The ratio of the mean frequency to the number of trials is therefore the probability itself. When p is small the mean error differs little from the square root $\sqrt{Np}$ of the mean frequency; and if p is nearly $= 1$, the mean error of the opposite event is nearly equal to $\sqrt{Nq}$. When the probability, p, is nearly equal to $\frac{1}{2}$, the mean error will be about $\frac{1}{2}\sqrt{N}$.

The law of error is not strictly typical, although the rational function of the r^{th} degree in $\lambda_r(m)$ vanishes for r different values of p between 0 and 1, the limits included, so that the deviation from the typical form must, on the whole, be small. If, however, we consider the relative magnitude of the higher half-invariants as compared with the powers of the mean error

$$\left.\begin{aligned} \lambda_3(m)\cdot(\lambda_2(m))^{-\frac{3}{2}} &= \frac{q-p}{\sqrt{Npq}} \\ \text{and} \quad \lambda_4(m)\cdot(\lambda_2(m))^{-2} &= \frac{q^2-4pq+p^2}{Npq} \end{aligned}\right\} \qquad (125)$$

the occurence of Npq in the denominators of the abridged fractions shows, not only that great numbers of repetitions, here as always, cause an approximation to the typical form, but also that, in contrast to this, the law of error in the cases of certainty and impossibility, when $q = 0$ and $p = 0$, becomes skew and deviates from the typical in an infinitely high degree, while at the same time the square of the mean errors becomes $= 0$. This remarkable property is still traceable in the cases in which the probability is either very small or very nearly equal to 1. In a hundred trials with the probability $= 99\frac{1}{2}$ per ct. the mean error will be about $= \sqrt{\frac{1}{2}}$. Errors beyond the mean frequency $99\frac{1}{2}$ cannot exceed $\frac{1}{2}$, and are therefore less than the mean error. The great diminishing errors must therefore be more frequent than in typical cases, and frequencies of 97 or 96 will not be rare in the case under consideration, though hey must be fully counter-balanced by numerous cases of 100 per ct. The law of error is consequently skew in a perceptible degree. In applications of adjustment to problems of probability, it is, from this reason, frequently necessary to reject extreme probabilities.

XV. THE FORMAL THEORY OF PROBABILITY.

§ 67. The formal theory of probability teaches us how to determine probabilities that depend upon other probabilities, which are supposed to be given. Of course, there are no mathematical rules specially applicable to computations that deal with probabilities, and there are many computations with probabilities which do not fall under the theory of probability, for instance, adjustments of probabilities. But in view of the direct application of probabilities, not only to games, insurances, and statistics, but to all conditions of life, it will be understood that special importance attaches to the marks which show that a computation will lead us to a probability as its result, as this implies in part or in the whole a determination of a law of errors. The formal theory of probabilities rests on two theorems, one concerning the addition of probabilities, the other concerning their multiplication.

I. The theorem concerning *the addition of probabilities* can, as all probabilities are positive numbers, be deduced from the usual definition of addition as a putting together: if a sum of probabilities is to be a probability itself, we must be allowed to look upon each of the probabilities that we are to add together as corresponding to its particular events. These events must mutually exclude one another, but must at the same time have a quality in common, to which, after the addition, our whole attention must be given. If the sum is to be the correct probability of events with this quality, the same quality must be found in no other event of the trial. An "either—or" is, therefore, the simple grammatical mark of the addition of probabilities. The event E_s, whose probability is $p_1 + p_2$, must occur, if *either* the result E_1, whose probability is p_1, *or* the quite different event E_2, whose probability is p_2, occurs, *and not* in any other case. If we require no other resemblance between the events whose probabilities are added together, than that they belong to the same trial, their sum must be the probability 1, certainty, because then all events of the trial are favourable. If p be the probability for a certain event, q the probability *against* the same, then we have $p + q = 1$, $q = 1 - p$. If n events of the same trial be equally probable, the probability of each being $= p$, then the aggregate probability of these events is $= np$.

II. The theorem concerning *the multiplication of probabilities* can, as all probabilities are proper fractions, be deduced from the definition of the multiplication of fractions, according to which the product is the same proportional of the multiplicand as the multiplier is of unity. Only as probabilities presuppose infinite numbers of trials, we shall commence by proving the corresponding proposition for relative frequencies.

