that he uses a conditional argument corresponding to the formula

$$E\{X\} = \{E(X|A) + E(X|\bar{A})\}/2.$$

Using modern notation, Pascal's procedure may be written as

$$\begin{aligned} e(0, n) &= 1 & \text{and} & & e(n, n) &= \frac{1}{2}, & n &= 1, 2, \dots, \\ e(a, b) &= \frac{1}{2}[e(a-1, b) + e(a, b-1)], & (a, b) &= 1, 2, \dots \end{aligned} \quad (1)$$

This is a simple example of a partial difference equation. It also provides a good illustration of the interdependence of mathematics and probability theory. To obtain an explicit expression for $e(a, b)$ the difference equation has to be solved, which Pascal did by means of his results for the arithmetic triangle. In the 18th century Pascal's recursive method became very popular

and helped to solve more difficult problems, which induced by Moivre, Lagrange, and Laplace to develop general methods for the solution of difference equations, as discussed in Chap. 23.

For large values of (a, b) , numerical recursion as shown in Table 5.3.1 is cumbersome. Possibly for this reason Pascal looked for a more expedient solution, which he found by comparing the recursion formula for the expectations with that for the numbers in the arithmetic triangle, noting that they differ only by a factor of $1/2$ and by the boundary conditions. He therefore proposes that

$$e(a, b) = \frac{1}{D_{a+b-1}} \sum_{j=a}^{a+b-1} t_{a+b-1-j, j}, \quad a+b=2, 3, \dots, \quad (2)$$

the numerator being equal to the last b terms of the diagonal sum and the denominator being equal to 2^{a+b-1} . Because of symmetry the last b terms equal the first b so that

$$e(a, b) = \frac{D_{a+b-1, b-1}}{D_{a+b-1}}. \quad (3)$$

Pascal proves this formula by induction. Since it obviously holds for $a+b=2$, he assumes that it holds for $a+b=k$ and proves that it holds for $a+b=k+1$. To find $e(a, k+1-a)$, he uses the recursion formula

$$\begin{aligned} e(a, k+1-a) &= \frac{e(a-1, k+1-a) + e(a, k-a)}{2} \\ &= \frac{D_{k-1, k-a} + D_{k-1, k-a-1}}{2D_{k-1}} \\ &= \frac{D_{k, k-a}}{D_k}, \end{aligned}$$

which follows from (2.7) and (2.10). This concludes the proof.

It will be seen that (2) and (3) may be written as

$$e(a, b) = \sum_{i=a}^{a+b-1} \binom{a+b-1}{i} \left(\frac{1}{2}\right)^{a+b-1} = \sum_{i=0}^{b-1} \binom{a+b-1}{i} \left(\frac{1}{2}\right)^{a+b-1}. \quad (4)$$

Pascal goes on to investigate some special cases. For $a < b$ he finds *the fraction of B's stake going to A*, assuming that both stake the same amount. Obviously, this fraction equals $2[e(a, b) - 1/2]$. By means of (3) and the

properties of the triangular numbers, he finds

$$2 \left[e(1, b) - \frac{1}{2} \right] = 1 - \frac{1}{D_{b-1}} = 1 - \left(\frac{1}{2} \right)^{b-1}, \quad (5)$$

$$2 \left[e(b-1, b) - \frac{1}{2} \right] = \frac{t_{b-1, b-1}}{D_{2b-2}} = \binom{2b-2}{b-1} \left(\frac{1}{2} \right)^{2b-2}, \quad (6)$$

$$2 \left[e(b-2, b) - \frac{1}{2} \right] = \frac{2t_{b-2, b-1}}{D_{2b-3}} = \binom{2b-3}{b-2} \left(\frac{1}{2} \right)^{2b-4}, \quad (7)$$

and he remarks that the last result is two times the last but one, or that A's gain by winning the first point is the same as the gain by winning the second. This means that he is considering *the fraction of B's stake going to A when A is winning a single game or point*. He does not, however, derive the general result

$$2[e(a, b) - e(a+1, b)] = \frac{t_{a, b-1}}{D_{a+b-1}} = \binom{a+b-1}{a} \left(\frac{1}{2} \right)^{a+b-1}, \quad (8)$$

which is easily proved by means of the recursion formula and (3). This result shows that the value of winning a point measured in terms of the opponent's stake equals the binomial probability (8).

