NEXUS Mathematicæ

ISSN: 2595-9824
v. 2, 2019, e19002.

A Generalization of the Car Versus Billy-Goat Problem

Uma Generalização do Problema do Carro Versus Bode

Jesus Carlos da Mota*

Abstract: This paper is a generalization of the famous and counter intuitive problem of car versus billy-goat, known as the Monty Hall problem. With the hypothesis that the number of cars plus the number of billy-goats is equal to the number of doors, and the TV host may open more than one door, we prove that the best strategy is always to switch the door. To obtain the analytical results we use only simple calculations of probability. We also did some numerical simulations that illustrate the obtained results. The programming code is included in the paper.

Keywords: Monty Hall problem. Several cars and several goats. Probability to win a car.

Resumo: Este artigo é uma generalização do famoso e contra intuitivo problema do carro versus bode, conhecido como o problema de Monty Hall. Considerando que o número de carros mais o número de bodes é igual ao número de portas, e que o apresentador de TV pode abrir mais de uma porta, provamos que a melhor estratégia é sempre trocar de porta. Para obter os resultados analíticos, usamos apenas cálculos simples de probabilidades. Fizemos ainda algumas simulações numéricas que ilustram os resultados obtidos. O código de programação está incluido no artigo.

Palavras-chave: Problema de Monty Hall. Vários carros e vários bodes. Probabilidade de ganhar um carro.

1 Introduction

The problem of the car versus the billy-goat, known as the Monty Hall problem, may be familiar to many readers. It is a probability problem which, at first, looks really

*Instituto de Matemática e Estatística, Universidade Federal de Goiás, jesusdamota@gmail.com

easy, but it tricks the intuition of most people. Because of that, it is an interesting problem to be shared in a coffee house with a group of friends.

The origin of the problem is a US television contest called Let's Make a Deal, aired in the 1970s, and for a couple times it was discussed in online text [1], in math journals [2, 5, 6, 8], in books [7], and even in psychology journals, see for exemple [3, 4].

The original problem is: on a TV show one candidate from the audience tries to win a car. The host shows him three doors to be chosen only one. One door hides the car and the two others hide the billy-goats. The candidate in order to win the car must pick the door which hides the car, if not, he wins a billy-goat. After some entertainment for the audience, common thing for this kind of show, the host asks the candidate to choose one door, without opening it. The host knows what each door hides and where is the car. Chosen the door, at least one of the other two doors is hiding a billy-goat. Continuing the game, the host opens one of the two remaining doors, it reveals one billy-goat. After that, he asks the candidate if he would like or would not like to choose the other door.

Throughout this text, sometimes, a door hiding a billy-goat is simply called billy-goat door. The question is, what is the best strategy for the candidate? By this question we mean, what has the greater probability to win the car, to keep the first door chosen, or to choose the other door? Here goes the answer already. If he does not switch to the other door, the chance of choosing the car is one in three and choosing a billy-goat is two in three. So, the probability of winning the car is $\frac{1}{3}$ and that of getting a billy-goat is $\frac{2}{3}$.

On the other hand, if the first choice is a billy-goat and the host opens the other billy-goat door, the candidate certainly will win the car by switching the door. Since the probability of choosing initially a billy-goat is $\frac{2}{3}$, then the probability of winning the car when switching doors is also $\frac{2}{3}$ and, as a consequence, the probability of getting a billy-goat is $\frac{1}{3}$. Therefor,e the best strategy to win the car is to switch doors.

An interesting question is: what happens if we have four doors or, more generally, if we have N doors and the TV host opened more than one door? What would be the best strategy, keep the first choice or change door? In this paper, we will discuss in detail this issue, including some numerical simulations.

The problem for N doors is formally described in the next section.

2 Several doors, several cars, and several billy-goats

In this section we discuss the problem in its various forms. The notation will be introduced throughout the discussion.

