한국학교수학회논문집 제 5권, 제 1호
Journal of the Korean School Mathematics Society
Volume 5, Number 1, 147-156, June 2002

TWO CLASS ACTIVITIES OF M&M CANDIES

G. Daniel Kim[1] and Sung Sook Kim[2]

1. Introduction

Jim Libby(2000) wrote an article of how to calculate the probability of drawing three M&Ms of a same color without replacement out of a given M&M Milk Chocolate Candies bag. Unlike the author assumed (also as the author suspected) in the article, the six different colors of M&Ms (Brown, Yellow, Red, Orange, Green, Blue) are not uniformly distributed, but are supposedly distributed according to the following probability distribution (Source: M&M web site at http://www.m-ms.com/cai/mms/faq.html).

Table 1. Probability Distribution of the six M&M colors

Color	Brown ($i=1$)	Yellow ($i=2$)	Red ($i=3$)	Orange ($i=4$)	Green ($i=5$)	Blue ($i=6$)
p_i	.30	.20	.20	.10	.10	.10

In subsequent, each color will be indexed by i according to Table 1. Calculating the probability of drawing three M&Ms of a same color according to the probability distribution takes extra considerations than what Libby(2000) did. In this article, the question considered by Libby(2000) is revisited, and extended to a more general question based on Table 1. Two classroom activities along with a TI calculator codes are introduced.

2. Notations and Developments

Despite M&M manufacturer aims at the six M& M colors distributed according to Table 1, due to its manufacturing processes there exists substantial deviation from their goals. In fact, significant deviation seem to exist not only in the proportions of the six colors in Table 1, but also in the total number of M&M candies contained

[1] Department of Mathematics Southern Oregon University Ashland, OR 97520, U. S. A.
[2] Department of Applied Mathematics Paichai University Doma-2-dong, Seo-gu, Daejeon 302-735, Korea

in each bag. Although it shouldn't be too difficult to investigate the deviations in more details, there seems to be no existing official reports on this issue yet. The presence of those deviations leads us to define the following random variables: For each $i = 1, \cdots, 6$,

M_i = the number of the ith color M&Ms contained in a package, and
N = the total number of M&M candies contained.

Then, $N = \sum_{i=1}^{6} M_i$, and $E(M_i|N) = p_i N$, $i = 1, 2, \cdots, 6$, where p_i is the probability distribution of the ith color. To begin with, we suppose that a M&M bag is randomly given, and draw n candies out of it, where n is assumed to be not so large number that satisfies $n \leq \min(M_1, M_2, \cdots, M_6)$. For convenience, let us denote the event of drawing n candies that results in a same color by A. We also define, for $i = 1, \cdots, 6$, Y_i to be the number of the ith color contained in the outcome of drawing n M&Ms. Then, notice that if $i \neq j$, $\{Y_i = n\}$ and $\{Y_j = n\}$ are mutually exclusive events. Therefore, the conditional probability of observing the event A given $M_1, M_2, \cdots, M_6$ is

$$\begin{aligned} P(A|N, M_1, M_2, \cdots, M_6) &= P(Y_1 = n, Y_2 = n, \cdots, \text{or } Y_6 = n) \\ &= P(Y_1 = n) + P(Y_2 = n) + \cdots + P(Y_6 = n) \\ &= \sum_{i=1}^{6} P(Y_i = n) \\ &= \sum_{i=1}^{6} \frac{\binom{M_i}{n}}{\binom{N}{n}} \qquad (1) \end{aligned}$$

If $n = 3$, it follows from the facts $E(M_i) = p_i N$ and $N = \sum_{i=1}^{6} M_i$ that

$$\begin{aligned} P(A|M_1, M_2, \cdots, M_6) &= \sum_{i=1}^{6} \frac{M_i(M_i-1)(M_i-2)}{N(N-1)(N-2)} \\ &= \frac{\sum_{i=1}^{6} M_i^3 - 3\sum_{i=1}^{6} M_i^2 + 2\sum_{i=1}^{6} M_i}{(\sum_{i=1}^{6} M_i)^3 - 3(\sum_{i=1}^{6} M_i)^2 + 2\sum_{i=1}^{6} M_i} \qquad (2) \end{aligned}$$