If, in $p = p_1 p_2$, p_1 is a relative frequency, it must relate to a trial T_1 which, repeated N times, has given favourable events in Np_1 cases; and if p_2, being also a relative frequency, takes the place of multiplier, then the corresponding trial T_2, if repeated Np_1 times, must have given $(Np_1)p_2$ favourable events. Now in the multiplication $p = p_1 p_2$, p must be the relative frequency of the compound trials which out of the total number of N repetitions have given $Np_1 p_2$ favourable events. The trials T_1 and T_2 must *both* have succeeded as conditional for the final event. As the number N can be taken as large as we please, the same proposition must hold good for probabilities.

The probability $p = p_1 p_2$, as the product of the probabilities p_1 and p_2, relates to the event of a compound trial, which is favourable only if both conditional trials, T_1 and T'_2, have given favourable events: first the trial T_1 must have had the event whose probability is p_1, and then the other trial T_2 must have succeeded in the event, whose probability, *on condition of success in* T_1, is p_2. However indifferent the order of the factors may be in the numerical computation it is nevertheless, if a probability is correctly to be found as the product of the probabilities of conditional events, necessary to imagine the conditional trials arranged in a definite order. To prove this very important proposition we shall suppose that both conditional trials are carried out in every case of the compound trial. Let both T_1 and T_2 have succeeded in a cases, while only T_1 has succeeded in b cases, only T_2 in c cases, and neither in d cases. Considering each of the two trials without any regard to the other, we therefore get $\frac{a+b}{a+b+c+d} = P_1$ and $\frac{a+c}{a+b+c+d} = P_2$ as the frequencies or probabilities of their favourable events. But in the multiplication for computation of the compound probability, P_1 and P_2 are applicable only as multiplicands; the correct result $p = \frac{a}{a+b+c+d}$ is found by $p = P_1 \cdot \frac{a}{a+b}$ or by $p = P_2 \cdot \frac{a}{a+c}$, according to the order in which the trials are executed, but *not* as $p = P_1 P_2$, unless $a : b = c : d$. But this proportion expresses that the frequency or probability of the trial T_2 is not affected by the event of the trial T_1. This proportionality is the mark of freedom, if we consider the multiplication of probabilities as the determination of the law of errors for a function of two observed values whose laws of errors are given.

Since impossibility is indicated by probability $= 0$, we see that the compound trial is impossible, if there is any of the conditional trials that cannot possibly succeed, i. e. if $p_1 = 0$ or $p_2 = 0$ in $p = p_1 p_2$. The condition of certainty (probability $= 1$) in a compound trial is certainty for the favourable events of all conditional trials; for as p_1 and p_2 as probabilities must be proper fractions, $p_1 p_2 = p = 1$ will be possible only when both $p_1 = 1$ and $p_2 = 1$.

Example 1. When the favourable events of all the conditional trials, n in number, have the same probability p, the compound event, which depends on the success of all these, has the probability p^n. If by every single drawing there is the probability of $\frac{1}{2}$ for "red" and $\frac{1}{2}$ for "black", the probability of 10 drawings *all* giving red will be $\frac{1}{1024}$.

Example 2. Suppose a pack of 52 cards to be so well shuffled that the probabilities of red and black may constantly be proportional to the remainder in the stock, then the probability of the 10 uppermost cards being red will be

$$= \frac{26}{52}\cdot\frac{25}{51}\cdot\frac{24}{50}\cdot\frac{23}{49}\cdot\frac{22}{48}\cdot\frac{21}{47}\cdot\frac{20}{46}\cdot\frac{19}{45}\cdot\frac{18}{44}\cdot\frac{17}{43} = \frac{\underline{|26}\,\underline{|42}\,\underline{|10}}{\underline{|52}\,\underline{|16}\,\underline{|10}} = \frac{\beta_{26}(10)}{\beta_{52}(10)} = \frac{19}{56588} = \frac{1}{2978},$$

the $\beta_n(x)$ being binomial functions.

Example 3. Compute the probability that a man whose age is a will be still alive after n years, and that he will die in one of the succeeding m years.

