Journal of Modern Applied Statistical Methods

Volume 3 | Issue 2 Article 18

11-1-2004

The President's Problem

Jann-Huei Jinn
Grand Valley State University, jinnj@gvsu.edu

Journal of Modern Applied Statistical Methods
November, 2004, Vol. 3, No. 2, 476-484

1538 – 9472/04/$95.00

Algorithms & Code
The President's Problem

Jann-Huei Jinn
Department of Statistics
Grand Valley State University

A solution is offered in response to a complex combination problem challenged by Blom, Englund, and Sandell (1998). The problem is to determine the probability that a random permutation of the word BILLCLINTON has no equal neighbors.

Key words: Combinatorics, equal neighbors, random permutations, run

Introduction

The problem is to determine the probability that a random permutation of the word BILLCLINTON has no equal neighbors. Choose an initial order of the letters in the word Billclinton, for example, IINNLLLBTCO. The problem is solved in three steps: Start with IINN, insert LLL, then insert B, T, C, and finally, insert O.

Methodology

Let X_1 be the number of equal neighbors in a random permutation of the four letters IINN. To obtain the solution, the probability function $P(X_1 = k)$ is needed. First, determine the probability function of the total number of runs (see references). Consider a random permutation of m 1's and n 0's. Denote by U the number of runs of 1's and by V the number of runs of 0's. The probability function $P(U = r, V = s)$ is needed. (Note that when $|r - s| > 1$, $P(U = r, V = s)$ =0.). The m 1's can be partitioned into r groups in $\binom{m-1}{r-1}$ ways.

Jann-Huei Jinn is a Professor in the Department of Statistics at Grand Valley State University. Allendale Michigan. Email: jinnj@gvsu.edu.

Similarly, the n 0's can be partitioned into s groups in $\binom{n-1}{s-1}$ ways. It is known that the number of permutations with r 1- runs and s 0-runs is the product of these two binomial coefficients when $|r - s|$ =1, and twice that product when $r = s$. Since the total number of permutations is $\binom{m+n}{m}$, obtained is

$$P(U = r, V = s) = \frac{\binom{m-1}{r-1}\binom{n-1}{s-1}}{\binom{m+n}{m}} \tag{1.1}$$

for r =1,2,…,m and s =1,2,…,n such that $|r - s|$ =1. If $r = s$, then

$$P(U = r, V = r) = \frac{2\binom{m-1}{r-1}\binom{n-1}{r-1}}{\binom{m+n}{m}} \tag{1.2}$$

Step 1. Consider the permutation of IINN, we know that m=n=2 and Table 1 gives the probabilities $P(U = r, V = s)$ for this case.

Table 1: m=n=2 Probability Function P(U=r,V=s).

r\s	1	2
1	2/6	1/6
2	1/6	2/6

The probability function $P(U+V=k)$ of the total number of runs is obtained from the previous distribution (1.1) by summation. The result when m=n=2 is given in Table 2.

Table 2: m=n=2 Probability Function P(U+V=k) of the Total Number of Runs

k	2	3	4
P(U+V=k)	2/6	2/6	2/6

Let W be the number of equal neighbors in the random permutation. The relation between runs and equal neighbors is W = m+n-U-V. Hence, when m=n=2, the probability of 2 equal neighbors is equal to the probability 2/6 of two runs, the probability of 1 equal neighbors is 2/6, and so on. Therefore, X_1=4-(U+V). The probability function of X_1 required for the solution of the Statistician problem is given in Table 3.

Table 3: Probability Function of X_1.

i	0	1	2
$P(X_1=i)$	2/6	2/6	2/6

Step 2. Insert LLL. Let X_2 be the number of equal neighbors among the seven letters, IINNLLL, obtained. Since

$$P(X_2=j)=\sum_i P(X_2=j|X_1=i)P(X_1=i)\,. \tag{1.3}$$

Consider j=0,1,2,3,4. Four letters B ,T ,C ,and O can be inserted in the case of "IINNLLL" to get no equal neighbors. The conditional probabilities required are given in Table 4.