This formula holds a little hurdle because it requires prior knowledge of $M_1, M_2, \cdots, M_6$ which are not constants, but random variables due to variation in packaging process. If our goal is to find the exact probability of the event A without prior counting of $M_1, M_2, \cdots, M_6$ of a given package, it would be unfortunately a formidable task because we will have to deal with not only the multivariate probability distribution of $(M_1, M_2, \cdots, M_6)$ (remember $M_1, M_2, \cdots, M_6$ are not independent random variables), but also the covariance of $(M_1, M_2, \cdots, M_6)$. This quickly becomes beyond the scope of high school or undergraduate introductory probability or statistics course. To make the problem more tractable without prior counting of

$M_1, M_2, \cdots, M_6$, let us attempt to estimate the formula of (2) based on the equation $E(M_i|N) = p_iN$. For this end, let $\hat{M}_i = p_iN$. Then, the numerator and the denominator of (2) can be estimated by substituting $\hat{M}_i$ for M_i as follows:

$$\sum_{i=1}^{6} \hat{M}_i^3 - 3\sum_{i=1}^{6} \hat{M}_i^2 + 2\sum_{i=1}^{6} \hat{M}_i = .046N^3 - .6N^2 + 2N \tag{3}$$

and

$$\left(\sum_{i=1}^{6} \hat{M}_i\right)^3 - 3\left(\sum_{i=1}^{6} \hat{M}_i\right)^2 + 2\sum_{i=1}^{6} \hat{M}_i = N^3 - 3N^2 + 2N \tag{4}$$

Then, (1)-(4) yield

$$P(A|M_1, M_2, \cdots, M_6) \approx P(A|\hat{M}_1, \hat{M}_2, \cdots, \hat{M}_6) = P(A|N) = \frac{.046N^2 - .6N + 2}{N^2 - 3N + 2} \tag{5}$$

Figure 1 is the graph of (5) as a function of N. Again, if our goal is to figure out the probability of the event A without counting of the number of candies, the formula of (5) would not answer the question because it requires prior information of N, the total number candies contained in the bag. Perhaps, N may be estimated by dividing the total weight of the M&M bag by the weight of a single M&M candy. This problem may be accommodated into a classroom activity as presented in Activity I.

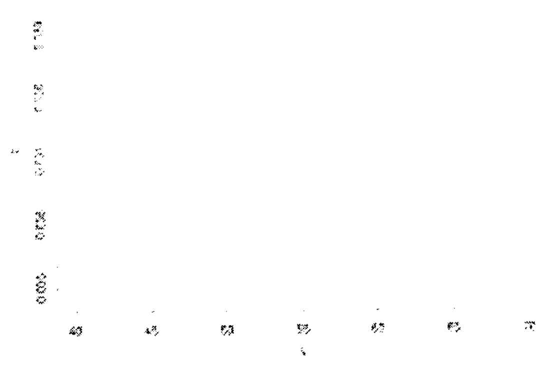


Figure 1

3. Two Activities

3.1. Activity I - Calculations and Comparisons

Activities that involve M&M candies seems to create always a joyful excitement in classrooms. In this activity, a simple activity is introduced to draw student's attention to the theories that were built so far. Before any activity begins, if possible, it should be pointed out again that the exact formula of (2) and the approximate formula of (5) were bridged by substituting $\hat{M}_i$ for M_i where $\hat{M}_i$ is an unbiased estimate of M_i given N. Perhaps, students may be curious about the validity and the reliability of the bridging (or substituting) process. The current activity consists of three parts. In part I, a strategy to estimate the approximate probability of (5) without any prior counting of candies is introduced. In part II and III, the exact probability of (2) and the approximate probability of (5) with prior counting of candies are considered, respectively. In this activity, it is assumed that students are grouped into two, and each group is given with a plain M&M milk candy bag.

Part 1: Use of formula (5) - an approximate probability without counting

(a) Grab a few M&M candies (for instance five or so), weigh each candy on a scale, average them out, and call this value w as an estimate of single M&M candy's weight. (If there was any known value of this, it may be used instead, and this step may be skipped)

(b) Calculate an estimate $\hat{N}$ of N by $\hat{N} = \frac{W}{w}$ where W, the weight of the M&M bag, is given in the back of M&M bag.

(c) Calculate the formula in (5) by substituting $\hat{N}$ for N.

Part 2: Use of formula (2)- the exact probability

(a) Open an M&M bag, and count $M_1, M_2, \cdots, M_6$.