Let us suppose that the candidate, now, may choose between N doors, $N \geq 3$, of which N_c hide cars and N_g hide billy-goats, with

$$N_c + N_g = N, \quad \text{where} \quad 2 \leq N_g < N, \quad N \geq 3.$$

When $N = 3$ the inequalities above shows that $N_g = 2$, as a consequence $N_c = 1$, and the problem returns to its original form, just like in the introduction.

We observe that we can also include empty doors, doors that do not hide a car nor a billy-goat. But, what makes the problem interesting, is the certain idea that the candidate will win a car or a billy-goat. Therefore, we are not going to consider empty doors. The host asks the candidate to choose one door, without opening it, and the prize will be whatever the door hides. Of course we are going to assume the candidate wants to win the car. The host knows what each door hides. Continuing the show, the host opens a door, and reveals a billy-goat. After that, he shoots the question:

– Would you like to choose another door?

The candidate is not sure what to do, then the host opens another door, revealing another billy-goat, and shoots the same question:

– Would you like to choose another door?

With the candidate still facing the doubt, the hosts continues to open billy-goat doors, and asking the same question over and over. After opening a total of N_o doors, all of those revealing a billy-goat, the host warns the candidate:

– This is your last chance! Would you like to choose another door? Obviously the number o doors opened is smaller than the total number of doors that hide billy-goats,

$$N_o < N_g = N - N_c.$$

The problem to be discussed here is: what is the best strategy for the candidate in order to maximize his chance of winning the car? To Switch or not to switch? And, if switching, when should he do so? After how many doors opened by the host?

3 Discussion of the solution

Let P_c be the probability of the candidate winning the car, and P_g the probability of winning a billy-goat. Since we are not considering empty doors, we have

$$P_c + P_g = 1.$$

Let us discuss the solution considering three cases:

1. Only one door hides a car, the others hide billy-goats, and the host opens only one door.

2. Only one door hides a car, the others hide billy-goats, but the hosts opens one or more doors.

3. Some doors hide cars, the others hide billy-goats, and the host can open more than one door.

In the three cases, we calculate, independently, the probability of the candidate winning the car and the likelihood of him winning a billy-goat. That is, we calculate, independently, P_c and P_g. Obviously, the problem would be solved by calculating only P_c or P_g, because $P_c + P_g = 1$. We do the two calculations only to pair the results.

Case 1. In this case

$$N_c = 1, \quad N_g = N - 1 \quad \text{and} \quad N_o = 1.$$

Probability keeping the first choice:

If the candidate does not switch to another door the probability of winning the car is $\frac{1}{N}$ and that of winning a billy-goat is $\frac{N-1}{N}$, that is,

$$P_c = \frac{1}{N} \quad \text{and} \quad P_g = \frac{N-1}{N},$$

because, only one of N doors hides the car and the others $N-1$ hide billy-goats.

Probability switching doors:

The candidate will win the car only if two simultaneous facts occur: choosing as his first door a billy-goat door, that probability is $\frac{N-1}{N}$, and after the host opens another billy-goat door, choose between the other $N-2$ doors still unopened, exactly the one which hides the car, that probability is $\frac{1}{N-2}$. Therefore, the probability of winning the car by switching doors is

$$P_c = \Big(\frac{N-1}{N}\Big)\Big(\frac{1}{N-2}\Big).$$

The candidate will win a billy-goat if one of the following facts occur: the first door chosen hides the car, and by changing his door he will get a billy-goat, that probability is $\frac{1}{N}$, or the first door chosen hides a billy-goat, that probability is $\frac{N-1}{N}$ and, after the host opens another door revealing a billy-goat, choose between the $N-2$ doors remaining another billy-goat door, that probability is $\frac{N-3}{N-2}$. Therefore, the probability of winning the billy-goat is

$$P_g = \frac{1}{N} + \Big(\frac{N-1}{N}\Big)\Big(\frac{N-3}{N-2}\Big).$$

As expected, an easy calculation shows that

$$P_c + P_g = \Big(\frac{N-1}{N}\Big)\Big(\frac{1}{N-2}\Big) + \frac{1}{N} + \Big(\frac{N-1}{N}\Big)\Big(\frac{N-3}{N-2}\Big) = 1.$$

When $N = 3$, notice that $P_c = \frac{2}{3}$ and $P_g = \frac{1}{3}$, results that match the original problem results, shown in the introduction.