Table 4: Some Conditional Probabilities $P(X_2=j|X_1=i)$

i\j	0	1	2	3	4
0	60/210	120/210	30/210	0	0
1	36/210	72/210	78/210	24/210	0
2	18/210	48/210	78/210	48/210	18/210.

All the probabilities given in Table 4 were calculated. Suppose that X_1=0, for example, ININ, the first L is inserted, the second L, and then the third L. It does not matter to insert the first L at any place, for example, if the first L is inserted to the right end of ININ, i.e., ININL, then !I!N!I!N!L! is obtained (where "!" represent the space to insert the second L). It is convenient to use a tree diagram to do illustration; see Figure 1.

Step 3. Insert B,T,C, and O (The order to insert).

Let X_3 be the number of equal neighbors among the eleven letters obtained.

$$P(X_3=0)=\sum_{j=0}^{4}P(X_3=0|X_2=j)P(X_2=j) \tag{1.4}$$

When X_2=0, for example, LILINLN, it does not matter where to insert B,T,C,O, there are no equal neighbors. Therefore, $P(X_3=0|X_2=0)=1$.

When X_2=1, for example, LIINLNL,

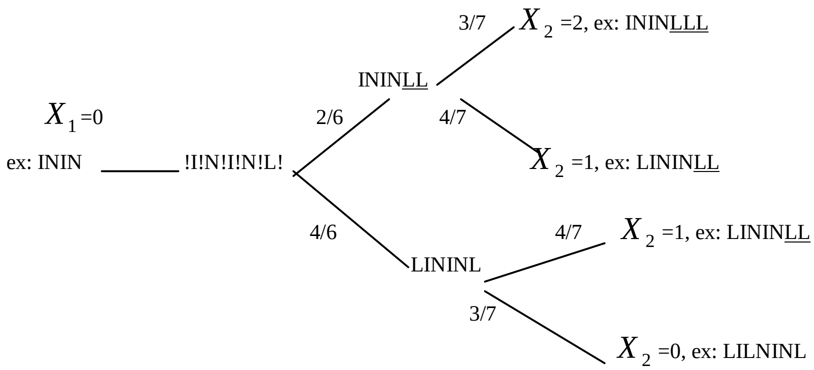
$$P(X_3=0|X_2=1)=\frac{1}{8}+\frac{7}{8}\left(\frac{1}{9}+\frac{8}{9}\times\frac{1}{10}+\frac{8}{9}\times\frac{9}{10}\times\frac{1}{11}\right)=\frac{2880}{7920}$$

(see Figure 2 - 4).

Conclusion

Thus, the solution to the Presidents problem, the probability that a random permutation of the word BILLCLINTON has no equal neighbors, is 39/110.

Figure 1: Tree diagram for calculation of the conditional probability $p(X_2 = i | X_1 = 0)$.



Therefore, $P(X_2 = 2 | X_1 = 0) = \frac{2}{6} \times \frac{3}{7} = \frac{6}{42} = \frac{30}{210}$,

$P(X_2 = 1 | X_1 = 0) = \frac{2}{6} \times \frac{4}{7} + \frac{4}{6} \times \frac{4}{7} = \frac{24}{42} = \frac{120}{210}$, and $P(X_2 = 0 | X_1 = 0) = \frac{4}{6} \times \frac{3}{7} = \frac{12}{42} = \frac{60}{210}$.