Color	Brown ($i=1$)	Yellow ($i=2$)	Red ($i=3$)	Orange ($i=4$)	Green ($i=5$)	Blue ($i=6$)	Total
M_i							

(b) Calculate the exact probability of (2) based on $M_1, M_2, \cdots, M_6$. (With calculators, it should be done without much difficulty)

Part 3: Use of formula (5)- an approximate probability with counting

(a) Calculate N by adding up $M_1, M_2, \cdots, M_6$.

(b) Calculate the formula in (5) with the N.

Discussion:

(a) Within your group, discuss what you have learned through the three steps of activity.

(b) Share your thoughts with other group. (This can be lead by teacher)

3.2. Activity II - Simulations and Comparisons

Good use of advanced technologies such as calculators and computers in classrooms can be a great benefit to students because it allows students to experience great deal of powerful message in mathematical science. Trend shows vast majority of high school and college students in USA are equipped with programmable graphing calculators such as TI calculators. In second part of this Activity II, it is assumed that students can assess either programmable calculators or statistical software in computer. This Activity II consists of two parts. In first part, each group conducts a simple hands on experiment to simulate the probability of the event A. In second part, each group will be asked to import a program into their calculators or computer, where the program can be supplied perhaps by their teacher somehow. A sample TI-83 program for the Part II is written and stored at http://www.sou.edu/math/faculty/kim/Bmm.8xp. This file is in .8xp file format(TI-83 file format), and TI-83 Graph Link software is needed to open the file and transfer it into your own TI-83 calculator. The program should fit most of other types of TI calculators such as TI-89, and 92+ as well with slight modification. Appendix 2 contains a few selective TI outputs of the program. In Figure 2 the program asks two initial conditions: (a) how many objects (in this article it would be the number of M&M candies contained in a package), and (b) total number of classes (it would be six for the six colors of a plain M&M package). In Figure 3 (which is a short cut) the program asks the probability distribution of the six classes. Figure 4 displays three things: (a) B = a simulated M&M bag where the integer i $(= 1, 2, \cdots, 6)$ stands for ith color which we reserved earlier. (b) Q = the number of M&Ms for each color contained in the bag, and (c) S = the proportion of each color. Figure 5 asks two questions: (a) Number of choose? (how many M&Ms you want to draw at once), (b) Number of sampling? (how many times you want to repeating the sampling). Finally, Figure 6 displays two quantities: (a) the exact solution based on the formula (2), (b) the proportion of same color drawn out of the repeated number of sampling.

Part 1: Hands on simulation

One student(Experimenter) in each group draws $n = 3$ candies out of the given M&M bag, and another person(Recorder) in the group records the outcome of the trial in Table 3. Repeat this trial total 20 times, switch the role of being experimenter and recorder, and repeat this trial another 20 times. Based on Table 3, calculate the percent of trials that result in a same color.

Trial Number	**Drew three M&Ms of a same color?**	
	Yes	No
1		
2		
⋮	⋮	⋮
40		
Total		

Part 2: Programmable calculator or computer software based simulation

The following is an outline of the programmable algorithm that simulates the probability of drawing three M&M candies.

a) Simulate a M&M candy bag of size N according to the probability distribution of the six colors in Table 1, where N is the total number of candies of the M&M bag each group has.

b) Simulate drawing three candies out of the simulated M&M bag in the previous step, and repeat this trial 40 times.

c) Calculates the percent of trials that result in a same color.

Discussion: Discuss what they learned from Part I and Part II in this activity. Also compare the findings from Part I and Part II to the findings from Activity I which are algebraical solutions. Since this discussion is the last part of the M&M activity, perhaps students can start eating as they share their thoughts and fun.

4. Summary

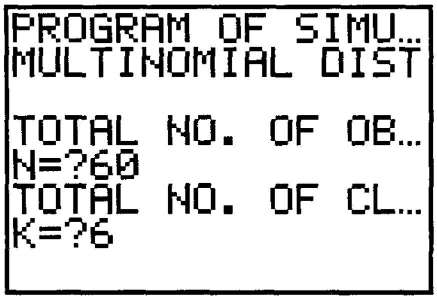
In this article we studied the probability of drawing n candies of a same color out of a M&M bag. For this, an exact formula was developed, and an approximate formula was considered for $n = 3$. In Activity I, how to estimate some unknown parameters involved in those two formulas were studied, and applied to calculate both the exact formula and the approximate formula. In Activity II, two different simulations were introduced to enhance the quality of student's understanding. Through the development of the theoretical formulas that was presented in this article and the activity I, students will be experiencing or re visiting some fundamental probability and statistical concepts such as conditional probability, conditional expectation, unbiased estimate. The two simulation works in Activity II will allow students to experience not only the power of mathematical theory, but also a great fun in mathematics classroom. The two activities we considered in section 3 are just ideas of how the activities can be conducted in a classroom. Teachers who wish to adopt the idea of the Activity I and II will have to design careful activity sheets to fit their students' needs and classroom environments.