For any number of doors greater or equal to three, the probability of winning the car by switching doors is greater than the probability of winning the car without switching doors, because

$$\Big(\frac{N-1}{N}\Big)\Big(\frac{1}{N-2}\Big) > \frac{1}{N}, \text{ since } \frac{N-1}{N-2} > 1.$$

Therefore the better strategy for the candidate to win the car is to switch doors.

Case 2. In this case,

$$N_c = 1, \quad N_g = N-1 \quad \text{and} \quad 1 \leq N_o < N_g.$$

Probability keeping the first choice:

If the candidate does not switch to another door, the number of billy-goat doors opened by the host will not influence the results. Therefore, the probabilities are the same as in Case 1, that is,

$$P_c = \frac{1}{N} \quad \text{and} \quad P_g = \frac{N-1}{N}.$$

Probability switching doors:

The probabilities are calculated supposing that the candidate switches doors right after the host opens N_o billy-goat doors, where $1 \leq N_o < N_g = N - N_c$.

The candidate will win the car if two simultaneous facts occur: choosing at first a billy-goat door, which has probability $\frac{N-1}{N}$ and, after the host opens N_o billy-goat doors, choosing, among the $N - N_o - 1$ doors remaining, the door that hides the car which has a probability of $\frac{1}{N-N_o-1}$. Therefore, the probability of the candidate winning the car by switching doors is

$$P_c = \Big(\frac{N-1}{N}\Big)\Big(\frac{1}{N-N_o-1}\Big). \tag{1}$$

The candidate, in this case, will win a billy-goat if one of the following occurs: they initially choose the door which hides the car, with a probability of $\frac{1}{N}$ and, later, switch to another door which will, inevitably, have a billy-goat, or else they initially choose a billy-goat door, with probability $\frac{N-1}{N}$ and, after the host opens N_o doors that hide billy-goats, they choose, among the $N - N_o - 1$ doors remaining, one that also hides a billy-goat, which will have a probability of $\frac{N-N_o-2}{N-N_o-1}$. Therefore the probability of winning a billy-goat is

$$P_g = \frac{1}{N} + \Big(\frac{N-1}{N}\Big)\Big(\frac{N-N_o-2}{N-N_o-1}\Big). \tag{2}$$

As expected, an easy calculation shows that

$$P_c + P_g = \Big(\frac{N-1}{N}\Big)\Big(\frac{1}{N-N_o-1}\Big) + \frac{1}{N} + \Big(\frac{N-1}{N}\Big)\Big(\frac{N-N_o-2}{N-N_o-1}\Big) = 1.$$

Equations (1) and (2) show that when $N_o = 1$ the results here match those of Case 1.

Regardless of the number of doors opened by the host, the probability of winning the car by switching doors is always grater than without switching, because

$$\Big(\frac{N-1}{N}\Big)\Big(\frac{1}{N-N_o-1}\Big) > \frac{1}{N}, \text{ since } \frac{N-1}{N-N_o-1} > 1.$$

Even better, the best strategy for the candidate to win the car is to wait to change to another door until the host says, "this is your last chance", by that we mean, wait for the host to open as many doors as he can. This is because, by Equation (1), the probability of winning the car by switching doors increases with the number of doors opened.

Now we are going to discuss the more general case.