Suppose that X_1=1, an example where this occurs is INNI. We obtain X_2=0 by separating the pair NN with the first L, the second L or the third L inserted, but not both or all three. It is convenient to use a tree diagram; see Figure 2. We obtain

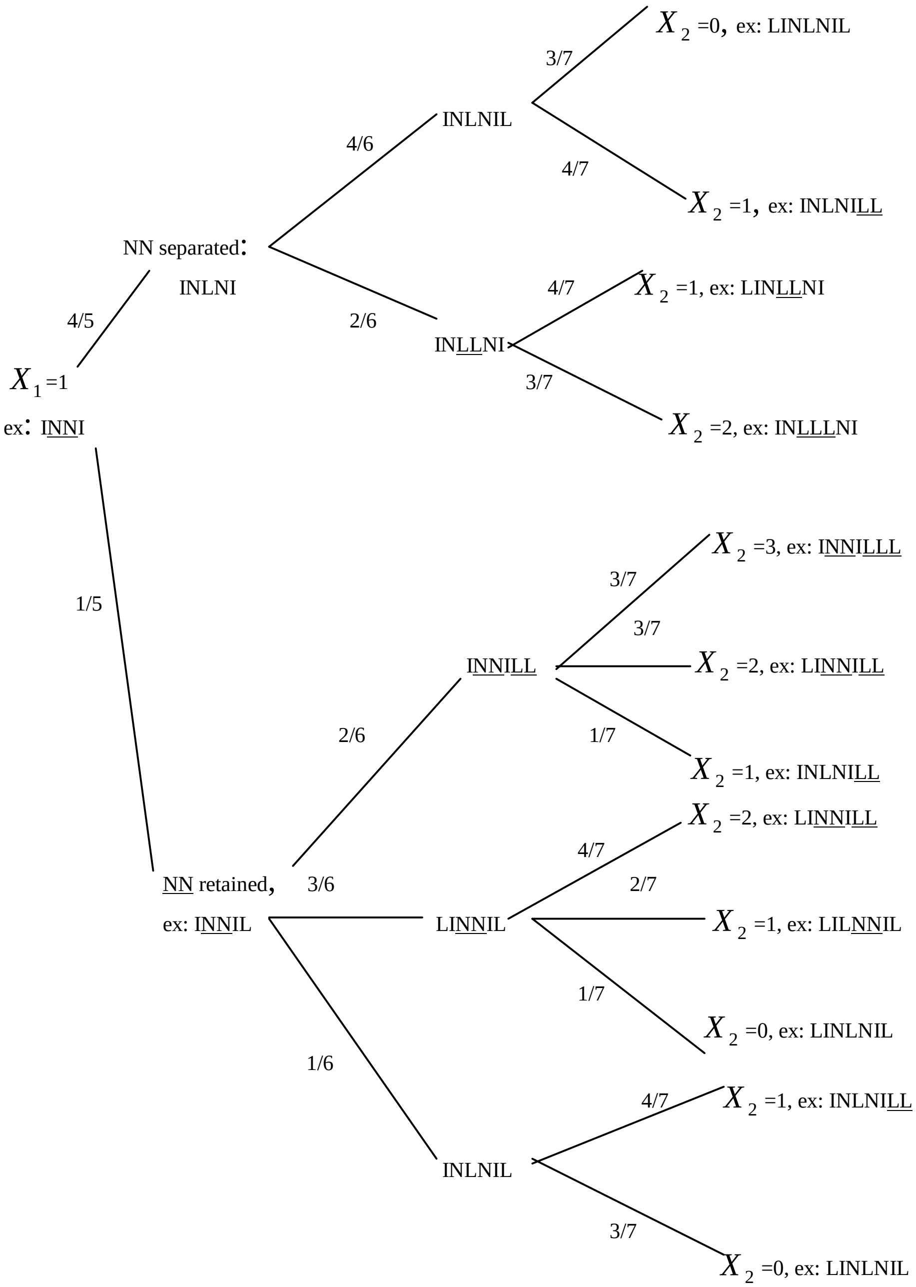
$$P(X_2 = 0 | X_1 = 1) = \frac{1}{5} \times \frac{4}{6} \times \frac{3}{7} + \frac{4}{5} \times \frac{3}{6} \times \frac{1}{7} + \frac{4}{5} \times \frac{1}{6} \times \frac{3}{7} = \frac{36}{210},$$