If we suppose that q_i is the probability that a man whose age is i will die before his next birthday, the probability that the man whose age is a will be alive at the end of n years will be

$$P_n = (1-q_a)(1-q_{a+1})\dots(1-q_{a+n-1}).$$

The probability Q_m of his *then* dying in either one or the other of the succeeding m years will be

$$Q_m = q_{a+n} + (1-q_{a+n})\{q_{a+n+1} + (1-q_{a+n+1})[q_{a+n+2} + \dots + (1-q_{a+n+m-2})q_{a+n+m-1}]\};$$

or

$$1 - Q_m = (1-q_{a+n})(1-q_{a+n+1})\dots(1-q_{a+m+n-1}).$$

The required probability of death after n years, but before the elapse of $n+m$ years, is consequently $P_n Q_m = P_n - P_{n+m}$.

The most convenient form for statements of mortality is not, as we here supposed, a table of the probabilities q_i for all integral ages i, but of the absolute frequencies l_i of the men from a large (properly infinitely large) population who will reach the age of i. After this $q_i = \frac{l_i - l_{i+1}}{l_i}$, $\left(1 - q_i = \frac{l_{i+1}}{l_i}\right)$ will only be a special case of the general answer:

$$P_n Q_m = \frac{l_{a+n} - l_{a+n+m}}{l_a}.$$

Example 4. We imagine a game of cards arranged in such a way that each player, in a certain order, gets two cards of the well-shuffled pack, and wins or loses according as the sum of the points on his two cards is eleven or not. For 5 players we use, for instance, only the cards 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 of the same colour.

What then is the probability of h players (named beforehand) getting 11 and not any of the $5-h$ others?

Secondly, what probability, r_k, is there that the k^{th} player in succession will be the first who gets 11?

Lastly, what is the probability, q, that none of the players will get 11?

It will be found perhaps that it is not quite easy to compute these probabilities directly. In such cases it is a good plan to reconnoitre the problem by first bringing out such results as present themselves quite easily and simnply, without considering whether they are just those we require. In this case, for instance, we take the probabilities, p_i, that each of the first i players will get 11.

We then attack the problem more seriously, and examine if there are not any simple functions of the probabilities we have found, p_i, which may be interpreted as probabilities of the same or similar sort as those inquired after.

$$q = \tfrac{344}{945} = 10(p_2 - p_3) + 5(p_4 - p_1) + p_0 - p_5.$$

§ 68. Repetitions of the same trial occur very frequently in problems solvable by the theory of probabilities, and should always by treated by means of a very simple and important law, the polynomial formula.

Let us suppose that the various events of the single trial may be indicated by colours, and that, in the single trial, the probability of white is w, of black b, and of red r.

The probability that we shall get in $x+y+z$ trials x white, y black, and z red results, in a given order, is then

$$w^x \cdot b^y \cdot r^z.$$

The number of the events of this kind that differ only in order, is the trinomial coefficient

$$\tau(x, y, z) = \frac{1 \cdot 2 \cdot 3 \ldots (x+y+z)}{1 \cdot 2 \ldots x \cdot 1 \cdot 2 \ldots y \cdot 1 \cdot 2 \ldots z},$$

which is the coefficient of the term $w^x \cdot b^y \cdot r^z$ in the development of $(w+b+r)^{(x+y+z)}$. And this same term

$$\tau(x, y, z)\, w^x \cdot b^y \cdot r^z \tag{126}$$

is the required probability of getting white x times, black y times, and red z times by $(x+y+z)$ repetitions.

When the probabilities of all possible single results are known and employed, so that $w+b+r+\ldots = 1$, and when the number of repetitions is n, we must consequently imagine $(w+b+r+\ldots)^n$ developed by the polynomial theorem, and the single terms of the development will then give us the probabilities of the different possible events of the repetitions without regard to the order of succession.

