# Thirteen Correct Solutions to the "Problem of Points" and Their Histories

PRAKASH GORROOCHURN

The Problem of Points (POP) is not only the first major problem of probability but it is also the one responsible for its foundation. Indeed, it is one of the questions the Frenchman Antoine Gombaud (1607-1684) (better known as the Chevalier de Méré),[1] posed to Blaise Pascal (1623–1662) in 1654. The latter discussed the problem with his friend Pierre de Fermat (1601–1665). Both men exchanged letters, and through their communication the theory of probability was officially born. A recent book recounting the exchange between Pascal and Fermat is by Devlin [10]. POP had such an impact that almost all major probabilists from then on, from Huygens to Laplace, had a say on it. Even before Pascal, several solutions to POP had been offered by the likes of Pacioli, Peverone, and Cardano. However, these were incorrect and have been well documented elsewhere [e.g., 16, p. 35;18;31, Chap. 1]. The aim of this article is to present thirteen *correct* methods of solution to POP while also briefly discussing their histories.

POP, in its general form for $n$ players, is as follows (POP-n):

*Players $P_1, P_2, \ldots, P_n$ play a game for a sum of money. The game is made up of several rounds such that their respective probabilities of winning one round are $p_1, p_2, \ldots, p_n$, where $\sum_i p_i = 1$. At some point, the players are respectively $s_1, s_2, \ldots, s_n$ rounds short of winning the game when the game suddenly stops. How should the sum of money be divided among the players?*

## Method of Enumeration

Consider the following simple example of POP-2, which we denote by POP-2': Two players A and B play a fair game such that the first player who wins a total of 6 rounds wins a prize. Suppose the game stops when A has won a total of 5 rounds and B has won a total of 3 rounds. How should the prize be divided between A and B?

To solve POP-2', we note that Player A is $s_1 = 1$ round short, and player B $s_2 = 3$ rounds short, of winning the prize. The maximum number of hypothetical remaining rounds is $(1 + 3) - 1 = 3$. In terms of equally likely sample points, the sample space of the game is

$$\Omega = \{A_1A_2A_3, A_1A_2B_3, A_1B_2A_3, A_1B_2B_3, B_1A_2A_3, B_1A_2B_3, B_1B_2A_3, B_1B_2B_3\}.$$

Here $A_1A_2B_3$, for example, denotes the event that A wins the first two remaining rounds and B wins the third. There

[1]Leibniz describes the Chevalier de Méré as "a man of penetrating mind who was both a player and a philosopher" [22, p. 539]. Pascal biographer Tulloch also notes [33, p. 66]: "Among the men whom Pascal evidently met at the hotel of the Duc de Roannez [Pascal's younger friend], and with whom he formed something of a friendship, was the well-known Chevalier de Méré, whom we know best as a tutor of Madame de Maintenon, and whose graceful but flippant letters still survive as a picture of the time. He was a gambler and libertine, yet with some tincture of science and professed interest in its progress."

 
DOI 10.1007/s00283-014-9461-5

are in all eight equally likely outcomes, only one of which ($B_1B_2B_3$) results in B hypothetically winning the game. Player A thus has a probability 7/8 of winning. The division ratio (DR) between A and B should therefore be 7:1.

This method of enumeration is credited to Fermat and is explicitly described in his September 25, 1654, letter to Pascal. At first, some scholars (e.g., Roberval) questioned Fermat's principle of reasoning in terms of the maximum number of hypothetical rounds, because it is very likely that the maximum will never be played [14, p. 30]. Indeed the sample space could have been written as $\Omega = \{A_1, B_1A_2, B_1B_2A_3, B_1B_2B_3\}$. Fermat anticipated this criticism and correctly pointed out that writing $\Omega$ in the latter form would not result in equally likely sample points. In the same September 25 letter, he writes:

> "[T]his fiction," of extending the game to a certain number of plays serves only to make the rule easy and (according to my opinion) to make all the chances equal; or better, more intelligibly to reduce all the fractions to the same denomination [30, p. 562].

Fermat's procedure is the easiest method of solution for simple examples of POP. For POP-n, the maximum number of remaining rounds is

$$\rho_n = \sum_{i=1}^{n} s_i - n + 1$$

and the number of sample points in $\Omega$ is

$$|\Omega| = n^{\rho_n}.$$

Thus, the method of enumeration quickly becomes intractable when either the number of players ($n$) or especially the maximum number of remaining rounds ($\rho_n$) increases. Moreover, it works only for players of equal skill (all $p_i$'s equal). These factors limit its usefulness.

## Method of Recursion

Pascal was aware of Fermat's method of enumeration [11] but he also recognized it could become very complicated. He therefore sought an alternative method. In his letter, dated July 29, 1654, he states:

> Your method is very sound and it is the first one that came to my mind in these researches, but because the trouble of these combinations was excessive, I found an abridgment and indeed another method that is much shorter and more neat, which I should like to tell you here in a few words; for I should like to open my heart to you henceforth if I may, so great is the pleasure I have had in our agreement. I plainly see that the truth is the same at Toulouse and at Paris.[2] [30, p. 548]

The "shorter and more neat" method Pascal is referring to in his letter is recursion. Pascal's method of recursion may be represented by a recursive tree as shown in Figure 1. Suppose we wish to solve POP-2' and determine

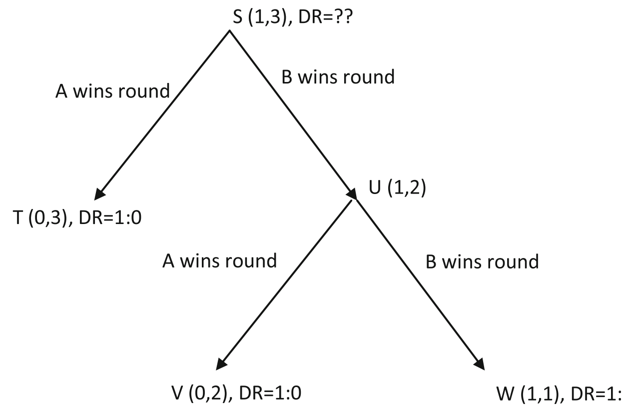


**Figure 1.** Tree Illustrating Pascal's Method of Recursion.

the DR at S(1, 3). To build the tree, we start from the root S and build the tree downward. Each left branch added corresponds to a round won by A (with probability 1/2); each right branch corresponds to a round lost by A (with probability 1/2). We continue building the tree downward until, for $i \neq 0$, we reach an $(i, 0)$ (corresponding to a DR = 0:1), or an $(0, i)$ (corresponding to a DR = 1:0), or an $(i, i)$ (corresponding to a DR = 1:1). Having done that, we now determine the DR at U(1,2): From U, A can either win the next round and win the game, or lose the next round and then be at equality with B. The probability of this is $(1/2)(1) + (1/2)(1/2) = 3/4$, i.e., the DR at U(1,2) is 3:1. We can finally now determine the DR at S(1,3): From S, A can either win the next round and win the game, or lose the game and be at U(1,2). The probability of this is $(1/2)(1) + (1/2)(3/4) = 7/8$, i.e., the DR at S(1,2) is 7:1.