$$P(X_2 = 1 | X_1 = 1) = \frac{1}{5} \times \frac{4}{6} \times \frac{4}{7} + \frac{1}{5} \times \frac{2}{6} \times \frac{4}{7} + \frac{4}{5} \times \frac{2}{6} \times \frac{1}{7} + \frac{4}{5} \times \frac{3}{6} \times \frac{2}{7} + \frac{4}{5} \times \frac{1}{6} \times \frac{4}{7} = \frac{72}{210},$$

$$P(X_2 = 2 | X_1 = 1) = \frac{1}{5} \times \frac{2}{6} \times \frac{3}{7} + \frac{4}{5} \times \frac{2}{6} \times \frac{3}{7} + \frac{4}{5} \times \frac{3}{6} \times \frac{4}{7} = \frac{78}{210}, \text{ and}$$

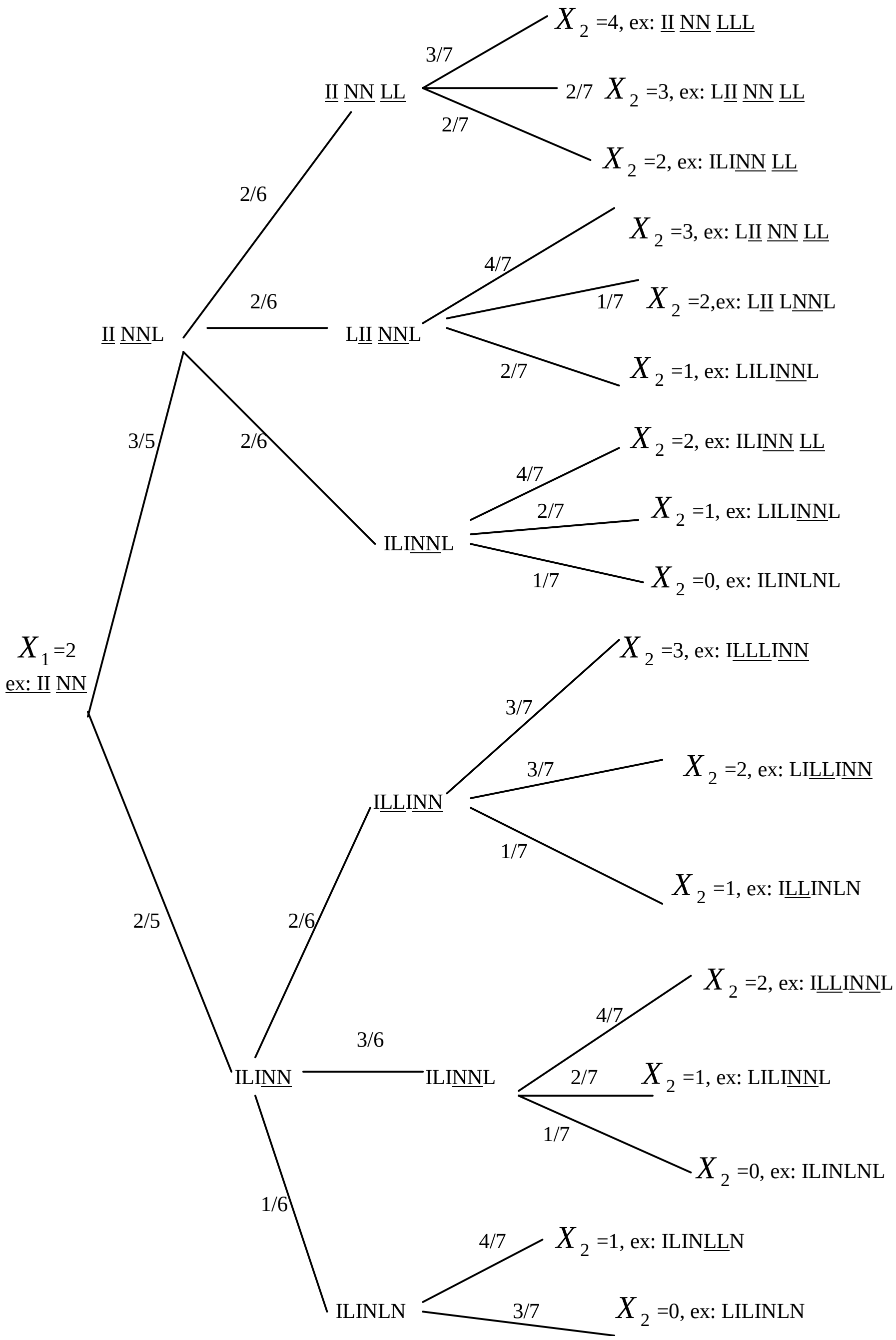
$$P(X_2 = 3 | X_1 = 1) = \frac{4}{5} \times \frac{2}{6} \times \frac{3}{7} = \frac{24}{210}.$$