Example 1. If the question is of the probability of getting, in 10 trials in which there are the three possible events of white, black, and red, even numbers x, y, and z of each colour, and if the probabilities of the single events are w, b, and r, respectively, then we must retain the terms of $(w+b+r)^{10}$ which have even indices, and we thus find:

$$w^{10}+45w^8(b^2+r^2)+210w^6(b^4+6b^2r^2+r^4)+210w^4(b^6+15b^4r^2+15b^2r^4+r^6)+$$
$$+45w^2(b^8+28b^6r^2+70b^4r^4+28b^2r^6+r^8)+b^{10}+45b^8r^2+210b^6r^4+210b^4r^6+45b^2r^8+r^{10} =$$
$$= \tfrac{1}{4}\left\{(w+b+r)^{10}+(-w+b+r)^{10}+(w-b+r)^{10}+(w+b-r)^{10}\right\} =$$
$$= \tfrac{1}{4}\left\{1+(1-2w)^{10}+(1-2b)^{10}+(1-2r)^{10}\right\}.$$

The probability, consequently, is always greater than $\tfrac{1}{4}$, but only a little greater, unless the probability of getting some of the events in a single trial, is very small.

Example 2. Peter and Paul play at heads-or-tails (i. e. probability $= \tfrac{1}{2}$ for and against). But Peter throws with 3 coins, Paul only with 2, and the one wins who gets the greatest number of "heads". If both get the same number of heads they throw again, as often as may be necessary. What is the probability that Peter will win?

If we write for Peter's probability for and against throwing heads $p_1 = \frac{1}{2}$ and $q_1 = \frac{1}{2}$, for Paul's $p_2 = \frac{1}{2}$ and $q_2 = \frac{1}{2}$, then we should develop $(p_1+q_1)^3 \cdot (p_2+q_2)^2$, and the terms in which the index of p_1 is greater than that of p_2, are in favour of Peter; those in which the indices are equal, give a drawn game; and those in which the index of p_2 is greater than that of p_1, are in favour of Paul. For the single game there is the probability

for Peter of $\frac{8}{16}$,
for a drawn game of $\frac{5}{16}$,
for Paul of $\frac{3}{16}$.

As the probabilities are distributed in the same way, when they play the games over again, we need not consider the possibilities of drawn games at all, and we find $\frac{8}{11}$ as Peter's final probability.

Example 3. A game which is won once out of four times, is repeated 10 times. What is the probability of winning at most 2 of these?

$$\frac{551124}{1048576}.$$

§ 69. It often occurs that we inquire in a general way concerning a probability, which is a function of one or more numbers. Often it is also easier to transform a special problem into such a one of a more general character, where the unknown is a whole table $p_1, p_2, p_3, \dots p_n$ of probabilities, the suffixes being the arguments of the table. And then we must generally work with implicit equations, $f(p_1, \dots p_n) = 0$, particularly such as hold good for an arbitrary value of n, i. e. with difference-equations. Integration of finite difference-equations is indeed of so great importance in the art of solving problems of the theory of probabilities, that we can almost understand that Laplace has treated this method almost as the one to be used in all cases, in fact as the scientific quintessence of the theory of probabilities.

Since finite difference-equations like differential equations cannot as a rule be integrated by known functions, we can in an elementary treatise deal only with the simplest cases, especially such as can be solved by exponential functions, namely the linear difference-equations with constant coefficients. As to these, it is only necessary to mention here that, when

$$c_m p_{n+m} + \dots + c_0 p_n = 0 \quad (n \text{ being arbitrary}),$$

the solution is given by

$$p_n = k_1 r_1^n + \dots + k_m r_m^n, \tag{127}$$

where $r_1, \dots r_m$ are the roots in the equation

$$c_m r^m + \dots + c_0 = 0,$$

while $k_1, \dots k_m$ are integration-constants whenever the corresponding roots occur singly; but rational integral functions with arbitrary constants, and of the degree $i-1$, if the corresponding root occurs i times.

I shall mention one other means, however, not only because it can really lead to the integration of many of the difference-equations which the theory of probabilities leads to, particularly those in which the exponential functions occur in connection with binomial functions and factorials, but also because it has played an important part in the conception of this book.

The late Professor L. Oppermann, in April 1871, communicated to me a method of transformation, which I shall here state with an unessential alteration.