Pascal's reasoning may be expressed as a difference equation as follows. Consider POP-2 such that the player who wins a total of $r$ rounds first collects the prize. Let the game stop suddenly when $A_1$ is short to win by $s_1$ rounds and $A_2$ is short to win by $s_2$ rounds, and let $A_1$'s probability of winning be $P^{(1)}_{s_1,s_2}$. Then

$$\begin{aligned} P^{(1)}_{s_1,s_2} &= \frac{1}{2}\left[P^{(1)}_{s_1-1,s_2} + P^{(1)}_{s_1,s_2-1}\right], \quad s_1, s_2 = 1, 2, \ldots, r, \\ P^{(1)}_{s_1,0} &= 0, \quad s_1 = 1, 2, \ldots, r-1, \\ P^{(1)}_{0,s_2} &= 1, \quad s_2 = 1, 2, \ldots, r-1. \end{aligned} \tag{1}$$

Although Pascal's recursive method may be applied to POP-n in general, the method becomes cumbersome for large $n$ or large $\rho_n$, and may be used only for the simplest cases.

## The Arithmetic Triangle

Pascal realized that his recursive method would quickly become unwieldy for large $s_1$ and $s_2$ in POP-2. Moreover, he was unable to use it when player $A_1$ is $s_2 - 1$ rounds short and player $A_2$ is $s_2$ rounds short. Therefore, he resorted to the Arithmetic Triangle[3] for a solution. He

[2] Pascal was residing in Paris while Fermat was in Toulouse.

[3] The Arithmetic Triangle was known well before Pascal and had also been used by Cardano in his *Opus novum* [1, 2]. It is called *Yang Hui's Triangle* in China in honor of the Chinese mathematician Yang Hui (1238–1298) who used it in 1261. Others have called it *Halayudha's Triangle* because the Indian writer Halayudha used it in the tenth century. The triangle was first called *Pascal's* triangle by Montmort in his *Essay d'Analyse sur les Jeux de Hazard* [25, p. 80], see Samueli and Boudenot [29, pp. 38–39]. For a modern treatment of the Pascal's Arithmetic Triangle, see Edwards [12] and Hald [16, pp. 45–54].

correctly identified the value of $s_1 + s_2 - 1$ with each row of the triangle, such that the corresponding entries give the number of ways $A_1$ can win 0, 1, 2, ..., rounds. Thus, for row $s_1 + s_2 - 1$, the $j^{th}$ entry counting from left is $\binom{s_1+s_2-1}{j}$, the number of ways $A_1$ can win $j$ rounds out of $s_1 + s_2 - 1$. Now, suppose player $A_1$ is short by $s_1$ rounds and player $A_2$ by $s_2$ rounds. Player $A_1$ wins if she wins any of the remaining $s_1, s_1 + 1, \ldots, s_1 + s_2 - 1$ rounds. Pascal showed that the number of ways this can happen is given by the sum

$$\binom{s_1+s_2-1}{s_1} + \binom{s_1+s_2-1}{s_1+1} + \ldots + \binom{s_1+s_2-1}{s_1+s_2-1} \equiv \binom{s_1+s_2-1}{0} + \binom{s_1+s_2-1}{1} + \ldots + \binom{s_1+s_2-1}{s_2-1},$$

which is the sum of the first $s_2$ entries in the Arithmetic Triangle for row $s_1 + s_2 - 1$. Similarly, player $A_2$ wins if she wins any of the remaining $s_2, s_2 + 1, \ldots, s_1 + s_2 - 1$ rounds. The number of ways this can happen is given by the sum

$$\binom{s_1+s_2-1}{s_2} + \binom{s_1+s_2-1}{s_2+1} + \ldots + \binom{s_1+s_2-1}{s_1+s_2-1},$$

which is the sum of the last $s_1$ entries in the Arithmetic Triangle for row $s_1 + s_2 - 1$. Pascal was thus able to provide the general DR for a fair game between $A_1$ and $A_2$ from the entries of his Arithmetic Triangle (counting from left):

$$\begin{aligned}&(\text{sum of the first } s_2 \text{ entries for row } s_1+s_2-1)\\ &\quad : (\text{sum of the last } s_1 \text{ entries for row } s_1+s_2-1)\end{aligned} \qquad (2)$$

Although Pascal solved only simple cases in his correspondence with Fermat, he was able to use mathematical induction to prove the previously mentioned general DR in his *Traité du Triangle Arithmétique* [27].[4] Applying this simple rule to POP-2, we have $s_1 = 1$, $s_2 = 3$, $s_1 + s_2 - 1 = 3$, and a DR of $(1 + 3 + 3){:}1 = 7{:}1$ between A and B, as required.

The Arithmetic Triangle may be applied to any POP-2 as long as the players are equally skilled. It may be used only for simple cases as the method becomes cumbersome when $\rho_2 = s_1 + s_2 - 1$ is large. It is not applicable to POP-n for $n > 2$.

## Binomial Distribution

Let us now generalize Pascal's idea when players $A_1$ and $A_2$ have probabilities $p_1$ and $p_2$ $(= 1 - p_1)$ of winning each round. Suppose $A_1$ and $A_2$ are $s_1$ and $s_2$ rounds, respectively, short of winning the prize, when the game suddenly stops. If the game had continued, the maximum number of possible more rounds would have been $s_1 + s_2 - 1$. Player $A_1$ wins the prize by winning any of $s_1, s_1 + 1, \ldots, s_1 + s_2 - 1$ rounds. Now $A_1$ can win $j$ rounds out of $s_1 + s_2 - 1$ rounds in $\binom{s_1+s_2-1}{j}$ ways, so $A_1$'s probability of winning the prize is

$$P^{(1)}_{s_1,s_2} = \sum_{j=s_1}^{s_1+s_2-1} \binom{s_1+s_2-1}{j} p_1^j p_2^{s_1+s_2-1-j}, \quad p_2 = 1 - p_1. \qquad (3)$$

Equation (3) first appeared in the second edition of Pierre Rémond de Montmort's (1678–1719) *Essay d'Analyse sur les Jeux de Hazard*[5] [26, pp. 244–245] as the first formula for POP-2. This solution had been communicated to Montmort by John Bernoulli (1667–1748) in a letter that is reproduced in the *Essay* [26, p. 295].

