AN APPLICATION OF THE HIPERGEOMETRIC SCHEME

Cristian-George CONSTANTINESCU[*], Ion CIOACĂ[**]

[*]"Henri Coandă" Air Force Academy, Braşov, Romania
(c.g.constantinescu@gmail.com)
[**]Secondary School, Buneşti, Romania (puiu_cioaca2007@yahoo.com)

DOI: 10.19062/1842-9238.2017.15.2.12

***Abstract**: The purpose of this scientific approach is to report an error. We probably are not the first to notice it, but it's very possible that we are the first to make it public. This paper presents the problem VIII.8, from the book "Problems of mathematics for engineers", by Rodica Trandafir, with the author's solution. It is demonstrated that this is wrong and the right solution is proposed using hypergeometric scheme.*

***Keywords:** hypergeometric scheme, probabilistic approach, distribution, Mathcad*

1. INTRODUCTION

1.1 The hypergeometric scheme (of the urn with the unreturned ball):

In a urn there are *A* white balls and *B* black balls, from which *n* balls are extracted, $n < N := A + B$, one by one, without returning the drawn ball into the urn (or equivalently, all the *n* balls are removed simultaneously). The probability that among the *n* balls we have *a* white balls and $n - a := b$ balls is [1]:

$$P(a,b) = \frac{C_A^a C_B^b}{C_{A+B}^{a+b}}, \text{ where } \max\{0, A+n-N\} \le a \le \min\{A, n\} \quad (1)$$

1.2 Generalization:

If there are balls colored in *m* colors in the urn: A_1 balls of c_1 color, A_2 balls of c_2 color, ..., A_m balls of c_m color and *n* balls are extracted one by one without returning the ball drawn into the urn (or, equivalently, all the *n* balls are removed simultaneously), then the probability that form the *n* balls we have a_1 balls of c_1 color, a_2 balls of c_2 color,... a_m balls of c_m color,...where $a_1 + a_2 + \ldots + a_m = n$ is:

$$P(a_1, a_2, \ldots, a_m) = \frac{C_{A_1}^{a_1} C_{A_2}^{a_2} \ldots C_{A_m}^{a_m}}{C_{A_1+A_2+\ldots+A_m}^{a_1+a_2+\ldots+a_m}} \quad (2)$$

2. THE WRONG SOLUTION

The problem in discussion is formulated as follows: *A urn contains 36 white balls and 12 black balls. One person pulls the balls one by one until 10 white balls are obtained. Calculate the mean value of the number of extracted black balls* [2].

The solution proposed by the author is detailed below.

We are in the case of the unbalanced ball. The probability of removing *10* white balls and *k* black balls is:

$$P = \frac{C_{36}^{10} \cdot C_{12}^{k}}{C_{48}^{10+k}} \tag{3}$$

The mean value of the number of extracted black balls is:

$$M = C_{36}^{10} \cdot \sum_{k=0}^{12} k \frac{C_{12}^{k}}{C_{48}^{10+k}} \tag{4}$$

It is true that we are in the case of the unreturned ball, but the probability computed above is wrong. It respects the condition of having 10 white balls, but does not respect what is meant by the context: the last ball drawn out is white. Probability P also includes the solution that the last ball is black, or when the 10th white ball is extracted, the experience ends. For example, we calculated the sum of the probabilities associated with the random variable Ω that counts the black balls drawn from the urn, according to the hypothesis of the problem:

$$\Omega : \begin{pmatrix} 0 & 1 & 2 & 3 & \ldots & 12 \\ \frac{C_{36}^{10} \cdot C_{12}^{0}}{C_{48}^{10+0}} & \frac{C_{36}^{10} \cdot C_{12}^{1}}{C_{48}^{10+1}} & \frac{C_{36}^{10} \cdot C_{12}^{2}}{C_{48}^{10+2}} & \frac{C_{36}^{10} \cdot C_{12}^{3}}{C_{48}^{10+3}} & \ldots & \frac{C_{36}^{10} \cdot C_{12}^{12}}{C_{48}^{10+12}} \end{pmatrix} \tag{5}$$

$$\sum_{k=0}^{12} p_k = C_{36}^{10} \cdot \sum_{k=0}^{12} \frac{C_{12}^{k}}{C_{48}^{10+k}} = \frac{36}{10 \cdot 26} \cdot \sum_{k=0}^{12} \frac{\frac{12}{k! \cdot (12-k)!}}{\left(\frac{48}{(10+k)! \cdot (38-k)!} \right)} = 1.324 \tag{6}$$

