The Monty Hall problem revisited

Christian Caamaño-Carrillo* and Manuel González-Navarrete†

Departamento de Estadística
Universidad del Bío-Bío.

Abstract

We propose a new approach to solve the classical Monty Hall problem in its general form. The solution is based on basic tools of probability theory, by defining three elementary events which decompose the sample space into a partition. The probabilities of each element of the partition allow us to compute the conditional and marginal probabilities of winning.

Keywords: Partition of a set, conditional probability, law of total probability, law of large numbers, central limit theorem.

1 Introduction

The Monty Hall problem appeared first in a letter by S. Selvin to the American Statistician [20, 21], it is a nice and controversial problem for introductory courses in probability, statistics, and game theory. The problem can be posed as follows:

You are playing a game on a TV show. There are three doors. One of them has a car in the backside, and the other ones have goats. You select one of the doors, say door No. 1. Before opening it, the host (who knows where the car is) opens another door, say No. 3, which has a goat. And then he gives you the possibility to change, allowing you to pick door No. 2. What do you do?

Nowadays, the answer for this problem is well known: the best strategy is changing the selected door. However, in the early 90’s a publication of M. Vos Savant [23] generated a controversy, because of the answer (solution) given to the problem. Essentially, the article exposed the intuitive arguments of the right (mathematical) solution proposed by Selvin [20]. Nevertheless, some mathematicians did not agree with Vos Savant’ explanation and stated that it does not matter if you change the selected door. Since there are two remaining doors, the probability of door No. 1 (or No. 2) having a car is 1/2. Subsequent articles [24, 25] discussed some opinions that defended the latter argument, these criticized Vos Savant even saying something like:

May I suggest that you obtain and refer to a standard textbook on probability before you try to answer a question of this type again?

In fact, a standard textbook on probability theory [2, 5, 17] is enough to confirm Vos Savant’ solution. Clearly, she was sure about that and in [24] encouraged the readers to check empirically her answer to the
Monty Hall problem. Even more, Vos Savant [23] suggests 1.000.000 doors rather than 3. In this case, there are 999.999 doors with goats behind them and one door with a prize. After the player picks a door, the host opens 999.998 of the remaining doors. On average, in 999.999 times out of 1.000.000, the remaining door will contain the prize.

In this sense, the literature presents several solutions [21, 14, 4, 9, 3, 11, 13, 8]. We remark that, the Monty Hall problem is relevant in the teaching of probability, because of the misperception that it can generate. In this sense, initial concepts from theory of probability can be studied by the statement of this problem. For instance, conditional probability and the law of total probability. However, a set theory approach is also of interest, because of the several possibilities to define the events of interest. Our proposal is based in the definition of the following three events:

- $E$: the contestant initially chooses the correct door.
- $C$: the contestant changes the selected door.
- $W$: the contestant wins the car.

At this point the notion of partition of a set is relevant to understand the solution that we propose (see for details [11]). A partition divides a set into separate and non-empty events. For instance, Figure 1 includes a Venn diagram representing a partition with eight events, where the union of these subsets is equal to the sample space. The same collection is shown in the tree diagram in Figure 2.

![Venn diagram and tree diagram](image.png)

**Figure 1:** Venn diagram associated to the three events above.

In this paper we discuss the implicit assumptions made by the different approaches in the literature and discussed their importance in the solution of the problem (see [10]). The advantage of our proposal is its simplicity. The definition of the three above events allow us to solve the problem with elementary probabilistic tools, turning the problem a candidate to introduce the concepts of conditional probability in a first undergraduate course or even more, in a lecture for secondary school.

The theoretical solutions are then checked by numerical simulations and the pseudo-codes are included to allow the replication of our results. The goal is to mimic the empirical evidence supporting the solutions, as required by Vos Savant in [23].

The rest of the paper is organized as follows. In Section 2 we include the formulation of the probability models and present their theoretical solutions; Section 3 is dedicated to the numerical simulations of the solutions, including the arguments to choose the numbers of iterations. Finally, Section 4 concludes providing final remarks.
2 The probability model and its solution: computing conditional and marginal probabilities of winning

We propose an approach which allows us to solve the MH problem in the original form and its generalizations as proposed in [21] and [24], that is, the existence of \( n \) doors, with \( n \geq 3 \).

We use the chain rule in probability to obtain, for instance

\[
P(E \cap C \cap W) = P(W \mid E \cap C) \cdot P(C \mid E) \cdot P(E),
\]

also for \( \{E \cap C \cap W^c\}, \{E \cap C^c \cap W\} \) and so on, for all eight subsets in the partition proposed by Venn diagram or for each of the branches in the tree diagram in Figure 2.