Case 3. In this case,

$$N_c \geq 1, \quad N_g = N - N_c \quad \text{e} \quad 1 \leq N_o < N_g.$$

Probabilities keeping the first choice:

If the candidate does not switch doors, since among the N doors there are N_c doors that hide cars and $N - N_c$ that hide billy-goats, the probability of winning a car is $\frac{N_c}{N}$ and that of winning a billy-goat is $\frac{N-N_c}{N}$. That is,

$$P_c = \frac{N_c}{N} \quad \text{and} \quad P_g = \frac{N - N_c}{N}. \tag{3}$$

Probability switching doors:

Like in Case 2, the probabilities are calculated supposing that the candidate switches doors right after the host opens N_o billy-goat doors, where $1 \leq N_o < N_g = N - N_c$.

The candidate will win the car if one of the following occur: they initially choose a billy-goat door, with probability $\frac{N-N_c}{N}$ and, after the host opens N_o billy-goat doors, they choose among the $N - N_o - 1$ doors remaining, a door that hides a car, which will have a probability of $\frac{N_c}{N-N_o-1}$, or else they initially choose a door that hides a car, with a probability of $\frac{N_c}{N}$ and, after N_o billy-goat doors are revealed by the host, they choose among the $N - N_o - 1$ remaining doors, one that hides another car, which will have a probability of $\frac{N_c-1}{N-N_o-1}$.

Therefore the probability of winning the car by switching doors is

$$P_c = \Big(\frac{N - N_c}{N}\Big)\Big(\frac{N_c}{N - N_o - 1}\Big) + \frac{N_c}{N}\Big(\frac{N_c - 1}{N - N_o - 1}\Big) = \frac{N_c(N-1)}{N(N - N_o - 1)}\,. \tag{4}$$

The candidate will win a billy-goat if one of the following occurs: initially they choose a billy-goat door, with probability $\frac{N-N_c}{N}$ and, after N_o doors opened, they choose, among the $N - N_o - 1$ doors remaining, another billy-goat door, which will have a probability of $\frac{N-N_c-N_o-1}{N-N_o-1}$, or else they initially choose a door hiding a car, with probability $\frac{N_c}{N}$ and, after N_o billy-goat doors are open, they choose among the $N - N_o - 1$ remaining doors one that hides a billy-goat, which will have a probability of $\frac{N-N_c-N_o}{N-N_o-1}$.

Therefore, the probability of winning a billy-goat by switching doors is

$$P_g = \Big(\frac{N - N_c}{N}\Big)\Big(\frac{N - N_c - N_o - 1}{N - N_o - 1}\Big) + \frac{N_c}{N}\Big(\frac{N - N_c - N_o}{N - N_o - 1}\Big)$$

$$= \frac{N(N - N_o - 1) - N_c(N-1)}{N(N - N_o - 1)} \,. \tag{5}$$

Again, as expected, an easy calculation shows that

$$P_c + P_g = \frac{N_c(N-1)}{N(N - N_o - 1)} + \frac{N(N - N_o - 1) - N_c(N-1)}{N(N - N_o - 1)} = 1 \,.$$

Equations (4) and (5) show that when $N_c = 1$ the results here match those of Case 2 and, when $N_o = 1$ and $N_c = 1$, they match the results of Case 1.

Obviously, by switching or not switching doors, if the number of doors hiding cars increases, the probability of winning a car also increases.

Just like in Case 2, the number of doors opened does not change the fact that the probability of winning a car is greater when the candidate switches doors, because,

$$\frac{N_c(N-1)}{N(N - N_o - 1)} > \frac{N_c}{N}, \text{ since } \frac{N-1}{N - N_o - 1} > 1.$$

Also like in Case 2, the best strategy for the candidate to win the car is to wait for the host to say, "this is your last chance", that is, after the host opens as many doors as the show allows him to open. This is because, by Equation (4), the probability of winning the car by switching doors increases with the number of doors opened before the switch.

Some numerical simulations are presented in the next section, showing the agreement with the theoretical exact solution.