A finite or infinite series of numbers

$$u_0, u_1, \dots u_n$$

can univocally be expressed by another:

$$\left.\begin{aligned}
w_0 &= u_0 + u_1 + u_2 + u_3 + u_4 + \dots \\
w_1 &= \quad - u_1 - 2u_2 - 3u_3 - 4u_4 - \dots \\
w_2 &= \qquad\qquad u + 3u_3 + 6u_4 + \dots \\
w_3 &= \qquad\qquad\quad - u_3 - 4u_4 - \dots \\
w_4 &= \qquad\qquad\qquad\qquad u_4 + \dots \\
w_z &= (-1)^z \Sigma \beta_p(z) u_p,
\end{aligned}\right\} \tag{128}$$

where the sum Σ may be taken from $-\infty$ to $+\infty$, provided that $u_p = 0$ when $p > n$. In order, vice versâ, to compute the u's by means of the w's, we have equations of just the same form:

$$\left.\begin{aligned}
u_0 &= w_0 + w_1 + w_2 + w_3 + w_4 + \dots \\
u_1 &= \quad - w_1 - 2w_2 - 3w_3 - 4w_4 - \dots \\
u_2 &= \qquad\qquad w_2 + 3w_3 + 6w_4 + \dots \\
u_3 &= \qquad\qquad\quad - w_3 - 4w_4 - \dots \\
u_4 &= \qquad\qquad\qquad\qquad w_4 + \dots \\
u_x &= (-1)^x \Sigma \beta_r(x) w_r.
\end{aligned}\right\} \tag{129}$$

Here, as in (17) and (18), the general dependency between the u_i and w_j can be expressed in a single equation, be means of an independent variable z. From (129) we get identically

$$u_0 + u_1 e^z + u_2 e^{2z} + \dots = w_0 + (1-e^z) w_1 + (1-e^z)^2 w_2 + \dots$$

If we here put $1 - e^z = e^y$, then $1 - e^y = e^z$ will reduce (128) to an equation of the same form.

If u_i is the frequency or probability of i taken as an observed value, then also

$$u_0 + u_1 e^z + u_2 e^{2z} + \dots = s_0 + \frac{s_1 z}{1} + \frac{s_2 z^2}{1 \cdot 2} + \dots =$$
$$= s_0 e^{\frac{\mu_1 z}{1} + \frac{\mu_2 z^2}{1.2} + \dots} = w_0 + (1-e^z) w_1 + (1-e^z)^2 w_2 + \dots$$

illustrate the relations of the values in Oppermann's transformation to the half-invariants and sums of powers. In particular we have

$$\mu_1 = -\frac{w_1}{w_0}$$
$$\mu_2 = 2\frac{w_2}{w_0} - \frac{w_1(w_0 + w_1)}{w_0^2}$$
$$\mu_3 = -6\frac{w_3}{w_0} + 6\frac{w_2(w_0 + w_1)}{w_0^2} - \frac{w_1(w_0 + w_1)(w_0 + 2w_1)}{w_0^3}.$$

If now $u_0, u_1, \dots u_n$ are a series of probabilities or other quantities which depend on their suffix according to a fixed law, and if we know this law only through a difference-equation, then Oppermann's transformation of course leads only to a difference-equation for $w_0, w_1, \dots w_n$ as function of their suffix. But it turns out that, in problems of probabilities, this equation pretty often is easier to deal with than the original one (for instance the more difficult ones in Laplace's collection of problems). If we can look upon a probability u_i as the functional law of errors for i as the observed value, then w expresses the same law of errors by symmetrical functions, and frequently we want nothing more. If we have to reverse the process to find u_i itself, the series are pretty simple if w is simple; but they are often less favourable for numerical computation, as they frequently give the unknown as a difference between much larger quantities. There exists a means of remedying this, but it would carry us too far to enter into a closer examination of the question here.

Example 1. I throw a die, and go on throwing till I either win by getting "one" twice, or lose by throwing "two" or "three". If the game is to be over at latest by the n^{th} throw, what is my probability of winning? If the number of throws is unlimited, what is the probability of another "one" appearing before any "two" or "three"?