Pascal's paper is purely mathematical, and he does not attempt any combinatorial interpretation of his results even though he mentions that the division rule may be found also by combinatorial methods. This topic is discussed in the correspondence with Fermat.

Several letters of the correspondence are missing, among them Pascal's first letter with the formulation of the problem and his solution and Fermat's answer in which he proposes to solve the problem by combinatorics. However, in Pascal's letter of 24 August 1654, he describes in detail Fermat's procedure. Fermat says that if the players lack a and b games, respectively, in winning then the play will be over after *at most* $a + b - 1$ further games. Imagining that these games are actually played, there will be 2^{a+b-1} equally possible outcomes and the cases favorable to A in relation to 2^{a+b-1} will give his fraction of the stake. For $(a, b) = (2, 3)$ there will be $2^4 = 16$ outcomes, which may be listed as AAAA, AAAB, AABA, AABB, ..., BBBB, the letters indicating the winner of each game. Since there are 11 cases favorable to A, Fermat's rule gives $e(2, 3) = 11/16$. In Pascal's letter of 29 July he makes the following comment on this procedure: "Your method is very safe and is the one which first came to my mind in this research; but because the trouble

of the combinations is excessive I have found an abridgement and indeed another method that is much shorter and more neat, which I would like to tell you here in a few words." Pascal then proceeds to give an exposition of his recursive method by means of three examples like those we have already discussed.

By example he also proves (5) and (6). Presumably he was working on his treatise when he wrote the letter because a month later, Fermat in a letter of 29 August acknowledges receipt of a copy of the *Traité*. Pascal's letter is not as clear, systematic, and comprehensive as the *Usage de Triangle Arithmétique*, but it contains the main ideas and a complete table for the case $b = 6, a = 1, 2, \dots, 6$ and a stake of 512. Presumably Pascal has first calculated $e(a, b)$ by recursion; he gives a table of $e(a, b) - 1/2$, and by differencing he gets a table of $e(a, b) - e(a + 1, b)$, i.e., A's share of B's stake for each game won. He does not understand certain features of this table, which indicates that he does not know the general result expressed by Equation (8).

Returning to the discussion of the combinatorial method, Pascal writes that he has communicated Fermat's solution to some colleagues in Paris and that Roberval, professor of mathematics, has objected to it because the division rule depends on games which possibly will not be played. Pascal replies that this is immaterial because the two procedures, the actual and the hypothetical play, will always give the same result for the two players, since the cases in which A wins or loses in the actual play will also be the cases in which he wins or loses in the hypothetical play. This is easily seen from the example above, since in each combination of four letters only one of the players can win, and the same winner comes out whether all four games are played or the play is stopped after the winning score.

In his reply of 25 September, Fermat approves of Pascal's explanation and adds two important remarks to convince Roberval. First he says that it does not affect the division how far the fictitious play is extended over $a + b - 1$ games. If in the example the play is extended to five instead of four games, the effect will be that each of the 16 combinations will be changed to two by adding an A or a B, and this does not change the relative number of cases in which A wins.

Although Fermat illustrates his second remark by an example with three players, we shall use it here in connection with the example above for $(a, b) = (2, 3)$. It is clear that A may win in two, three, or four games, corresponding to the arrangements AA, ABA and BAA, and ABBA, BABA, BBAA, which each contains two A's with an A at the last place. For two games there are 2^2 equally possible outcomes with one favorable to A; for three games there are 2^3 outcomes with two favorable ones; and for four games, 2^4 outcomes with three favorable ones. The chances for A to win

may thus be found as the sum

$$e(2, 3) = \frac{1}{2^2} + \frac{2}{2^3} + \frac{3}{2^4} = \frac{11}{16},$$

in agreement with the result previously found. For the case (a, b) it is clear that A may win at either the a th, the $(a + 1)$ st, ..., or the $(a + b - 1)$ st game. The number of different arrangements of a A's and i B's with an A at the last place equals the number of combinations of $a - 1$ A's and i B's, so that

$$e(a, b) = \sum_{i=0}^{b-1} \binom{a-1+i}{a-1} \left(\frac{1}{2}\right)^{a+i}. \quad (9)$$

Fermat does not give this general result explicitly, but he writes

And this rule is good and general for all cases, so that without recourse to fictitious games, the actual combinations in each number of games give the solution and show what I said at the outset, that the extension to a certain number of games is nothing else than the reduction of various fractions to the same denominator.