3. THE CORRECT SOLUTION

Below the correct solution is presented. It can be observed that we have modified the data of the problem, which obviously does not influence the correctness of the solution:

In a urn we have 20 white balls and 80 black balls. Balls are extracted from the urn (without turning back the extracted ball) until 10 white balls are obtained. Calculate the mean and variance of the number of extracted black balls.

The correct solution is detailed below.

It is considered a urn that contains N_A white balls and N_B black balls, from which we extract the balls until n_A white balls are obtained, where $0 \le n_A \le N_A$

Let Ω be the variable describing the number of extracted black balls. It can be seen that ***the last ball*** drawn in the experience ***can only be white***. It follows that event A: $\Omega = k$ (we extracted *k* black balls to get the n_A white balls, so a total amount of $k + n_A$ balls were extracted from the urn) is written as the intersection of the following two events, obviously independent ones:

A_1: in the first $k + n_A - 1$ extractions $n_A - 1$ white balls and *k* black balls were obtained (regardless of their order);

A_2 : in the last extraction a white ball was obtained.

$$\left.\begin{array}{c} A = A_1 \cap A_2 \\ A_1, A_2 \text{ independente} \end{array}\right\} \Rightarrow P(A) = P(A_1 \cap A_2) = P(A_1)\cdot P(A_2) = P(\Omega = k) \tag{7}$$

Event A_1 is described by the hypergeometric scheme associated with the urn containing N_A white balls and N_B black balls, from which we extract $k + n_A - 1$ balls, such that $n_A - 1$ balls are white and k balls are black.

If $N = N_A + N_B$, it results:

$$P(A_1) = \frac{C_{N_A}^{n_A-1} \cdot C_{N_B}^{k}}{C_N^{n_A+k-1}} \tag{8}$$

Event A_2 is described by the hypergeometric scheme associated with the urn with $N_A - n_A + 1$ white ball and $N_B - k$ black balls, from which we draw a ball and it must be white.

$$P(A_2) = \frac{C^1_{N_A-n_A+1} \cdot C^0_{N_B-k}}{C^1_{N_A-k-n_A+1}} = \frac{N_A - n_A + 1}{N_A - k - n_A + 1} \tag{9}$$

$$P(\Omega = k) = \frac{C_{N_A}^{n_A-1} C_{N_B}^{k}}{C_N^{n_A+k-1}} \cdot \frac{N_A - n_A + 1}{N - k - n_A + 1}; k = \overline{0, N_B} \tag{10}$$

In the proposed problem we have the particular case: $N_A = 20$; $N_B = 80$; $n_A = 10$. The following random variable Ω distribution array results:

$$\Omega : \begin{pmatrix} 0 & 1 & 2 & \ldots & k & \ldots & 80 \\ \frac{C_{20}^9 C_{80}^0}{C_{100}^9} \cdot \frac{11}{91} & \frac{C_{20}^9 C_{80}^1}{C_{100}^{10}} \cdot \frac{11}{90} & \frac{C_{20}^9 C_{80}^2}{C_{100}^{11}} \cdot \frac{11}{89} & \ldots & \frac{C_{20}^9 C_{80}^k}{C_{100}^{n_A+k-1}} \cdot \frac{11}{91-k} & \ldots & \frac{C_{20}^9 C_{80}^{80}}{C_{100}^{89}} \cdot \frac{11}{11} \end{pmatrix} \tag{11}$$