Four results are to be distinguished from one another. At any throw, say the i^{th}, the game can in general be won, lost, half won (by only one "one"), or drawn. Let the probability of the i^{th} throw resulting in a win be p_i, of the same resulting in a loss be q_i, in half win s_i, and in a drawn game be r_i, then $p_1 = 0$, $q_1 = \frac{1}{3}$, $s_1 = \frac{1}{6}$, and $r_1 = \frac{1}{2}$. Thus the probability of a second throw is $\frac{2}{3}$, and, generally, the probability of an i^{th} throw $s_{i-1} + r_{i-1}$. It is easy to express p_i, q_i, r_i, and s_i in terms of r_{i-1} and s_{i-1}, and also

p_{i-1}, q_{i-1}, r_{i-1}, and s_{i-1} in terms of r_{i-2} and s_{i-2}, etc. By elimination then the difference-equations can be found.

When we replace p or q or s or r by x the difference-equation can be written in the common form

$$x_i - x_{i-1} + \tfrac{1}{4}x_{i-2} = 0,$$

which is integrated as

$$x_i = (a+bi)\,2^{-i};$$

for r we have the simpler form

$$r_i = \tfrac{1}{2}r_{i-1}.$$

When, by the probabilities of the first throws, we have determined the constants, we get

$$p_i = \frac{i-1}{9}\,2^{-i},$$

$$q_i = \frac{2i+4}{9}\,2^{-i},$$

$$s_i = \frac{i}{3}\,2^{-i},$$

and

$$r_i = 2^{-i}.$$

We then have the formulae $P_n = p_1 + \dots + p_n$ and $Q_n = q_1 + \dots + q_n$, for the probabilities of making the winning or losing throw, and we get

$$\frac{P_n}{P_n+Q_n} = \frac{1}{3}\cdot\frac{(2^n-1)-n}{3(2^n-1)-n} \quad\text{and}\quad \frac{P\infty}{P\infty+Q\infty} = \frac{1}{9}.$$

Example 2. In a game the probability of winning is ϖ. The same game is repeated a great many, n, times. If it then happens at least once in this series that m successive games are won, you get a prize. What is the probability p_n of this? In a game of dice, where $\varpi = \frac{1}{6}$, what is the probability of getting a series of 5 "sixes" in 10000 throws?

It will be simplest to find the probability, $q_r = 1 - p_r$, that the prize will not be got in the first r repetitions. The difference-equation for this is

$$q_{r+m+1} - q_{r+m} + (1-\varpi)\,\varpi^m q_r = 0 \qquad \text{(a)}$$

or

$$\varpi^r c_0 = q_{r+m} - (1-\varpi)\left\{q_{r+m-1} + \varpi q_{r+m-2} + \dots + \varpi^{m-2} q_{r+1} + \varpi^{m-1} q_r\right\} = 0, \qquad \text{(b)}$$

where (b) is the first integral of (a). (As well as (a) we can directly demonstrate (b). How?). Hence

$$q_r = c_1\rho_1^r + \dots + c_m\rho_m^r,$$

where $c_1, \dots c_m$ are constants, which as well as $c_0 = 0$ must be determined by means of

$q_0 = q_1 = \ldots = q_{m-1} = 1$, $q_m = 1 - \varpi^m$, and ρ_1 to ρ_m are the roots of an irreducible equation of the m^{th} degree, which is got from

$$\rho^{m+1} - \rho^m = \varpi^{m+1} - \varpi^m \tag{c}$$

by dividing out $\rho - \varpi$. The largest of these roots (for small ϖ's or large m's) will be only a little less than 1; a small negative root occurs when m is even; the others are always imaginary, and they are also small.

In the actual computation it is highly desirable to avoid the complete solution of (c). This can be done, and this problem will illustrate a most important artifice. We may use the difference-equation to compute a single value of the unknown function by means of those which are known to us from the conditions of the problem, and then successive values of the unknown function by means of those already obtained; here, for instance, (b) enables us to get q_{m+1} in terms of $q_1, \ldots q_m$. Then we get q_{m+2}, either by again applying (b) to $q_2, \ldots q_{m+1}$ or by applying (a) to q_1 and q_{m+1} (or best in both ways for the sake of the check), etc.

