A probability problem suitable for Problem-Based Learning

Un problema de probabilidad adecuado para el Aprendizaje Basado en Problemas

J. L. González-Santander[1*]

[1] Department of Mathematics. University of Oviedo.
[*] Correspondecia: University of Oviedo. Department of Mathematics. Calle Federico García Lorca, 18. 33007 Oviedo. Asturias. Spain. *E-mail:* gonzalezmarjuan@uniovi.es. ORCID ID: 0000-0001-5348-4967

ABSTRACT

We propose a simple probability problem for undergraduate level. This problem involves different branches of Mathematics, such as graph theory, linear algebra and hypergeometric sums hence it is quite suitable to be used as an example of Problem-Based Learning. In addition, the problem allows several variations so that it may be introduced to different groups of students at the same time.

KEYWORDS: *Problem-Based Learning, vertex matrix, Chu-Vandermonde summation formula.*

RESUMEN

Se propone un sencillo problema de probabilidad para nivel de grado universitario. Este problema involucra distintas ramas de las matemáticas, como la teoría de grafos, el álgebra lineal o las sumas hipergeométricas, por lo que resulta muy apto para ser usado como aprendizaje basado en problemas. Además, el problema admite diversas variantes, por lo que puede ser propuesto a diferentes grupos de estudiantes simultáneamente.

PALABRAS CLAVE: *Aprendizaje basado en problemas, matriz de incidencia, fórmula de suma de Chu-Vandermonde.*

INTRODUCTION

We propose a simple probability problem whose solution involves different mathematical areas. The main aim of this problem is to show to the students that although these areas are taught in separate subjects, in fact, they are quite connected to real problems. Therefore, it is quite suitable for Problem-Based Learning [1]. Next, we present the statement of the problem.

Problem If we choose a sequence of m natural numbers, a_1, a_2, …, a_m, at random, such that $1 \leq ai \leq n$, what is the probability of obtaining $a_1 \leq a_2 \leq \cdots \leq a_m$?

According to Laplace's first principle of probability [2], the problem leads straightforwardly to a combinatorial problem to count the number of "favoured events". This combinatorial problem can be tackled by using the vertex matrix of Graph Theory. Since we have to calculate a certain power of the vertex matrix, we may resort to Linear Algebra diagonalization. However, in this case, diagonalization fails, so that we must perform heuristic reasoning wherein we calculate a hypergeometric sum. Therefore, the solution to this problem entails Probability Theory, Combinatorics, Graph Theory, Linear Algebra, and hypergeometric sums.

Therefore, the problem stated above is quite suitable for Problem-Based Learning at undergraduate level, wherein the solution unfolds in different stages and students are encouraged to go over the properties and relations of different topics. Moreover, the problem admits different variations for which, the students may apply similar strategies.

FROM PROBABILITY TO LINEAR ALGEBRA

According to Laplace's first principle of probability, the probability p of a certain event to occur is the ratio of the "favoured events" f to the total possible events N, assuming that all events are equally likely. The latter assumption is reasonable since the sequence of natural numbers of the problem are chosen at random.

$$p = \frac{f}{N}. \tag{1}$$

The number of total possible events N is straightforwardly calculated by using the variations with repetition formula. Indeed, for each natural number a_i we have n possible natural numbers. Since we have a sequence of m numbers, then

$$N = VR_n^{\,m} = n^m \tag{2}$$

The number of favoured events is more complicated to calculate because it is subject to some restrictions, i.e. $a_1 \leq a_2 \leq \cdots \leq a_m$. For simplicity, we start considering $n = m = 3$. In this case, the 10 favoured events are listed below.

1-1-1, 1-1-2, 1-1-3, 1-2-2, 1-2-3,
1-3-3, 2-2-2, 2-2-3, 2-3-3, 3-3-3.

We can visualize these sequences as the allowed paths of length 2 in the directed graph shown in figure 1.
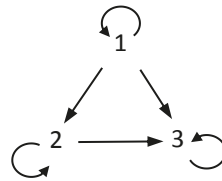
Fig.1. Graph for 3 vertices

The element matrix r_{ij} of the *vertex matrix R* is equal to the number of edges joining vertices *i* and *j* [3]. Thereby, the vertex matrix of the graph of figure 1 is

$$R_3 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Clearly, the vertex matrix for $n \in \mathbb{N}$ is the following triangular matrix of order *n*.

$$R_n = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}. \tag{3}$$

The number of different paths going from vertex *i* to vertex *j* of length *q* is the matrix element s_{ij}, where $S = R_n^{\,q}$, [4]. In our case, $q = m - 1$. Therefore, the number of favourable events *f* is given by the sum of all the matrix elements of $R_n^{\,m-1}$,

$$f = \sum_{i,j=1}^{n} \left(R_n^{m-1} \right)_{ij}. \tag{4}$$

In order to calculate the matrix $R_n^{\,m-1}$, we may attempt to get R_n in diagonal form, thus

$$R_n = PDP^{-1} \rightarrow \quad R_n^{\,m-1} = PD^{m-1}P^{-1},$$

where *P* is a square matrix with the eigenvectors set in columns and *D* is a diagonal matrix with the eigenvalues set in the main diagonal [4]. However, R_n cannot be obtained in diagonal form since

R_n has got one eigenvalue, $\lambda = 1$, and one eigenvector, $\overrightarrow{v} = (1,0,\ldots,0)$. Therefore, let us make another attempt with the following definition.