It may be observed that in this case the sum of probabilities is equal to 1:

$$Na := 80 \qquad Nb := 20 \qquad na := 10$$

$$\sum p_k = \sum_{k=0}^{Nb} \frac{\frac{Na!}{(na-1)!\cdot(Na-na+1)!} \cdot \frac{Nb!}{k!\cdot(Nb-k)!}}{\left(\frac{(Na+Nb)!}{(na+k-1)!\cdot(Na+Nb-na-k+1)!}\right)} \cdot \frac{Na-na+1}{Na+Nb-k-na+1} = 1 \tag{12}$$

$$m := M(\Omega) = 11 \cdot C_{20}^9 \cdot \sum_{k=0}^{80} k \frac{C_{80}^k}{C_{100}^{9+k}(91-k)} \tag{13}$$

$$m = 11 \cdot \frac{20}{9! \cdot 11} \cdot \sum_{k=0}^{80} \frac{k \cdot \frac{80}{k! \cdot (80-k)!}}{(91-k) \cdot \frac{100}{(9+k)! \cdot (91-k)!}} = 38.095 \tag{14}$$

$$M(\Omega^2) = 11 \cdot C_{20}^{9} \cdot \sum_{k=0}^{80} k^2 \frac{C_{80}^{k}}{C_{100}^{9+k}(91-k)} \tag{15}$$

$$M(\Omega^2) = 11 \cdot \frac{20}{9! \cdot 11} \cdot \sum_{k=0}^{80} \frac{k^2 \cdot \frac{80}{k! \cdot (80-k)!}}{(91-k) \cdot \frac{100}{(9+k)! \cdot (91-k)!}} = 1.543 \cdot 10^{3} \tag{16}$$

$$\sigma^2 := D^2(\Omega) = M(\Omega^2) - m^2 \tag{17}$$

$$\sigma^2 =$$

$$11 \cdot \frac{20}{9! \cdot 11} \cdot \left[\sum_{k=0}^{80} \frac{k^2 \cdot \frac{80}{k! \cdot (80-k)!}}{(91-k) \cdot \frac{100}{(9+k)! \cdot (91-k)!}} - 11 \cdot \frac{20}{9! \cdot 11} \cdot \left[\sum_{k=0}^{80} \frac{k \cdot \frac{80}{k! \cdot (80-k)!}}{(91-k) \cdot \frac{100}{(9+k)! \cdot (91-k)!}} \right]^2 \right] = 91.61 \tag{18}$$

In formulas 12-18 the calculations were performed by using the Mathcad program.

CONCLUSIONS

Different extensions of this problem can be formulated and some of them can be reducible to the case of the two-color urn. One of them might be the following: *In a urn we have N_1 balls of c_1 color, N_2 balls of c_2 color, ... and N_m balls of c_m color. Balls from the urn (without returning the extracted ball) are extracted until n_p balls of c_p color are obtained. Calculate the average and dispersion of the number of other color balls extracted*. The problem is reduced to that presented at point 3, considering the urn with N_p balls of c_p color and $N - N_p$ balls of another color.

Another observation that can be made: the probabilistic approach is a method of calculating sums with combinations. For example, this problem has shown that:

$$\sum_{k=0}^{N_B} \frac{C_{N_A}^{n_A-1} C_{N_B}^{k}}{C_{N_A+N_B}^{n_A+k-1}} \cdot \frac{N_A - n_A + 1}{N - k - n_A + 1} = C_{N_A}^{n_A-1}(N_A - n_A + 1) \sum_{k=0}^{N_B} \frac{C_{N_B}^{k}}{C_{N_A+N_B}^{n_A+k-1}} \cdot \frac{1}{N - k - n_A + 1} = 1 \tag{19}$$

REFERENCES

[1] Gh. Mihoc, N. Micu, *Probability Theory and Mathematical Statistics*, Didactic and Pedagogical Publishing House, 1980.

[2] R. Trandafir, *Problems of Mathematics for Engineers*, Technical Publishing House, 1977.