Figure 2: Tree diagram for calculation of the conditional probability $P(X_2 = i | X_1 = 1)$.



Suppose that X_1=2, an example is II NN. Refer to Figure 3.

Figure 3: Tree diagram for calculation of the conditional probability $P(X_2 = i | X_1 = 2)$



$$P(X_2=0|X_1=2)=\frac{3}{5}\times\frac{2}{6}\times\frac{1}{7}+\frac{2}{5}\times\frac{3}{6}\times\frac{1}{7}+\frac{2}{5}\times\frac{1}{6}\times\frac{3}{7}=\frac{18}{210},$$

$$P(X_2=1|X_1=2)=\frac{3}{5}\times\frac{2}{6}\times\frac{2}{7}+\frac{3}{5}\times\frac{2}{6}\times\frac{2}{7}+\frac{2}{5}\times\frac{2}{6}\times\frac{1}{7}+\frac{2}{5}\times\frac{3}{6}\times\frac{2}{7}+\frac{2}{5}\times\frac{1}{6}\times\frac{4}{7}=\frac{48}{210},$$

$$P(X_2=2|X_1=2)=\frac{3}{5}\left[\frac{2}{6}\left(\frac{2}{7}+\frac{1}{7}+\frac{4}{7}\right)\right]+\frac{2}{5}\left(\frac{2}{6}\times\frac{3}{7}+\frac{3}{6}\times\frac{4}{7}\right)=\frac{78}{210},$$

-6-

$$P(X_2=3|X_1=2)=\frac{3}{5}\left(\frac{2}{6}\times\frac{2}{7}+\frac{2}{6}\times\frac{4}{7}\right)+\frac{2}{5}\times\frac{2}{6}\times\frac{3}{7}=\frac{48}{210},\text{ and } P(X_2=4|X_1=2)=\frac{3}{5}\times\frac{2}{6}\times\frac{3}{7}=\frac{18}{210}.$$

Based on (1.3) and combining Table 3 and Table 4, we obtain

$$P(X_2=0)=\sum_{i=0}^{2}P(X_2=0|X_1=i)\ P(X_1=i)=\frac{2}{6}\times\left(\frac{60}{210}+\frac{36}{210}+\frac{18}{210}\right)=\frac{228}{1260}.$$ Similarly,