Appendices

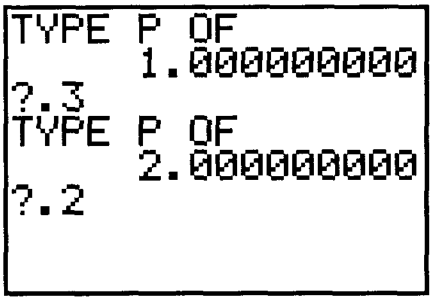
Appendix 1: TI codes for Activity II (Part 2)

```
:ClrHome :ClrList ⌊P, ⌊U, ⌊B, ⌊T, ⌊Q
:Disp “PROGRAM OF SIMULATING” :Disp “MULTINOMIAL DIST” :Disp “ ”
:Disp “TOTAL NO. OF OBJECTS” :Prompt N
:Disp “TOTAL NO. OF CLASSES” :Prompt K
:For(I,1,K)
:Disp “TYPE P OF ”,I :Input X :X→ ⌊P(I)
:End
:If sum(⌊P)≠ 1
:Then :ClrHome :Disp “ERROR...”:Disp “ ” :Disp “SUM OF P”:Disp “MUST BE 1” :Stop :End
:⌊P(1)→ ⌊T(1)
:For(I,1,K-1) :⌊P(I+1)+⌊T(I) → ⌊T(I+1) :End
:rand(N)→ ⌊U
:For(I,1,N) :1→ J
:While ⌊U(I) ≥ ⌊T(J) :J+1 → J :End
:J → ⌊B(I) :End
:Disp ⌊B :Disp “LOOK ⌊B FOR MM BAG” :Disp “ ”
:For(I,1,K) :0→ ⌊Q(I) :End
:For(I,1,N) :1→ J
:While ⌊B(I)≠ J :J+1 → J :End
:1+⌊Q(J)→ ⌊Q(J) :End
:For(I,1,K):⌊Q(I)/N → ⌊S(I) :END
:Disp “LOOK ⌊Q and ⌊S FOR ” :Disp “SUMARY OF BAG” :Disp “ ”
:Disp “NO. OF CHOOSE?” :Prompt C :Disp “NO. OF SAMPLING?” :Prompt R
:0 → X
:For(I,1,R)
:randInt(1,N,C)→ ⌊C :0→ Y
:For(J,2,C)
:If ⌊B(⌊C(1))=⌊B(⌊C(J)) :Y+1→ Y :End
:If Y=C-1 :X+1→ X :End
:Disp “P. OF SAME COLOR” :Disp X/R :Disp “ ”
:Disp “EXACT SOLUTION” :sum(⌊Q)→ S
:Disp (sum(⌊Q∧3)-3sum(⌊Q∧2)+2S)/(S∧3-3S∧2+2S)
```

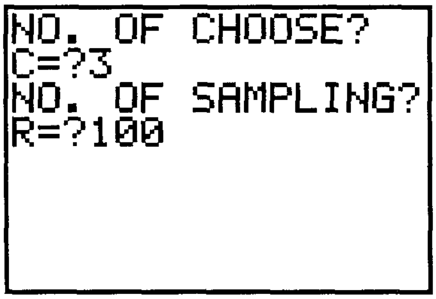
Appendix 2: A few selective outputs of TI codes in Appendix 1



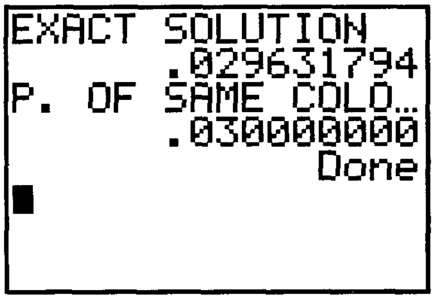
(a) Figure 2



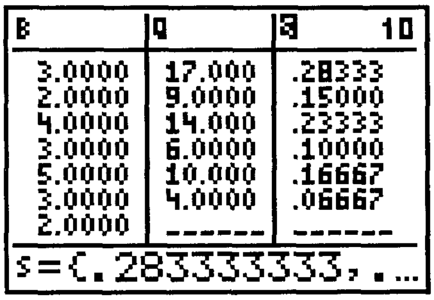
(b) Figure 3



(c) Figure 4



(d) Figure 5



(e) Figure 6

Reference

Libby, J (2000), "M&M Probability", *TOMT: The Oregon Mathematics Teacher*

Wackerly, D., Mendenhall, W., Scheaffer, R., "Mathematical Statistics with Applications," Duxburry Press, 5th edition, 1996