The binomial method is one of the best methods for solving any POP-2, but is difficult to extend to POP-n for $n > 2$. To see this, let us consider POP-3: Players $A_1$, $A_2$, and $A_3$ are $s_1$, $s_2$, and $s_3$ rounds, respectively, short of winning a prize, when the game suddenly stops. To calculate $P^{(1)}_{s_1,s_2,s_3}$, $A_1$ must not only win $s_1$ rounds, but also she must do so *before* $A_2$ and $A_3$ win $s_2$ and $s_3$ rounds, respectively. This suggests a *waiting time* approach to solving such problems, which is precisely what the next method does.

## Negative Binomial Distribution

Having received Fermat's method of enumeration for two players, Pascal incorrectly stated that Fermat's method was not applicable to a game with three players (POP-3). In the letter of August 24, 1654, Pascal says:

> When there are but two players, your theory which proceeds by combinations is very just. But when there are three, I believe I have a proof that it is unjust that you should proceed in any other manner than the one I have. [30, p. 554]

In his September 25, 1654, letter to Pascal, Fermat explains why his method of counting is actually correct and gives the following alternative method to obtain the DR. Fermat's reasoning is based on the waiting time for a given number of "successes." Let us generalize his idea using modern notation. Note that both Pascal and Fermat considered only fair games, but here we shall assume $A_1$ and $A_2$'s probabilities of winning one round are $p_1$ and $p_2 = 1 - p_1$, respectively. Note that $A_1$ is $s_1$ rounds short of winning the game and the maximum number of possible more rounds is $s_1 + s_2 - 1$. To obtain $P^{(1)}_{s_1,s_2}$, we can either focus on the number of additional rounds won, as in the binomial case, or we can do the following: we watch the game only until $A_1$ wins $s_1$ additional rounds (and "don't worry" about the rest of the game). Proceeding in this manner, we see that the $s_1$ rounds can be won out of $s_1$ rounds, or out of $(s_1 + 1)$ rounds,..., or out of $(s_1 + s_2 - 1)$ rounds. Now, for $A_1$ to win $s_1$ rounds out of $(s_1 + j)$ rounds $(j = 0, 1, \ldots, s_2 - 1)$, she must win $s_1 - 1$ rounds out of $s_1 - 1 + j$ rounds and then also win the $(s_1 + j)^{th}$ round. Thus $A_1$'s probability of winning can also be written as

[4] In Pascal's *Oeuvres Completes* Vol. II [28, pp. 434–436].

[5] An Essay on the Analysis of Games of Chance.

$$P^{(1)}_{s_1,s_2} = p_1^{s_1}\sum_{j=0}^{s_2-1}\binom{s_1-1+j}{s_1-1}p_2^j, \quad p_2 = 1-p_1. \qquad (4)$$

The prize should therefore be divided between $A_1$ and $A_2$ with DR $=P^{(1)}_{s_1,s_2} : (1-P^{(1)}_{s_1,s_2})$. This second reasoning is based on a *binomial waiting time* or negative binomial distribution. The latter is also sometimes called the Pascal distribution, although it is Fermat who first actually made use of it for the case $p = 1/2$. Eq. (4) first appeared in the second edition of Montmort's *Essay d'Analyse sur les Jeux de Hazard*[6] [26, p. 245] as the second formula for POP-2.

Together with the binomial distribution, the negative binomial distribution is the most frequently proposed method for solving POP-2. Moreover, the negative binomial distribution offers a natural extension of POP for more than two players [17, p. 442]. Let us take the case of POP-3. Player $A_1$ is $s_1$ rounds short of winning and the maximum number of possible more rounds is $s_1+s_2+s_3-2$. To win the game, she can win either on the $s_1^{\text{th}}$ round, or the $(s_1+1)^{\text{th}}$ round, …, or the $(s_1+s_2+s_3-2)^{\text{th}}$ round. Thus, she must win $s_1-1$ rounds out of the total $s_1-1+j$ $(j=0, 1, \ldots, s_2+s_3-2)$ rounds with players $A_2$ and $A_3$ also winning $v_2 (< s_2)$ and $j-v_2 (< s_3)$ rounds, and then finally player $A_1$ winning the $(s_1+j)$th round. This event has probability

$$\sum_{v_2=j-s_3+1}^{s_2-1}\binom{s_1-1+j}{s_1-1, v_2, j-v_2}p_1^{s_1}p_2^{v_2}p_3^{j-v_2} = \binom{s_1-1+j}{s_1-1}\sum_{v_2=j-s_3+1}^{s_2-1}\binom{j}{v_2}p_1^{s_1}p_2^{v_2}p_3^{j-v_2},$$

$$p_3 = 1-p_2-p_1.$$

Summing over all $j = 0, 1, \ldots, s_2+s_3-2$, we obtain the probability of $A_1$ winning the game as

$$P^{(1)}_{s_1,s_2,s_3} = \sum_{j=0}^{s_2+s_3-2}\binom{s_1-1+j}{s_1-1}\sum_{v_2=j-s_3+1}^{s_2-1}\binom{j}{v_2}p_1^{s_1}p_2^{v_2}p_3^{j-v_2},$$

$$p_3 = 1-p_2-p_1.$$

The probability of $A_2$ winning the game can be obtained by interchanging $p_1$, $s_1$ with $p_2$, $s_2$, respectively,

$$P^{(2)}_{s_1,s_2,s_3} = \sum_{j=0}^{s_1+s_3-2}\binom{s_2-1+j}{s_2-1}\sum_{v_2=j-s_3+1}^{s_1}\binom{j}{v_2}p_1^{v_2}p_2^{s_2+j}p_3^{j-v_2},$$

$$p_3 = 1-p_2-p_1.$$

We can use $P^{(1)}_{s_1,s_2,s_3}$ and $P^{(2)}_{s_1,s_2,s_3}$ above to solve the POP-3 problem (which we will denote by POP-3') that we alluded to at the start of this Section: Players $A_1$, $A_2$, and $A_3$ of equal skill are 1, 2, 2 rounds, respectively, short of winning a prize, when the game suddenly stops. How should the prize be divided? Using the formulas above, we have $P^{(1)}_{s_1,s_2,s_3} = 17/27$, $P^{(2)}_{s_1,s_2,s_3} = 5/27$, and also $P^{(3)}_{s_1,s_2,s_3} = 5/27$.