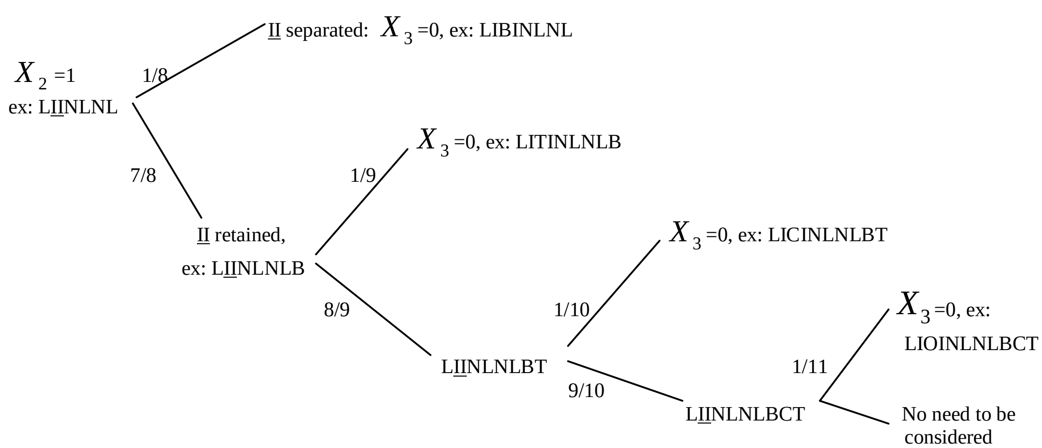
$$P(X_2=1)=\frac{2}{6}\times\left(\frac{120+72+48}{210}\right)=\frac{480}{1260},\ P(X_2=2)=\frac{2}{6}\times\left(\frac{30+78+78}{210}\right)=\frac{372}{1260},$$

$$P(X_2=3)=\frac{2}{6}\times\left(\frac{24+48}{210}\right)=\frac{144}{1260},\text{ and } P(X_2=4)=\frac{2}{6}\times\frac{18}{210}=\frac{36}{1260}.$$

Table 5: The Probability Function of X_2

j	0	1	2	3	4
$P(X_2=j)$	228/1260	480/1260	372/1260	144/1260	36/1260

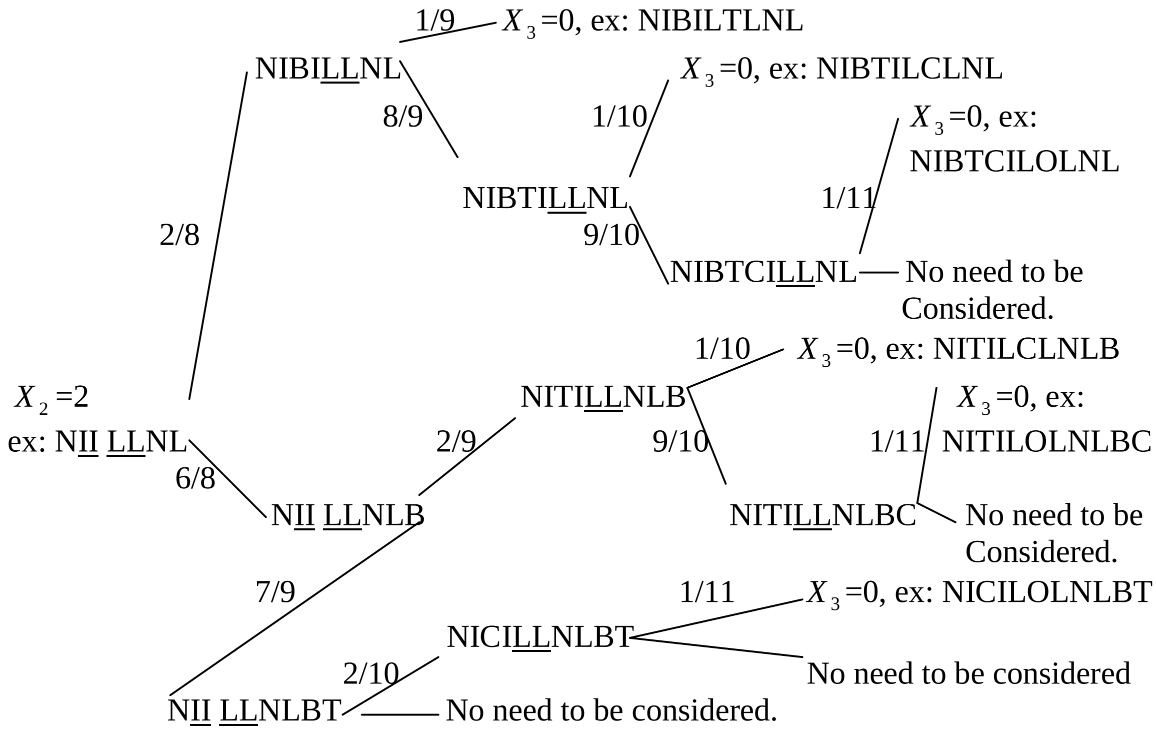
Figure 4: Tree diagram for calculation of the conditional probability $P(X_3=0|X_2=1)$.