4 Numerical simulation

Using the software MATLAB®, we developed an algorithm for numerical simulations of the problem in Case 3. Just out of curiosity, for those interested, we provide the code for the algorithm at the end of this text.

Ten rounds were simulated, each round simulates 1000 candidates trying to win the car. In this case, the number of simulations, denoted by m, is m = 1000.

Table 1 shows the results for the original problem given in the introduction, where the total number of doors is $N = 3$, the total number of cars is $N_c = 1$, the total number of billy-goats is $N_g = 2$ and the number of billy-goat doors opened by the host is $N_o = 1$.

Table 2 shows the simulation results when the total number of doors is $N = 8$, the number of cars is $N_c = 4$, the number of billy-goats is $N_g = 4$, and three values for the number of billy-goat doors opened by the host, $N_o = 1$, $N_o = 2$ and $N_o = 3$.

Table 1: Simulation for the following data: number of doors $N = 3$; number of cars $N_c = 1$; number of billy-goats $N_g = 2$; number of doors opened by the host $N_o = 1$. Columns NCnc, NGnc, NCyc, and NGyc, are respectively, the number of candidates who won a car by not switching doors, won a billy-goat by not switching doors, won a car by switching doors, won a billy-goat by switching doors. The first line of numbers corresponds to waited accurate values. The others ten lines of numbers correspond to the results of the ten rounds of simulation, each round with 1000 candidates.

NCnc	NGnc	NCyc	NGyc
333	667	667	333
354	646	646	354
339	661	661	339
339	661	661	339
362	638	638	362
342	658	658	342
299	701	701	299
373	627	627	373
350	650	650	350
327	673	673	327
335	665	665	335

In each one of the Tables 1 and 2, the column NCnc is the number of candidates who won a car by not switching doors; NGnc is the number of candidates who won a billy-goat by not switching doors; NCyc is the number of candidates who won a car by switching doors; and NGyc is the number of candidates who won a billy-goat by switching doors. The first line of numbers corresponds to the accurate expected values, calculated by multiplying the corresponding probability by the number of simulations m, and taking the integer number closer to that product. For example, the number of cars expected in 1000 simulations is equal to the integer number closer to $1000 \times P_c$, where for Case 3, P_c is given by (3) if the candidate does not switch doors, or given by (4) if the candidate switches doors. The last ten lines of numbers correspond to the results of the ten rounds simulated.

The programming code used for numerical simulations is shown in the next section.

Table 2: Simulation for the following data: number of doors $N = 8$; number of cars $N_c = 4$; number of billy-goats $N_g = 4$; and three values for the number of doors opened by the host $N_a = 1$, $N_a = 2$, and $N_a = 3$. Columns NCnc, NGnc, NCyc, and NGyc, are respectively, the number of candidates who won a car by not switching doors, won a billy-goat by not switching doors, won a car by switching doors, won a billy-goat by switching doors. The first line of numbers corresponds to waited accurate values. The others ten lines of numbers correspond to the results of the ten rounds of simulation, each round with 1000 candidates.

Na = 1				Na = 2				Na = 3			
NCnc	NGnc	NCyc	NGyc	NCnc	NGnc	NCyc	NGyc	NCnc	NGnc	NCyc	NGyc
500	500	583	417	500	500	700	300	500	500	875	125
501	499	595	405	492	508	676	324	489	511	865	135
501	499	572	428	488	512	709	291	487	513	867	133
498	502	572	428	501	499	681	319	488	512	882	118
499	501	543	457	487	513	694	306	502	498	873	127
478	522	572	428	504	496	701	299	516	484	881	119
492	508	577	423	489	511	709	291	492	508	888	112
515	485	583	417	503	497	690	310	511	489	873	127
505	495	559	441	492	508	708	292	493	507	892	108
501	499	582	418	524	476	711	289	489	511	875	125
521	479	580	420	484	516	677	323	497	503	888	112