An alternative form of solution, also using the waiting argument, was offered by Abraham de Moivre (1667–1754) [5, Problem 8;6]. The first general formula for POP-n appears in de Moivre [7; 8, Problems 6 and 69; 9, Problem 6] as

$$P^{(1)}_{s_1,\ldots,s_k} = \sum_{0\le x_i\le s_i-1, i=2,\ldots,k} \times\binom{s_1-1+x_2+\ldots+x_n}{s_1-1, x_2, \ldots, x_n}p_1^{s_1}p_2^{x_2}\ldots p_n^{x_n},$$

$$\sum_{i=1}^{n}p_i = 1. \qquad (5)$$

The waiting-time argument offers a powerful method for solving POP-n. The solution in (5) is appealing and may easily be recommended.

## Inverse Probability

In *Mémoire sur la probabilité des causes par les événement,* Pierre-Simon Laplace (1749–1827) considered a Bayesian variation of POP [20, p. 39]. In POP-2, for example, he considered two players such that the player who wins a total of $r$ rounds first collects the prize. Let the game stop suddenly when $A_1$ is short of winning by $s_1$ rounds and $A_2$ is short of winning by $s_2$ rounds, and let $A_1$'s probability of winning be $P^{(1)}_{s_1,s_2}$. Now Laplace assumes the probability $p_1$ that $A_1$ wins a round is unknown and has an *a priori* Unif(0, 1) distribution, i.e., $f(p_1) = 1$ for $0 < p_1 < 1$. Let $F$ be the event "$A_1$ wins the game," and $G$ be the event "$A_1$ wins $r-s_1$ rounds and $A_2$ wins $r-s_2$ rounds." Then Bayes's Theorem yields

$$f(p_1|G) = \frac{f(G|p_1)f(p_1)}{\int_0^1 f(G|p_1)f(p_1)dp_1} = \frac{p_1^{r-s_1}(1-p_1)^{r-s_2}}{\int_0^1 p_1^{r-s_1}(1-p_1)^{r-s_2}dp_1} = \frac{p_1^{r-s_1}(1-p_1)^{r-s_2}}{B(r-s_1+1, r-s_2+1)},$$

where

$$B(\alpha,\beta) = \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

is the B-function, and

$$\Gamma(\alpha) = \int_0^\infty e^{-u}u^{\alpha-1}du$$

is the $\Gamma$-function. Therefore,

$$\Pr\{F|G\} = \int_0^1 \Pr\{F|G,p_1\}f(p_1|G)dp_1 = \int_0^1 P^{(1)}_{s_1,s_2}f(p_1|G)dp_1.$$

Using the last expression for $f(p_1 | G)$ and the formula for $P^{(1)}_{s_1,s_2}$ from Eq. (3), we have

[6] In the first edition of 1708, Montmort discussed the problem of points, but only for a fair game [25, pp. 165–178].

$$\Pr\{F|G\} = \frac{\int_0^1 \sum_{j=s_1}^{s_1+s_2-1} \binom{s_1+s_2-1}{j} p_1^j (1-p_1)^{s_1+s_2-1-j} p_1^{r-s_1}(1-p_1)^{r-s_2} dp_1}{B(r-s_1+1, r-s_2+1)}$$
$$= \sum_{j=s_1}^{s_1+s_2-1} \left[ (s_1+s_2-1j) \frac{B(r-s_1+j+1, r+s_1-j)}{B(r-s_1+1, r-s_2+1)} \right]$$
$$= \frac{1}{\binom{2r}{s_1+s_2-1}} \sum_{j=0}^{s_2-1} \binom{r+j}{r-s_1} \binom{r-j-1}{r-s_2}. \tag{6}$$

If we solve POP-2' using the above, we obtain a probability of $A_1$ winning of

$$\Pr\{F|G\} = \frac{1}{\binom{12}{3}} \sum_{j=0}^{2} \binom{6+j}{5} \binom{5-j}{3} = \frac{10}{11},$$

(and a probability of 1/11 for $A_2$) [32]. On the other hand, when we assume the two players are equally skilled ($p_1 = p_2 = 1/2$), the probability that $A_1$ wins is 7/8, as we have shown several times before.

Laplace's method of inverse probability can hardly be recommended unless there is a strong justification for using a Unif(0, 1) prior for $p_1$. In most cases, such a prior is based on the principle of indifference, and carries all the criticisms attributed to the principle. The strongest of these pertains to the inconsistencies that arise if $p_1$ was to be transformed to a one-to-one function $f(p_1)$ and a Unif(0, 1) prior assigned to $f(p_1)$ [see, e.g., 14, p. 135].

## Difference Equations

In *Recherches sur les suites récurrentes*, Joseph-Louis Lagrange (1736–1813) presented a solution to POP through difference equations [19]. Lagrange first considers POP-2 and proceeds to solve for $P_{s_1,s_2}^{(2)}$ by using an equation similar to Eq. (1) (except that now the players are not assumed to have equal skills):

$$P_{s_1,s_2}^{(2)} = p_1 P_{s_1-1,s_2}^{(2)} + p_2 P_{s_1,s_2-1}^{(2)}, \quad s_1, s_2 = 1, 2, \ldots, r; p_1 + p_2 = 1,$$
$$P_{s_1,0}^{(2)} = 1, \qquad s_1 = 1, 2, \ldots, r-1,$$
$$P_{0,s_2}^{(2)} = 0, \qquad s_2 = 1, 2, \ldots, r-1. \tag{7}$$

Lagrange assumes a general solution of the form

$$P_{s_1,s_2}^{(2)} = \gamma \alpha^{s_1} \beta^{s_2}, \tag{8}$$

where $\alpha$, $\beta$, $\gamma$ are arbitrary constants. Substituting the above in Eq. (7), we obtain

$$\beta = \frac{p_2}{1 - p_1/\alpha},$$

so that Eq. (8) becomes

$$P_{s_1,s_2}^{(2)} = \gamma \alpha^{s_1} p_2^{s_2} \left\{ 1 + s_2 \frac{p_1}{\alpha} + \frac{s_2(s_2+1)}{2!} \frac{p_1^2}{\alpha^2} + \ldots + \right\}$$
$$= p_2^{s_2} \left\{ \gamma \alpha^{s_1} + s_2 p_1 \gamma \alpha^{s_1-1} + \frac{s_2(s_2+1)}{2!} p_1^2 \gamma \alpha^{s_1-2} + \ldots \right\}$$
$$= p_2^{s_2} \left\{ P_{s_1,0}^{(2)} + s_2 p_1 P_{s_1-1,0}^{(2)} + \frac{s_2(s_2+1)}{2!} p_1^2 P_{s_1-2,0}^{(2)} + \ldots \right\}.$$

Now, for $i = 1, 2, 3, \ldots$, we have $P_{i,0}^{(2)} \equiv 1$ and $P_{0,i-1}^{(2)} \equiv 0$, so that the above becomes

$$P_{s_1,s_2}^{(2)} = p_2^{s_2} \left\{ 1 + s_2 p_1 + \frac{s_2(s_2+1)}{2!} p_1^2 + \ldots + \frac{s_2(s_2+1)(s_2+2)\ldots(s_2+s_1-2)}{(s_1-1)!} p_1^{s_1-1} \right\}.$$

This coincides in form to the solution provided by the negative binomial argument in Eq. (4) (with $p_2$, $s_2$ replaced by $p_1$, $s_1$, respectively).