When X_2 =2, for example, NII LLNL, $P(X_3=0|X_2=2)=\frac{2}{8}\times\left(\frac{1}{9}+\frac{8}{9}\times\frac{1}{10}+\frac{8}{9}\times\frac{9}{10}\times\frac{1}{11}\right)$

$$+\frac{6}{8}\times\left(\frac{2}{9}\times\frac{1}{10}+\frac{2}{9}\times\frac{9}{10}\times\frac{1}{11}\right)+\frac{6}{8}\times\frac{7}{9}\times\frac{2}{10}\times\frac{1}{11}=\frac{864}{7920}$$ (see Figure 5).

Figure 5: Tree diagram for calculation of the conditional probability $P(X_3 = 0 | X_2 = 2)$.

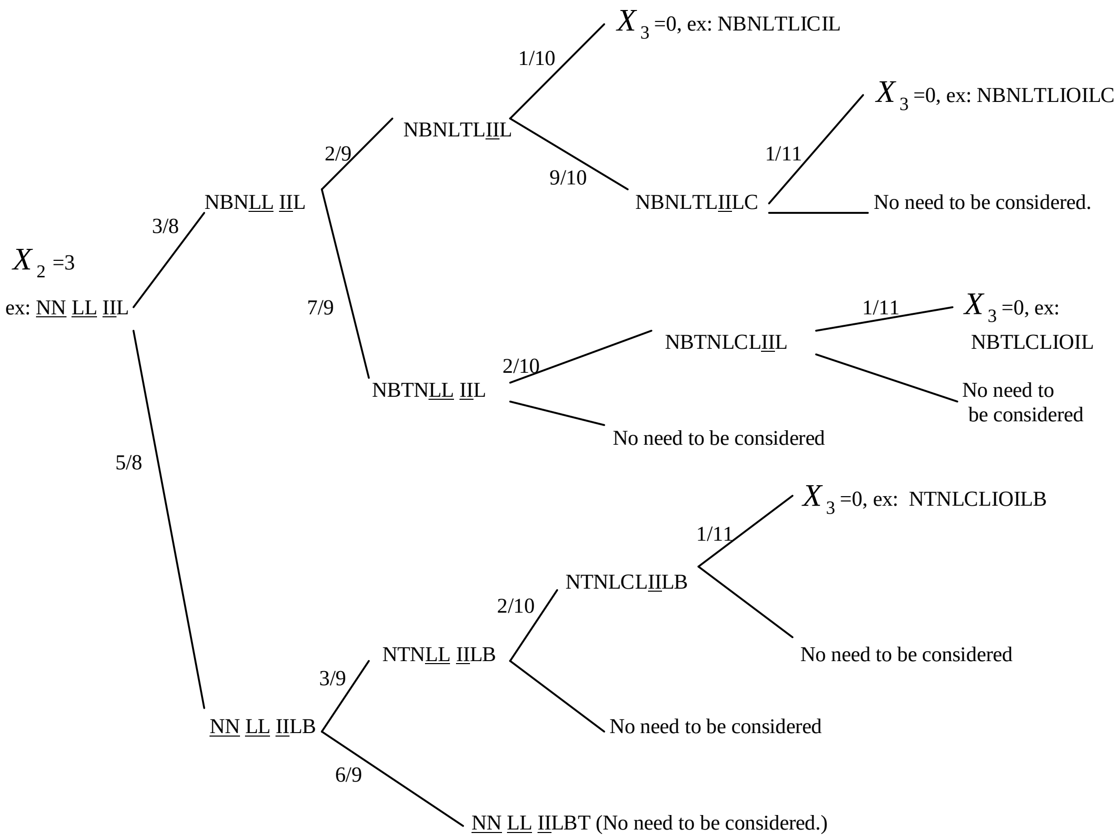


When X_2 =3, for example, NN LL IIL,

$$P(X_3 = 0 | X_2 = 3) = \frac{3}{8} \times \left(\frac{2}{9} \times \frac{1}{10} + \frac{2}{9} \times \frac{9}{10} \times \frac{1}{11} + \frac{7}{9} \times \frac{2}{10} \times \frac{1}{11} \right) + \frac{5}{8} \times \frac{3}{9} \times \frac{2}{10} \times \frac{1}{11} = \frac{192}{7920}$$

(see Figure 6).

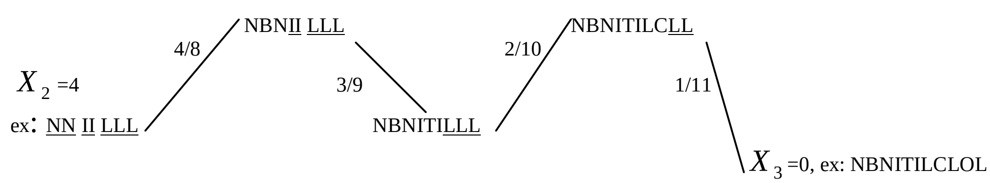
Figure 6: Tree diagram for calculation of the conditional probability $P(X_3 = 0 | X_2 = 3)$.



When X_2 =4, for example, NN II LLL, $P(X_3 = 0 | X_2 = 4) = \frac{4}{8} \times \frac{3}{9} \times \frac{2}{10} \times \frac{1}{11} = \frac{24}{7920}$