4.1 Programming code

```
function [] = car_billy-goat(N, Nc, Ng, No, m)
% N  = number of doors,  m  = number of simulations
% Nc = number of cars, Ng = number of billy-goats
% No = number of goots door oppened by the host
% NCnc = number of candidates who won a car no changing doors
% NCyc = number of candidates who won a car changing doors
% NGnc = number of candidates who won a billy-goat no changing  doors
% NGyc = number of candidates who won a billy-goat changing doors
% NCnc_esp = number of cars expected no changing doors
% NCyc_esp = number of cars expected changing doors
% NGnc_esp = number of billy-goats expected no changing doors
% NGyc_esp = number of billy-goats expected changing doors
simulacoes=zeros(10,4);
if Nc+Ng~=N
  disp('Error, number of cars plus number of billy-goats is...
```

```
        different from the number of doors')
elseif No == Ng
  disp('No = Ng, changing doors the candidate always win.')
  disp('N_cars_with_change="m", N_billy-goats_with_change=0.')
elseif No > Ng
  disp('Error, No > Ng, choose No < Ng.')
else
  for i=1:10
    xc=Nc/N;  yc=(Nc-1)/(N-No-1);  yb=Nc/(N-No-1);
    PCst=Nc/N; PBst=(N-Nc)/N; PCct=Nc*(N-1)/(N*(N-No-1));
    PBct=(N*(N-No-1)-Nc*(N-1))/(N*(N-No-1));
    NCnc_esp = round(m*PCst);  NGnc_esp = round(m*PBst);
    NCyc_esp = round(m*PCct);  NGyc_esp = round(m*PBct);
    NCnc=0;  NGnc=0;  NCyc=0;  NGyc=0;
    for j=1:m
      x = rand(1,1);  y = rand(1,1);
      if x <= xc
        NCnc=NCnc+1;
        if y<=yc
          NCyc=NCyc+1;
        else
          NGyc = NGyc+1;
        end
      else
        NGnc = NGnc+1;
        if y<=yb
          NCyc=NCyc+1;
        else
          NGyc = NGyc+1;
        end
      end
    end
    simulation(i, :)=[NCnc NGnc NCyc NGyc];
  end
  sol_esperada=[NCnc_esp NGnc_esp NCyc_esp NGyc_esp]
  simulation
end
```

References

[1] ELLIS, K. M., **The Monty Hall problem**. Copyright: 1995. Available at: https://www.montyhallproblem.com. Accessed: 15 Nov. 2019.

[2] GILLMAN, L., The car and the goats, **American Mathematical Monthly**, v. 99, no. 1, p. 3-7, 1992.

[3] GILOVICH, T., MEDVEC, V. H., CHEN, S. Commission, omission, and dissonance reduction: coping with regret in the "Monty Hall" problem, **Personality and Social Physichology Bulletin**, v. 21, n. 2, p. 182-190, 1995.

[4] GRANBERG, D. Croos-cultural comparison of responses to the Monty Hall dilemma, **Social Behavior and Personality**, v. 27, n. 4, p. 431-448, 1999

[5] MORGADO, A. C. Os dois bodes (in Portuguese), **Revista do Professor de Matemática, SBM** v. 33, p. 26-29, 1997.

[6] PUZA, B. D., PITT, D. G. W., O'NEILL, T. J. The Monty Hall three doors problem, **Teaching Statistics**, v. 27, n. 1, p. 11-15, 2005.

[7] ROSENHOUSE, J. **The Monty Hall problem:** the remarkable story of math's most contentious brainteaser, Oxford University Press, Inc., New York, 2009.

[8] SALDANHA, N. C. Como perder amigos e enganar pessoas (in Portuguese), **Eureka, Olimpíada Brasileira de Matemática**, v. 1, p. 41-50, 1998.

Submetido em 05 ago. 2019.
Aceito em 15 nov. 2019.