Definition. Consider the following set of sequences $a_n^{(0)}, a_n^{(1)}, \ldots a_n^{(m)}$ such that

$$
\begin{cases}
a_n^{(0)} = 1, & \forall n = 1, 2, \ldots \\
a_n^{(m+1)} = \displaystyle\sum_{k=1}^{n} a_k^{(m)}
\end{cases}
\tag{5}
$$

Theorem. The $m + 1$ power of the R_n matrix is given by

$$
R_n^{m+1} =
\begin{pmatrix}
a_1^{(m)} & a_2^{(m)} & \cdots & a_n^{(m)} \\
0 & a_1^{(m)} & \cdots & a_{n-1}^{(m)} \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & a_1^{(m)}
\end{pmatrix}
\tag{6}
$$

Proof. Let us prove (6) by induction. For $m = 0$, equation (6) follows straightforwardly

$$
R_n =
\begin{pmatrix}
a_1^{(0)} & a_2^{(0)} & \cdots & a_n^{(0)} \\
0 & a_1^{(0)} & \cdots & a_{n-1}^{(0)} \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & a_1^{(0)}
\end{pmatrix}
=
\begin{pmatrix}
1 & 1 & \cdots & 1 \\
0 & 1 & \cdots & 1 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 1
\end{pmatrix}.
$$

Assume that (6) is true for *m*. Let us prove that is also true for *m* + 1 we have

$$
R_n^{m+2} = R_n^{m+1} R_n = \begin{pmatrix} a_1^{(m)} & a_2^{(m)} & \cdots & a_n^{(m)} \\ 0 & a_1^{(m)} & \cdots & a_{n-1}^{(m)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_1^{(m)} \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}
$$

$$
= \begin{pmatrix} \sum_{k=1}^{1} a_k^{(m)} & \sum_{k=1}^{2} a_k^{(m)} & \cdots & \sum_{k=1}^{n} a_k^{(m)} \\ 0 & \sum_{k=1}^{1} a_k^{(m)} & \cdots & \sum_{k=1}^{n-1} a_k^{(m)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sum_{k=1}^{1} a_k^{(m)} \end{pmatrix} = \begin{pmatrix} a_1^{(m+1)} & a_2^{(m+1)} & \cdots & a_n^{(m+1)} \\ 0 & a_1^{(m+1)} & \cdots & a_{n-1}^{(m+1)} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_1^{(m+1)} \end{pmatrix},
$$

as we wanted to prove. ∎

Therefore, according to (4) and (6), we can sum up all the diagonals of the R_n^{m-1} matrix, to obtain

$$
f = \sum_{i,j=1}^{n} \left(R_n^{m-1} \right)_{ij} = \sum_{k=1}^{n} (n-k+1) a_k^{(m-2)}. \tag{7}
$$

In summary, our problem would be solved, if we could calculate the set of sequences $a_n^{(0)}, a_n^{(1)}, \dots$ $a_n^{(m)}$ satisfying (5). The latter leads us to calculate a hypergeometric sum.

A USEFUL HYPERGEOMETRIC SUM

Let us define first a hypergeometric series [5].

Definition. A hypergeometric series is a series $\sum c_k$ such that

$$
\frac{c_{k+1}}{c_k} = \frac{(k+\alpha_1)\cdots(k+\alpha_p)x}{(k+\beta_1)\cdots(k+\beta_q)(k+1)}. \tag{8}
$$

so that

$$\sum_{k=0}^{\infty} c_k = c_0 \sum_{k=0}^{\infty} \frac{(\alpha_1)_k \cdots (\alpha_p)_k}{(\beta_1)_k \cdots (\beta_q)_k} \frac{x^k}{k!} =: c_0 \, _pF_q \left(\begin{matrix} \alpha_1 \cdots \alpha_p \\ \beta_1 \cdots \beta_q \end{matrix} \middle| x \right),$$

(9)

where $(\alpha)_k = \Gamma(\alpha+k)/\Gamma(\alpha)$ denotes the Pochhammer symbol and $\Gamma(z)$ is the gamma function.