Using difference equations, Lagrange also considers the POP-3 and POP-n case. His general solution coincides in form with Eq. (5). However, the method of difference equations was historically less elegant than, and soon superseded by, the next method.

## Probability Generating Functions

Although Laplace had previously also used difference equations to solve POP, in the *Théorie Analytique des Probabilités* [21, p. 207] he solved POP-n through the method of probability generating functions. Let $P_{s_1,s_2,\ldots,s_n}^{(1)}$ be player $A_1$'s probability of winning the game. Generalizing Eq. (1) to $n$ players, we obtain

$$P_{s_1,s_2,\ldots,s_n}^{(1)} = p_1 P_{s_1-1,s_2,\ldots,s_n}^{(1)} + p_2 P_{s_1,s_2-1,\ldots,s_n}^{(1)} + \ldots + p_n P_{s_1,s_2,\ldots,s_n-1}^{(1)}.$$

where

$$P_{0,s_2,\ldots,s_n}^{(1)} = 1, \quad s_2, \ldots, s_n > 0,$$
$$P_{s_1,0,\ldots,s_n}^{(1)} = 0, \quad s_1, s_3, \ldots, s_n > 0,$$
$$\ldots$$
$$P_{s_1,s_2,\ldots,0}^{(1)} = 0, \quad s_1, \ldots, s_{n-1} > 0,$$
$$\sum_{i=1}^{n} p_i = 1.$$

We define the multivariate probability generating function $G(z_1, z_2, \ldots, z_n) \equiv G$ by

$$G = \sum_{s_1=1}^{\infty} \sum_{s_2=1}^{\infty} \ldots \sum_{s_n=1}^{\infty} P_{s_1,s_2,\ldots,s_n}^{(1)} z_1^{s_1} z_2^{s_2} \ldots z_n^{s_n}$$
$$= \sum_{s_1=1}^{\infty} \sum_{s_2=1}^{\infty} \ldots \sum_{s_n=1}^{\infty} \Big( p_1 P_{s_1-1,s_2,\ldots,s_n}^{(1)} + p_2 P_{s_1,s_2-1,\ldots,s_n}^{(1)} + \ldots + p_n P_{s_1,s_2,\ldots,s_n-1}^{(1)} \Big) z_1^{s_1} z_2^{s_2} \ldots z_n^{s_n}$$
$$= p_1 z_1 G + p_1 z_1 \sum_{s_2=1}^{\infty} \ldots \sum_{s_n=1}^{\infty} z_2^{s_2} \ldots z_n^{s_n} + (p_2 z_2 + \ldots + p_n z_n) G. \tag{9}$$

Therefore,

$$G=\frac{p_1z_1\sum\limits_{s_2=1}^{\infty}\ldots\sum\limits_{s_n=1}^{\infty}z_2^{s_2}\ldots z_n^{s_n}}{1-p_1z_1-p_2z_2-\ldots-p_nz_n}$$
$$=\frac{p_1z_1z_2\ldots z_n}{(1-z_2)(1-z_3)\ldots(1-z_n)(1-p_1z_1-p_2z_2-\ldots-p_nz_n)}. \tag{10}$$

(Laplace's expression for $G$ is slightly different because he starts the summations for $G$ in Eq. (9) at zero). Thus, the probability that player $A_1$ wins the game is

$$P^{(1)}_{s_1,s_2,\ldots,s_n}=\text{coef. of } z_1^{s_1}z_2^{s_2}\ldots z_n^{s_n}\text{ in } \frac{p_1z_1z_2\ldots z_n}{(1-z_2)(1-z_3)\ldots(1-z_n)(1-p_1z_1-p_2z_2-\ldots-p_nz_n)}.$$

In general, the probability that player $A_j$ wins the game is

$$P^{(j)}_{s_1,s_2,\ldots,s_n}=\text{coef. of } z_1^{s_1}z_2^{s_2}\ldots z_n^{s_n}\text{ in } \frac{p_j\prod\limits_{i=1}^{n}z_i}{\left(1-\sum\limits_{i=1}^{n}p_iz_i\right)\prod\limits_{\substack{i=1\\ i\neq j}}^{n}(1-z_i)}. \tag{11}$$

In POP-2' we had, for two players, $p_1=p_2=1/2$ and $s_1=1$, $s_2=3$. Therefore the probability that A wins is

$$\text{coef. of } z_1z_2^3 \text{ in } \frac{z_1z_2/2}{(1-z_2)(1-z_1/2-z_2/2)}=\frac{7}{8},$$

giving a division ratio of 7:1 for A and B, as we obtained before.

In POP-3', we had, for three players, $p_1=p_2=p_3=1/3$ and $s_1=1,\ s_2=2,$ and $s_3=2$. Therefore, the probability that $A_1$ wins is

$$\text{coeff. of } z_1z_2^2z_3^2 \text{ in } \frac{z_1z_2z_3/3}{(1-z_2)(1-z_3)(1-z_1/3-z_2/3-z_3/3)}=\frac{17}{27}.$$

The probability that $A_2$ wins is

$$\text{coeff. of } z_1z_2^2z_3^2 \text{ in } \frac{z_1z_2z_3/3}{(1-z_1)(1-z_3)(1-z_1/3-z_2/3-z_3/3)}=\frac{5}{27},$$

resulting in a division ratio of 17/27:5/27:5/27 = 17:5:5, again the same answer we obtained previously.

The method of probability generating functions is one of the elegant methods for solving POP. It may be used for any number of players with arbitrary skills and, with today's computational power, it is fairly easily applicable.