**Assumptions 2.1.** We solve the MH problem assuming that initially the contestant selects the door uniformly at random, that is, \( P(E) = \frac{1}{n} \), with \( n \geq 3 \). Moreover, we suppose that the contestant does not know if the initially chosen door has the car and takes the decision to change the door with a given probability \( p \in [0, 1] \), then, \( P(C \mid E) = P(C \mid E^c) = p \).

Then, we compare the conditional probabilities

\[
P(W \mid C) \text{ and } P(W \mid C^c),
\]

of winning the prize by changing and winning by staying, respectively and concluding which strategy is better. In addition, we obtain the marginal probability of winning, \( P(W) \).

This assumption generalizes previous solutions which implicitly assumed equal probabilities, \( p = 1/2 \). That is, the contestant has not a preference to change or not, making his (her) decision at random. For details, see comments in Section 4.
We solve the MH problem and generalize it in the sense proposed by Vos Savant [23]. Then we suppose there are \( n \) doors and consider two different strategies assumed by the host to open door(s) after the contestant initial choice. In the first case, in Section 2.2, the host opens \((n - 2)\) doors, then the contestant has to choose from the remaining 2 doors. In the second case, Section 2.3, the host opens just one door and then the contestant has to choose from the remaining \((n - 1)\) doors. In the original MH problem, with 3 doors, both strategies coincide.

In Section 3 we will develop numerical simulations of both strategies, aiming to check the theoretical solution, as suggested by Vos Savant to her readers in [23].

2.1 Original MH problem

There are 3 doors, then by Assumptions 2.1,

\[
P(E) = \frac{1}{3} \text{ and } P(C|E) = p. \tag{2.3}
\]

Therefore, the probability in (2.1) is computed by

\[
P(E \cap C \cap W) = 0 \cdot p \cdot \frac{1}{3}, \tag{2.4}
\]

given that if the contestant chooses the correct door and changes it, he is finally selecting a wrong door, that is, losing with probability 1. Now, we obtain

\[
P(E \cap C \cap W^c) = P(W^c|E \cap C) \cdot P(C|E) \cdot P(E) = 1 \cdot p \cdot \frac{1}{3}, \tag{2.5}
\]

Moreover,

\[
P(E^c \cap C^c \cap W^c) = P(W^c|E^c \cap C^c) \cdot P(C^c|E^c) \cdot P(E^c) = \frac{2}{3}(1 - p),
\]

\[
P(E \cap C^c \cap W) = P(W|E \cap C^c) \cdot P(C^c|E) \cdot P(E) = \frac{1}{3}(1 - p), \tag{2.6}
\]

\[
P(E^c \cap C \cap W) = P(W|E^c \cap C) \cdot P(C|E^c) \cdot P(E^c) = \frac{2}{3}p,
\]

and by similar arguments

\[
P(E \cap C^c \cap W^c) = P(E^c \cap C^c \cap W) = P(E^c \cap C \cap W^c) = 0. \tag{2.7}
\]

Therefore, the conditional probability of winning by changing the door is

\[
P(W|C) = \frac{P(W \cap C)}{P(C)} = \frac{P(E \cap C \cap W) + P(E^c \cap C \cap W)}{P(C)} = \frac{0 + \frac{2}{3}p}{p} = \frac{2}{3}, \tag{2.8}
\]

and the conditional probability of winning by keeping the door is

\[
P(W|C^c) = \frac{P(W \cap C^c)}{P(C^c)} = \frac{P(E \cap C^c \cap W) + P(E^c \cap C^c \cap W)}{P(C^c)} = \frac{\frac{1}{3}(1 - p) + 0}{1 - p} = \frac{1}{3}. \tag{2.9}
\]

In this sense, the probabilities in (2.8) and (2.9) are in accordance with the solution given by Selvin [20] and Vos Savant [23]. That is, the best strategy is changing the initially selected door.

Finally, by using the law of total probability
\[ P(W) = P(W|C)P(C) + P(W|C^c)P(C^c) \]
\[ = \frac{1}{3} + \frac{1}{3}p. \quad (2.10) \]

This probability of winning is a linear function of the parameter \( p \) (probability of change). In particular, the intercept and the slope of this function are the same (1/3). This fact implies that as \( p \) grows the probability in (2.10) increases. See Figure 3.

2.2 MH with \( n \) doors, the host leaves two closed doors

There are \( n \) doors and the assumption is kept.

\[ P(E) = \frac{1}{n} \quad \text{and} \quad P(C|E) = p. \quad (2.11) \]

In this case, all conditional probabilities \( P(W|E \cap C), P(W|E \cap C^c), P(W|E^c \cap C) \) and \( P(W|E^c \cap C^c) \) remain the same as in previous case. Finally, the conditional probabilities of winning by changing or keeping the door are, respectively

\[ P(W|C) = \frac{n-1}{n} \quad \text{and} \quad P(W|C^c) = \frac{1}{n}. \quad (2.12) \]

Therefore \( P(W|C) = P(E^c) \) and \( P(W|C^c) = P(E) \), given the dichotomy in the final selection, that is, one door with a car and another with a goat.