(see Figure 7).

Figure 7: Tree diagram for calculation of the conditional probability $P(X_3 = 0 | X_2 = 4)$.



Therefore,

$$P(X_3 = 0) = P(X_3 = 0 | X_2 = 0) P(X_2 = 0) + P(X_3 = 0 | X_2 = 1) P(X_2 = 1) +$$

$$P(X_3 = 0 | X_2 = 2) P(X_2 = 2) + P(X_3 = 0 | X_2 = 3) P(X_2 = 3) + P(X_3 = 0 | X_2 = 4) P(X_2 = 4)$$

$$= 1 \times \frac{228}{1260} + \frac{2880}{7920} \times \frac{480}{1260} + \frac{864}{7920} \times \frac{372}{1260} + \frac{192}{7920} \times \frac{144}{1260} + \frac{24}{7920} \times \frac{36}{1260} = \frac{39}{110}.$$

This is the answer to the Bill Clinton (President) problem.

References

Blom, G., Englund, J., & Sandell, D. (1998). The Mississippi Problem. *The American Statistician*, *52*, 49-50.

Blom, G., Holst, L., & Sandell, D. (1994). *Problems and Snapshots from the World of Probability*, New York: Springer-Verlag.

Constantine, G. M. (1987). *Combinatorial Theory and Statistical Design*. New York: Willey.

David, F. N., & Barton, D. E. (1962). *Combinatorial Chance*, London: Griffin.

Mood, A. M. (1940). The Distribution Theory of Runs, *The Annals of Mathematical Statistics*, *11*, 367-392.