## Incomplete B-Function

In his *Cours de Calcul des Probabilités* [24, p. 65], Antoine Meyer (1801–1857) proposed an alternative solution to POP. Let us consider POP-2, for example. Meyer's lengthy demonstration essentially boils down to showing the well-known relationship between the distribution function of the binomial and the incomplete B-function [e.g., see 15, p. 674]:

$$\sum_{i=\alpha}^{\alpha+\beta-1}\binom{\alpha+\beta-1}{i}p^i(1-p)^{\alpha+\beta-1-i}=\frac{B_p(\alpha,\beta)}{B(\alpha,\beta)}.$$

In the above, $B_p$ is the incomplete B-function:

$$B_p(\alpha,\beta)=\int_0^p u^{\alpha-1}(1-u)^{\beta-1}du.$$

Using Eq. (3), we have

$$P^{(1)}_{s_1,s_2}=\frac{B_{p_1}(s_1,s_2)}{B(s_1,s_2)}=\frac{\Gamma(s_1+s_2)}{\Gamma(s_1)\Gamma(s_2)}\int_0^{p_1}u^{s_1-1}(1-u)^{s_2-1}du. \tag{12}$$

For POP-2, the above gives $P^{(1)}_{s_1,s_2}=7/8$ as we obtained previously. Now, by making the substitution $u=1/(1+t)$, the formula in (12) becomes

$$P^{(1)}_{s_1,s_2}=\frac{\Gamma(s_1+s_2)}{\Gamma(s_1)\Gamma(s_2)}\int_{p_2/p_1}^{\infty}\frac{t^{s_2-1}}{(1+t)^{s_1+s_2}}dt,\quad p_1+p_2=1.$$

By using the above form for $P^{(1)}_{s_1,s_2}$, Meyer is able to extend the formula for POP-n:

$$P^{(1)}_{s_1,\ldots,s_n}=\frac{\Gamma(s_1+\ldots+s_n)}{\Gamma(s_1)\ldots\Gamma(s_n)}\times\int_{p_2/p_1}^{\infty}\ldots\int_{p_n/p_1}^{\infty}\frac{t_2^{s_2-1}\ldots t_n^{s_n-1}}{(1+t_2+\ldots+t_n)^{s_1+\ldots+s_n}}dt_2\ldots dt_n,\quad \sum_{i=1}^{n}p_i=1. \tag{13}$$

Let us solve POP-3' using the above method. We have, for three players, $p_1=p_2=p_3=1/3$ and $s_1=1,\ s_2=2,$ and $s_3=2$. Therefore,

$$P^{(1)}_{s_1,s_2,s_3}=\frac{4!}{0!1!1!}\int_1^{\infty}\int_1^{\infty}\frac{t_2t_3}{(1+t_2+t_3)^5}dt_2dt_3=\frac{17}{27},$$
$$P^{(2)}_{s_1,s_2,s_3}=\frac{4!}{0!1!1!}\int_1^{\infty}\int_1^{\infty}\frac{t_3}{(1+t_2+t_3)^5}dt_2dt_3=\frac{5}{27}.$$

For $P^{(2)}_{s_1,s_2,s_3}$ in the above, we have used $P^{(1)}_{s_1,s_2,s_3}$ with $p_1$, $s_1$ and $p_2$, $s_2$ interchanged.

Meyer's method is general, computationally satisfactory, and has much to recommend it. It may be preferred to the method of generating functions because it does not require extracting coefficients.

## Variation on Negative Binomial Distribution: I

In *Sur le Problème des Partis* [23], Paul Mansion (1844–1919) gave an alternative argument for POP, which results in a relatively simple form of the solution. Mansion considers POP-2 and uses the solution based on the negative binomial distribution in Eq. (4):

$$P^{(1)}_{s_1,s_2} = p_1^{s_1}\sum_{j=0}^{s_2-1}\binom{s_1-1+j}{s_1-1}(1-p_1)^j.$$

The latter can be written as

$$\begin{aligned}P^{(1)}_{s_1,s_2} &= p_1^{s_1}\sum_{j=0}^{s_2-1}\binom{s_1-1+j}{s_1-1}(1-p_1)^j\\ &= p_1^{s_1}\sum_{j=0}^{s_2-1}\sum_{i=0}^{j}\binom{s_1-1+j}{s_1-1}\binom{j}{i}(-1)^i p_1^i.\end{aligned}$$

The coefficient of $(-1)^i p_1^{i+s_1}$ is

$$\begin{aligned}\sum_{j=i}^{s_2-1}\binom{s_1-1+j}{s_1-1}\binom{j}{i} &= \sum_{j=i}^{s_2-1}\binom{s_1-1+j}{j-i}\binom{s_1-1+i}{i}\\ &= \binom{s_1-1+i}{i}\sum_{j=i}^{s_2-1}\binom{s_1-1+j}{j-i}\\ &= \binom{s_1-1+i}{i}\binom{s_1+s_2-1}{s_2-1-i}\\ &= \frac{1}{B(s_1,s_2)}\cdot\frac{1}{s_1+i}\binom{s_2-1}{i}.\end{aligned}$$

Hence,

$$P^{(1)}_{s_1,s_2} = \frac{\Gamma(s_1+s_2)}{\Gamma(s_1)\Gamma(s_2)}\sum_{i=0}^{s_2-1}(-1)^i\frac{1}{s_1+i}\binom{s_2-1}{i}p_1^{i+s_1}. \quad (14)$$

Although Mansion's method is relatively simple for the POP-2 case, it is less elegant for $n > 2$ and cannot be recommended for such cases.

## Variation On Negative Binomial Distribution: II

Eugène Charles Catalan (1814–1894) provided an alternative solution to POP in 1878 with interesting interpretations [3]. Like Mansion, Catalan also starts with the solution based on the negative binomial distribution for POP-2 (see Eq. (4)):

$$P^{(1)}_{s_1,s_2} = p_1^{s_1}\sum_{j=0}^{s_2-1}\binom{s_1-1+j}{s_1-1}(1-p_1)^j,$$

$$P^{(2)}_{s_1,s_2} = p_2^{s_2}\sum_{j=0}^{s_1-1}\binom{s_2-1+j}{s_2-1}(1-p_2)^j, \quad p_2 = 1-p_1.$$

Now,

$$p_1^{-s_1} = (1-p_2)^{-s_1} = \sum_{j=0}^{\infty}\binom{s_1-1+j}{s_1-1}p_2^j,$$

so that

$$1 = p_1^{s_1}\sum_{j=0}^{\infty}\binom{s_1-1+j}{s_1-1}p_2^j. \quad (15)$$

Because $P^{(1)}_{s_1,s_2} + P^{(2)}_{s_1,s_2} = 1$, we have

$$P^{(1)}_{s_1,s_2} + P^{(2)}_{s_1,s_2} = p_1^{s_1}\sum_{j=0}^{\infty}\binom{s_1-1+j}{s_1-1}p_2^j.$$