It is evident that the table of the numerical values of the function which we can form in this way, cannot easily become of any great extent or give us exact information as to the form of the function. But we are able to interpolate, and, when the general form of the function is known (as here), we may be justified in using extrapolation also. In our example we need only continue the computations above described until the term in $q_r = c_1\rho_1^r + \ldots$, corresponding to the greatest root ρ_1, dominates the others to such a degree that the first difference of $\operatorname{Log} q_r$ becomes constant, and the computation of q_r for higher indices can then be made as by a simple geometrical progression. In the numerical case $q_r = 1{\cdot}004078 \times (0{\cdot}9998928)^r$; $1 - q_{10000} = 0{\cdot}6577$.

Example 3. A bag contains n balls, a white and $a - n$ black ones. A ball is drawn out of the bag and a black ball then placed in it, and this process is repeated y times. After the y^{th} operation the white and black balls in the bag are counted. Find the probability $u^x(y)$ that the numbers of white balls will then be x and the black ones $n - x$.

We have

$$u_x(y) = \frac{n-x}{n} u_x(y-1) + \frac{x+1}{n} u_{x+1}(y-1)$$

and

$$u_x(0) = 0, \text{ except } u_a(0) = 1.$$

By Oppermann's transformation we find

$$w_z(y) = (-1)^z \Sigma \beta_x(z) \cdot u_x(y),$$

Σ taken from $x = -\infty$ to $x = +\infty$, or

$$w_z(y) = (-1)^z \Sigma \frac{n-x}{n} \beta_x(z) \cdot u_x(y-1) + (-1)^z \Sigma \frac{x+1-z}{n} \beta_{x+1}(z) \cdot u_{x+1}(y-1).$$

The limits of x under Σ being infinite, $x+1$ can be replaced by x, consequently

$$w_z(y) = \frac{n-z}{n} w_z(y-1).$$

This difference-equation, in which y is the variable, may easily be integrated. As we have, further,

$$w_z(0) = (-1)^z \beta_a(z),$$

we get

$$w_z(y) = (-1)^z \cdot \beta_a(z) \cdot \left(\frac{n-z}{n}\right)^y.$$

By Oppermann's inverse transformation we find now:

$$u_x(y) = (-1)^x \Sigma \beta_z(x) \cdot (-1)^z \cdot \beta_a(z) \cdot \left(\frac{n-z}{n}\right)^y,$$

Σ taken from $x = -\infty$ to $x = +\infty$. This expression

$$u_x(y) = \beta_a(x) \Sigma (-1)^{z+x} \cdot \beta_{a-x}(z-x) \cdot \left(\frac{n-z}{n}\right)^y$$

has the above mentioned practical short-comings, which are sensible particularly if n, $a-x$, or y are large numbers; in these cases an artifice like that used by Laplace (problem 17) becomes necessary. But our exact solution has a simple interpretation. The sum that multiplies $\beta_a(x)$ in $u_x(y)$, is the $(a-x)^{\text{th}}$ difference of the function $\left(\frac{n-z}{n}\right)^y$, and is found by a table of the values $\left(\frac{n-a}{n}\right)^y$, $\left(\frac{n-a+1}{n}\right)^y$, ... $\left(\frac{n-x-1}{n}\right)^y$, $\left(\frac{n-x}{n}\right)^y$, as the final difference formed by all these consecutive values. We learn from this interpretation that it is possible, if not easy, to solve this problem without the integration of any difference-equation, in a way analogous to that used in § 67, example 4.

If we make use of $w_z(y)$ to give us the half-invariants μ_1, μ_2 for the same law of errors as is expressed by $u_x(y)$, then we find for the mean value of x after y drawings

$$\lambda_1(y) = a\left(\frac{n-1}{n}\right)^y$$

and for the square of the mean error

$$\lambda_2(y) = a\left(\left(\frac{n-1}{n}\right)^y - \left(\frac{n-2}{n}\right)^y\right) + a^2\left(\left(\frac{n-2}{n}\right)^y - \left(\frac{n-1}{n}\right)^{2y}\right).$$

XVI. THE DETERMINATION OF PROBABILITIES A PRIORI AND A POSTERIORI.

§ 70. The computations of probabilities with which we have been dealing in the foregoing chapters have this point in common that we always assume one or several probabilities to be given, and then deduce from them the required ones. If now we ask, how