M&M 쵸코렛을 이용한 교실에서의 통계활동

G. Daniel Kim[1] · 김 성 숙[2]

요 약

Jim Libby(2000)는 임의의 M&M 밀크 쵸코렛 봉지에서 꺼낸 것을 다시 집어넣지 않고 3개의 쵸코렛을 꺼낼 때 같은 색이 나올 확률을 계산하는 방법의 논문을 썼다. Libby는 그의 논문에서 갈색, 노랑, 빨강, 주황, 초록, 파랑의 6개 색의 M&M 밀크 쵸코렛들이 똑같은 비율로 분포되었다고 가정하였다. 그러나 실제로 M&M 쵸코렛의 여섯 가지 색깔의 확률분포는 똑같은 비율이 아니다. M&M회사는 홈페이지(http://www.m-ms.com/cai/mms/faq.html)를 통해 실제적인 6개 색의 M&M 밀크 쵸코렛들의 분포는 갈색이 30%, 노랑과 빨간색이 각각 20% 주황, 초록과 파랑은 각각 10%의 분포라고 밝히고 있다.

이 논문에서 우리는 Libby가 생각하였던 문제를 실제적인 6개 색의 M&M 밀크 쵸코렛들의 확률 분표에 의거하여 다시 생각해보며, 또한 3개의 쵸코렛대신 n개의 쵸코렛을 꺼낸다고 가정하여 더욱 일반적인 결론을 유도한다. 또한 유도한 이 정확한 확률 공식과 근사 공식을 활동을 통해 점검하고 학생들이 주도적으로 지금까지 배워온 이론들을 점검할 수 있게 하였다. 활동을 시작하기 전에 정확한 확률 공식과 근사 공식과의 관계를 설명하고 기본적인 확률과 통계의 개념을 다시 정립할 수 있도록 하였다. Piaget가 '지식이란 학습자에 의해 능동적으로 구성되는 것이지 환경으로부터 수동적으로 받아들이는 것은 아니다'라고 했듯이, 활동을 통한 학습은 학생들을 능동적으로 만들기 때문에 학생들이 지식을 구성해 갈 수 있다. 활동을 간단히 소개하면 다음과 같다.

활동I에서는 초코렛을 세어서 근사 확률을 추정하는 방법이 소개된다. 어떻게 매개변수가 두 공식에 관련이 되는지를 측정하고 두 공식을 사용하여 정확한 확률과 근사 확률을 계산하여 비교해본다. 각 조원들과 이 세 과정에서 무엇을 배웠나 토론하고 다른 조들과 배운 것을 나눈다.

활동II는 두 과정으로 나누어진다. 첫 번 과정은 각 그룹의 한 학생이 주어진 쵸코렛 봉지에서 3개의 쵸코렛을 꺼낸다. 다른 학생은 표 3에 나온 결과를 기록한다. 계속하여 20번씩 한다. 다시 학생을 바꾸어 20번 계속한다. 같은 색깔의 쵸코렛이 나온 확률은 계산하기 위한 간단한 실험이고 두 번째 과정은 각 조가 웹사이트나 선생님으로부터 제공받은 프로그램을 다운로드 받아 하는 시뮬레이션이다. 이 실험 후에 학생들이 이 두 활동을 통해 무엇을 배웠는지 토론해보고 또 두 활동을 비교해 볼 수 있다. 마지막으로 M&M 쵸코렛을 먹는 것으로 활동을 마칠 수 있을 것이다. 활동 II에 나오는 두 시뮬레이션은 학생들이 수학 이론의 힘을 깨달을 뿐 아니라 수학 교실에서 큰 재미를 느끼게 될 것이다. 이 논문에서 그래픽 계산기로 할 수 있는 프로그램을 소개하였다.

1) Southern Oregon University
2) 배재대학교