Subtracting Eq. (4) from the above equation, $P^{(2)}_{s_1,s_2}$ can be expressed as an infinite series:

$$P^{(2)}_{s_1,s_2} = p_1^{s_1}\sum_{j=s_2}^{\infty}\binom{s_1-1+j}{s_1-1}p_2^j.$$

Similarly,

$$P^{(1)}_{s_1,s_2} = p_2^{s_2}\sum_{j=s_1}^{\infty}\binom{s_2-1+j}{s_2-1}p_1^j. \quad (16)$$

Catalan then makes two interesting remarks (one obvious, the other not so obvious):

- The right side of Eq. (15) is the sum of the probabilities that player $A_1$ wins $s_1 - 1$ rounds out of $s_1 - 1 + j$ ($j = 0, 1, 2, \ldots$), and then wins the $(s_1 + j)^{th}$ round. It is thus certain that $A_1$ will eventually win the $s_1$ rounds that she is short of, if the game were to continue indefinitely.
- Eq. (16) gives the probability that player $A_2$ wins $s_2$ rounds out of a total of at least $s_1 + s_2$ rounds. This is the same as the probability of $A_1$ winning $s_1$ rounds out of a total of $s_1 + s_2 - 1$.

Catalan does not discuss POP-n for $n > 2$, because the aim of his article was to give an alternative interpretation of $P^{(1)}_{s_1,s_2}$ in terms of infinite series. The formulas for $P^{(1)}_{s_1,s_2}$ and $P^{(2)}_{s_1,s_2}$ cannot be recommended for computational purposes because convergence is relatively slow (e.g., for POP-2', we need to sum the first 16 terms to obtain $P^{(1)}_{s_1,s_2} = .875$ to 3 d.p.)

## Variation on Binomial Distribution: I

In addition to the two traditional approaches for solving POP-2 (in earlier sections), George Chrystal (1851–1911) provided an interesting third argument for $P^{(1)}_{s_1,s_2}$ in the second part of his famous *Algebra: An Elementary Text* [4, pp. 556–558].[7] Unlike the two previous authors, Chrystal's derivation is actually based on the binomial distribution. Player $A_1$ wins at exactly $j$ rounds out of $s_1 + s_2 - 1$ rounds with probability

$$\begin{aligned}\binom{s_1+s_2-1}{j}p_1^j(1-p_1)^{s_1+s_2-1-j} &= \binom{s_1+s_2-1}{j}p_1^j\\ &\times\left[1-\binom{s_1+s_2-1-j}{1}p_1+\ldots+(-1)^i\right.\\ &\times\left.\binom{s_1+s_2-1-j}{i}p_1^i-\ldots+(-1)^{s_1+s_2-1-j}p_1^{s_1+s_2-1-j}\right]\\ &= \binom{s_1+s_2-1}{j}p_1^j-\binom{s_1+s_2-1}{j}\binom{s_1+s_2-1-j}{1}p_1^{j+1}\end{aligned}$$

[7]Chrystal's text is especially famous for statisticians because the author chose to omit inverse probability from it, and that was used by none other than Fisher to buttress his arguments against this particular use of probability [13, p. 29].

$$+\ldots+(-1)^i\binom{s_1+s_2-1}{j}\binom{s_1+s_2-1-j}{i}p_1^{j+i}$$
$$+\ldots+(-1)^{s_1+s_2-1-j}\binom{s_1+s_2-1}{j}p_1^{s_1+s_2-1}$$
$$=\binom{s_1+s_2-1}{j}p_1^j-\binom{j+1}{1}\binom{s_1+s_2-1}{j+1}p_1^{j+i}$$
$$+\ldots+(-1)^i\binom{j+i}{i}\binom{s_1+s_2-1}{j+i}p_1^{j+i}$$
$$+\ldots+(-1)^{s_1+s_2-1-j}\binom{s_1+s_2-1}{s_1+s_2-1-j}p_1^{s_1+s_2-1}.$$

Therefore, the probability that player $A_1$ wins at least $s_1$ rounds out of $s_1+s_2-1$ rounds can be obtained by summing the above from $j=s_1$ to $j=s_1+s_2-1$. In the latter summation, the coefficient of $\binom{s_1+s_2-1}{i+s_1}p_1^{i+s_1}$ is

$$(-1)^i\sum_{j=s_1}^{s_1+i}\binom{s_1+i}{j}(-1)^{j-s_1}.$$

Except for the $(-1)^i$, the above is actually the coefficient of $x^i$ in $(1+x)^{-1}(1+x)^{i+s_1}=(1+x)^{i+s_1-1}$, and is thus equal to $\binom{i+s_1-1}{i}$. Therefore, the coefficient of $\binom{s_1+s_2-1}{i+s_1}p_1^{i+s_1}$ is $(-1)^i\binom{i+s_1-1}{i}$. Hence, Chrystal is able finally to give the following expression for $P^{(1)}_{s_1,s_2}$ :

$$P^{(1)}_{s_1,s_2}=\sum_{i=0}^{s_2-1}(-1)^i\binom{s_1+s_2-1}{i+s_1}\binom{i+s_1-1}{i}p_1^{i+s_1}. \quad (17)$$

The interesting form that $P^{(1)}_{s_1,s_2}$ takes is best seen when expanded:

$$P^{(1)}_{s_1,s_2}=\binom{s_1+s_2-1}{s_1}p_1^{s_1}-\binom{s_1+s_2-1}{s_1+1}\binom{s_1}{1}p_1^{s_1+1}$$
$$+\binom{s_1+s_2-1}{s_1+2}\binom{s_1+1}{2}p_1^{s_1+2}$$
$$+\ldots.+(-1)^{s_2-1}\binom{s_1+s_2-1}{s_2-1}p_1^{s_1+s_2-1}.$$

Chrystal does not consider POP-n for $n>2$. Although Eq. (17) provides interesting insights, Chrystal's approach cannot easily be extended for more players and is recommended only for POP-2.

## Markov Chains

In his recent book *Chapters in Probability*, Craig Smorynski has offered yet an alternative method of solution for POP through Markov chains [31, p. 372]. For POP-n, suppose the players are $s_{1,t}, s_{2,t}, \ldots, s_{n,t}$ rounds short of winning at some time $t$ during the game. Because the current state depends only on what happened on the previous round (i.e., at time $t-1$), it is natural to model $S_{n,t}=(s_{1,t}, s_{2,t}, \ldots, s_{n,t})$ as a Markov chain.