Lemma. The following summation formula holds true:

$$G_s(t) := \sum_{k=0}^{t-1} \frac{(k+s)!}{k!} = \frac{(s+t)!}{(s+1)(t-1)!}.$$

(10)

Proof. Perform variable substitution $\ell = t-1-k$ in (8) to obtain

$$G_s(t) = \sum_{\ell=0}^{t-1} \frac{(t-1-\ell+s)!}{(t-1-\ell)!} = \sum_{\ell=0}^{\infty} \frac{(t-1-\ell+s)!}{(t-1-\ell)!},$$

(11)

where the upper limit of the sum can be extended to infinity because $1/n! = 0$ for $n \in \mathbb{Z}^-$. According to (8-9), rewrite (11) as a hypergeometric series. Indeed,

$$\begin{cases} \dfrac{c_{\ell+1}}{c_\ell} = \dfrac{(\ell-t+1)(\ell+1)}{(\ell-t+1-s)(\ell+1)}, \\ c_0 = \dfrac{(t-1+s)!}{(t-1)!}. \end{cases}$$

Thereby,

$$G_s(t) = \frac{(t-1+s)!}{(t-1)!} \, _2F_1 \left(\begin{matrix} -(t-1),1 \\ -t+1-s \end{matrix} \middle| 1 \right).$$

(12)

Now, apply Chu-Vandermonde summation formula [5],

$$_2F_1 \left(\begin{matrix} -k,\alpha \\ \gamma \end{matrix} \middle| 1 \right) = \frac{(\gamma-\alpha)_k}{(\gamma)_k}, \quad k = 1,2,\ldots$$

and the following property of the Pochhammer symbol, [6] in combination with the factorial property of the gamma function, $\Gamma(n+1)=n!$, $n \in \mathbb{N}$, [6],

$$(-x)_k = (-1)^k (x-k+1)_k = (-1)^k \frac{\Gamma(x+1)}{\Gamma(x-k+1)} = (-1)^k \frac{x!}{(x-k)!},$$

thus, after some simplification, (12) becomes

$$G_s(t) = \frac{(t-1+s)!}{(t-1)!} \frac{(-t-s)_{t-1}}{(1-t-s)_{t-1}} = \frac{(t+s)!}{(s+1)(t-1)!},$$

as we wanted to prove. ∎

Corollary. The following summation formula holds true:

$$\sum_{k=1}^{t} \binom{k+s-1}{s} = \binom{t+s}{s+1}.$$

(13)

Proof. Divide both sides of (10) by $s!$ and rewrite the result in terms of binomial coefficients, shifting the sum index to start at $k = 1$. ∎

Notice that (13) fits in the set of sequences (5) we were looking for. Indeed, take

$$a_n^{(m)} = \binom{n+m-1}{m},$$

(14)

thus

$$\left| \begin{array}{l} a_n^{(0)} = \binom{n-1}{0} = 1, \quad \forall n = 1, 2, \ldots \\ a_n^{(m+1)} = \binom{n+m}{m+1} = \sum_{k=1}^{n} \binom{k+m-1}{m} = \sum_{k=1}^{n} a_n^{(m)}. \end{array} \right.$$

Once we have obtained (14), we calculate the number of favourable events from (7), thereby,

$$f = \sum_{k=1}^{n} (n-k-1) \binom{k+m-3}{m-2} = \frac{1}{(m-2)!} \sum_{k=0}^{n-1} (n-k) \frac{(k+m-2)!}{k!}.$$

Applying (10), we have

$$f = \frac{1}{(m-2)!}\Big[\, n\, G_{m-2}\,(n) - G_{m-1}\,(n-1)\Big]$$

$$= \frac{(m+n-2)!}{(m-2)!(n-2)!}\left(\frac{n}{(n-1)(m-1)} - \frac{1}{m}\right)$$

$$= \frac{(m+n-1)!}{m!(n-1)!} = \binom{m+n-1}{m}. \tag{15}$$

Curiously, note that, according to (14), $f = a_n^{(m)}$. As a consistency test, we recover the number of favoured events for $n = m = 3$, i.e. $f = a_3^{(3)} = \dfrac{5!}{3!2!} = 10$, as aforementioned in the Introduction.

Finally, substitute (2) and (15) in (1), to obtain the probability,

$$p = \frac{(n+m-1)!}{m!(n-1)!\, n^m}.$$

CONCLUSIONS

We have proposed a simple probability problem whose solution involves Graph Theory, Linear Algebra and hypergeometric sums, hence it is quite suitable for Problem-Based Learning. This problem admits several variations. For instance, we can modify the restrictions of the favoured events, i.e. $a_1 < a_2 < \cdots < a_m$, so that these variations may be proposed to different groups of students at the same time.

Also, it connects different branches of Mathematics in a very natural way. This is important in an undergraduate level of education since these topics are usually studied separately, but many times they appear intertwined in real problems.

LITERATURE CITED

[1]   Boud D, Feletti G. The Challenge of Problem-Based Learning. London: Psychology Press; 1997.
[2]   Laplace PS. Essai philosophique sur les probabilités. 5th ed. Cambridge: Cambridge University Press; 2009.
[3]   Brualdi RA. Introductory to Combinatorics. 5th. ed. Pearson Educational; 2009.
[4]   Anton H, Rorres C. Elementary Linear Algebra. 9th ed. John Wiley & Sons; 2005.
[5]   Andrews GE, Askey R, Roy R. Special Functions. Encyclopedia of Mathematics 71. Cambridge: Cambridge University Press; 1999.
[6]   Oldham K, Myland J, Spanier J. An Atlas of Functions. 2nd ed. New York: Springer; 2009.