Hence, Pascal and Fermat both solved the problem of points by combinatorial methods; Pascal also solved it by recursion and showed that the explicit solution of the difference equation is identical to the combinatorial solution; finally Fermat used what is called today *a waiting-time argument* to find the solution (9), which erroneously has been called Pascal's distribution.

It is thus clear that Pascal solved the problem of points by means of *the binomial distribution* for $p = \frac{1}{2}$. Possibly there will not be general agreement about the statement above that Fermat solved the problem by means of *the negative binomial distribution* (9), as it is called today. This result is usually ascribed to Montmort (see §14.1). The reader should check that (4) and (9) give identical results.

To summarize and clarify the relations between the important concepts and results of the correspondence, we have shown the arithmetic triangle in a random walk diagram in Fig. 5.3.1.

The random walk starts at $(0, 0)$. If A wins a point the path goes one step to the right, and if B wins it goes one step up, the two paths being equally likely. It follows that only the states lying on a diagonal are equally likely and that the triangular numbers (binomial coefficients, number of combinations) give the number of paths from $(0, 0)$ to the state in question, as shown on the k th diagonal. Player A wins the stake if he gets a points

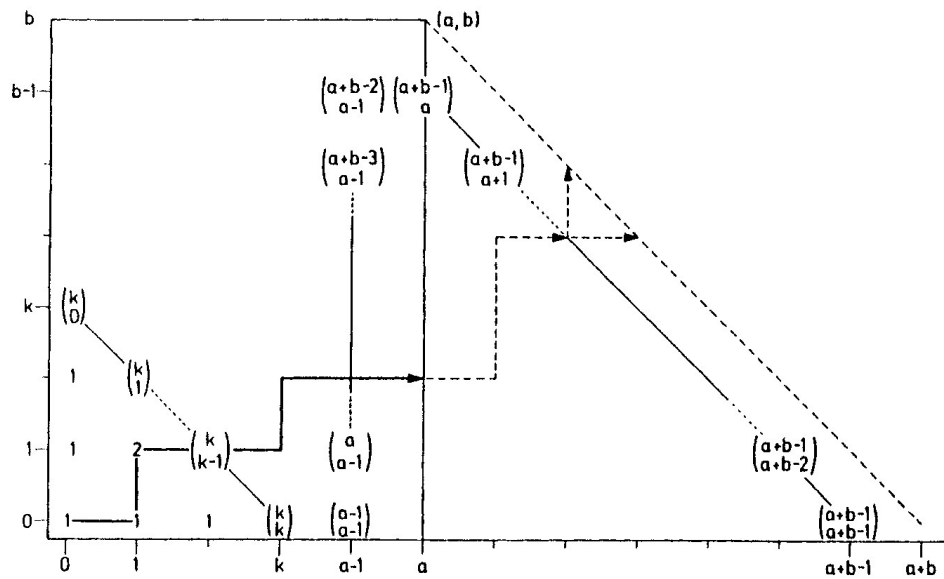


Fig. 5.3.1. Random walk diagram for the problem of points. The number of paths from $(0, 0)$ to (a, b) equals $\binom{a+b}{a}$; A wins if he gets a points before B gets b .

before player B gets b , which means that A wins if the random walk ends at the vertical line from $(a, 0)$ to $(a, b-1)$. These states are, however, not equally likely, and the number of paths can therefore not be added to find the number of chances for A. Consider now a fictitious continuation of the play from the b winning states on the vertical until $a+b-1$ games have been played; i.e., until the random walk ends at the $(a+b-1)$ st diagonal. It is easy to see that the paths going through the winning states and the paths ending at the $(a+b-1)$ st diagonal for the abscissas a to $a+b-1$ are the same. The number of chances for A may therefore be found by adding the number of paths ending at this diagonal, which leads to Pascal's result (2). The dashed part of the path represents the fictitious games, and Fermat's argument about the continuation of the fictitious play has been illustrated by the $(a+b)$ th diagonal. Fermat's result (9) corresponds to the number of paths ending at the vertical from $(a-1, 0)$ to $(a-1, b-1)$ and continuing horizontally to the vertical through $(a, 0)$. The result (8) may also be derived immediately from the diagram by comparing the $(a+b-1)$ st and the $(a+b)$ th diagonals.

The problem of the *division rule for three players*, A, B, and C, lacking 1, 2, and 2 games, respectively, is discussed in Pascal's letter of 24 August, presumably after some preliminary remarks by Fermat in a letter no longer extant. Pascal writes, "When there are only *two* players, your combinatorial