Let us now use this idea to solve POP-2'. The players A and B are $s_1=1$ and $s_2=3$ short of winning the game. By reasoning on what happens after (1, 3), we see that the Markov chain $\{S_{3,t}\}$ has states {(0,1), (0, 2), (0, 3), (1,0), (1,1), (1,2), (1,3)}, and transition probability matrix $\mathbf{P}$:

| | (0, 1) | (0, 2) | (0, 3) | (1, 0) | (1, 1) | (1, 2) | (1, 3) |
|---|---|---|---|---|---|---|---|
| (0, 1) | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| (0, 2) | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| (0, 3) | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| (1, 0) | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| (1, 1) | 1/2 | 0 | 0 | 1/2 | 0 | 0 | 0 |
| (1, 2) | 0 | 1/2 | 0 | 0 | 1/2 | 0 | 0 |
| (1, 3) | 0 | 0 | 1/2 | 0 | 0 | 1/2 | 0 |

The maximum number of additional rounds is 3. By calculating $\mathbf{aP}^3$, where $\mathbf{a}=[0,0,0,0,0,0,1]$ (corresponding to the starting state (1, 3)), we obtain the probabilities of being in the various states after 3 additional rounds as

$$\mathbf{aP}^3=\left[\frac{1}{8},\frac{1}{4},\frac{1}{2},\frac{1}{8},0,0,0\right].$$

Therefore, $P^{(1)}_{s_1,s_2}=1/8+1/4+1/2=7/8$ and $P^{(2)}_{s_1,s_2}=1/8$, as required.

The Markov method is quite elegant and matrix computations are not a problem currently. It works for any POP-n although the transition matrix does become extremely large when the number of players or additional rounds increases.

## Conclusion

POP is truly one of the most beautiful problems of probability and actually founded the discipline. As we have shown, it is also one of great diversity in terms of the number of ways it has lent itself to a resolution by some of the greatest mathematicians of the past centuries. We suspect that there are even more lines of attack than the thirteen we have presented here

**ACKNOWLEDGMENTS**

I express my thanks to Craig Smorynski and an anonymous reviewer for their helpful suggestions in improving the manuscript.

Department of Biostatistics
Columbia University
Room 620
New York, NY 10032
USA
e-mail: pg2113@columbia.edu

**REFERENCES**

[1] C. B. Boyer, Cardan and the Pascal Triangle. *Amer. Math. Monthly* 57 (1950) 387–390.

[2] G. Cardano, Opus novum de proportionibus numerorum, motuum, ponerum, sonorum, aliarumque rerum mensurandum, Basel, 1570.

[3] E. Catalan, Sur le problème des partis. *Nouv. Corr. Math. IV* (1878) 8–11.
[4] G. Chrystal, Algebra: An Elementary Text-Book for the Higher Classes of Secondary Schools and for Colleges (Part II), Adam and Charles Black, Edinburgh, 1889.
[5] A. de Moivre, De Mensura Sortis, seu, de Probabilitate Eventuum in Ludis a Casu Fortuito Pendentibus. *Phil. Trans.* 27 (1711) 213–264.
[6] A. de Moivre, The Doctrine of Chances, or a Method of Calculating the Probability of Events in Play, Millar, London, 1718.
[7] A. de Moivre, Miscellanea analytica de seriebus et quadraturis, Touson & Watts, London, 1730.
[8] A. de Moivre, The Doctrine of Chances, or a Method of Calculating the Probability of Events in Play, Millar, London, 1738.
[9] A. de Moivre, The Doctrine of Chances, or a Method of Calculating the Probability of Events in Play, Millar, London, 1756.
[10] K. Devlin, The Unfinished Game: Pascal, Fermat, and the Seventeenth-Century Letter That Made the World Modern, Basic Books, New York, 2008.
[11] A. W. F. Edwards, Pascal and the problem of points. *Int. Stat. Rev.* 50 (1982) 259–266.
[12] A. W. F. Edwards, Pascal's Arithmetical Triangle: The Story of a Mathematical Idea, John Hopkins University Press (Originally published by Charles Griffin & Company Limited, London, 1987), 2002.
[13] R. A. Fisher, Statistical Methods and Scientific Inference, Hafner Publising Company, New York, 1956.
[14] P. Gorroochurn, Classic Problems of Probability, Wiley, NJ, 2012.
[15] A. Hald, Statistical Theory with Engineering Applications, Wiley, NY-London, 1952.
[16] A. Hald, A History of Probability and Statistics and Their Applications Before 1750, Wiley, New Jersey, 1990.
[17] K. Jordan, Chapters in the Classical Calculus of Probability, Akadémiaio Kiadó, Budapest, Hungary, 1972.
[18] M. G. Kendall, The beginnings of a probability calculus. Biometrika 43 (1956) 1–14.
[19] J.-L. Lagrange, Recherche sur les suites récurrentes dont les termes varient de plusieurs manières différentes, ou sur l'intégration des équations linéaires aux différences finies et partielles; et sur l'usage de ces équations dans la théorie des hasards. *Mém. Acad. Berlin* 6, 1775 (1777) 183-272; Oeuvres IV, pp. 151–251.
[20] P.-S. Laplace, Mémoire sur la probabilité des causes par les événements. *Mémoire de l'Académie Royale des Sciences de Paris (savants étrangers)* 6 (1774) 621-656; OC 8, 27–65.
[21] P.-S. Laplace, Théorie Analytique des Probabilités, Mme Ve Courcier, Paris, 1812.
[22] G. W. Leibniz, New Essays Concerning Human Understanding, The Macmillan Company, New York (original work written in 1704 and published in 1765), 1896.
[23] P. Mansion, Sur le problème des partis. *Mém. de. Belg.* XXI. F Hayez (1870).
[24] A. Meyer, Cours de Calcul des Probabilités, F. Hayez, Bruxelles, 1874.
[25] P. R. d. Montmort, Essay d'Analyse sur les Jeux de Hazard, Quillau, Paris, 1708.
[26] P. R. d. Montmort, Essay d'Analyse sur les Jeux de Hazard, Quillau, Paris, 1713.
[27] B. Pascal, Traité du Triangle Arithmétique, avec Quelques Autres Petits Traités sur la Même Matière, Desprez, Paris (English translation of first part in Smith (1929), pp. 67–79), 1665.
[28] B. Pascal, Oeuvres Complètes de Blaise Pascal (Tome Second), Librairie de L. Hachette et Cie, Paris, 1858.
[29] J. J. Samueli, J. C. Boudenot, Une Histoire des Probabilités des Origines à 1900, Ellipses, Paris, 2009.
[30] D. E. Smith, A Source Book in Mathematics, McGraw-Hill Book Company, Inc., New York, 1929.
[31] C. Smorynski, Chapters in Probability, College Publications, London, 2012.
[32] L. Takács, The problem of points. *Math. Scient.* 19 (1994) 119–134.
[33] P. Tulloch, Pascal, William Blackwood and Sons, London, 1878.