Finally

\[ P(W) = \frac{1}{n} + \frac{(n-2)}{n}p. \quad (2.13) \]

As in previous case, the probability of winning is a linear function of parameter \( p \). Of course, for \( n = 3 \), (2.13) equals to (2.10). However, in this situation, both the intercept and the slope depend on the number of doors. In this sense, as \( n \) grows to infinity, the intercept goes to zero and the slope is going to one. In other words, for a large number of doors, the probability of winning coincides with the probability of change. That is, \( \lim_{n \to \infty} P(W) = p \). For details see Figure 4.

2.3 MH with \( n \) doors, the host opens one door

This problem was mentioned in [21], credited to D.L. Ferguson. There are \( n \) doors and the host opens one door. In this case

\[ P(E) = \frac{1}{n} = 1 - P(E^c), \quad (2.14) \]

and by Assumptions 2.1 \( P(C|E) = P(C|E^c) = p \). In addition, we obtain

\[ P(W|E \cap C) = P(W|E^c \cap C^c) = P(W^c|E \cap C^c) = 0. \quad (2.15) \]

Which in turn implies that

\[ P(W^c|E \cap C) = P(W^c|E^c \cap C^c) = P(W|E \cap C^c) = 1. \quad (2.16) \]
Now, by assuming that the contestant decides to change the door and chooses uniformly at random the new
door, we compute
\[ P(W|E^c \cap C) = \frac{1}{n-2}, \] (2.17)
that is, there is one door with the car belonging to the \( n-2 \) remaining doors (eliminating the opened and
the initially selected doors). Therefore, we also have
\[ P(W^c|E^c \cap C) = 1 - P(W|E^c \cap C) = \frac{n-3}{n-2}, \] (2.18)

Finally, we compute the conditional probabilities of winning by changing or keeping the initial door,
given by
\[
P(W|C) = \frac{P(W \cap C)}{P(C)} = \frac{P(E \cap C \cap W) + P(E^c \cap C \cap W)}{P(C)} = \frac{0 + \frac{n-1}{n(n-2)}p}{\frac{p}{n}} = \frac{n-1}{n(n-2)};
\] (2.19)
and
\[
P(W|C^c) = \frac{P(W \cap C^c)}{P(C^c)} = \frac{P(E \cap C^c \cap W) + P(E^c \cap C^c \cap W)}{P(C^c)} = \frac{\frac{1}{n}(1-p) + 0}{1-p} = \frac{1}{n}, \] (2.20)
respectively. In the case of the probability of winning, we obtain
\[
P(W) = P(W|C)P(C) + P(W|C^c)P(C^c)
= p \cdot \frac{n-1}{n(n-2)} + (1-p) \frac{1}{n}
= \frac{1}{n} + \frac{1}{n(n-2)}p.
\] (2.21)

Note that, as in previous cases, this probability is a linear function of parameter \( p \). In this sense, as \( n \) grows
to infinity, the intercept and the slope go to zero. That is, for a large number of doors, the probability of
winning goes to zero independently of the contestant decision to change the initially selected door or not.
Then, \( \lim_{n \to \infty} P(W) = 0 \). For details see Figure 5.

![Figure 3: Empirical probabilities of winning for the model in Section 2.1, the straight line is the function of parameter \( p \) in equation (2.10). Several numbers of steps to check the law of large numbers.](image)
Figure 4: Examples of empirical probabilities of winning for the problem in Section 2.2. In last column we use the optimal $l_0$ given by CLT. Upper part with 5 doors. Bottom part with 10 doors.
3 Simulation studies

The main goals of this section are twofold: on the one hand, we analyze the performance of the strategies to check the theoretical solution. On the other hand, we include the arguments to choose the number of iterations that guarantees the empirical probabilities of winning in (2.10), (2.13) and (2.21) converge to theoretical solutions by the strong law of large numbers (SLLN). The critical values are obtained using the bounds provided by the central limit theorem and Chebyshev inequality.

We develop numerical simulations for the three models in Section 2. In particular, we analyse the probabilities of winning in (2.10), (2.13) and (2.21), as functions of parameter $p$. The empirical probabilities converge to theoretical values by the SLLN. In other words, let a sequence of independent Bernoulli random variables $X_i$, $i = 1, \ldots, l$, where

$$X_i = \begin{cases} 1, & \text{if the } i\text{-th contestant win the car,} \\ 0, & \text{otherwise.} \end{cases}$$

By the SLLN we have

$$\frac{1}{l} \sum_{i=1}^{l} X_i \overset{a.s.}{\to} \mathbb{E}(X_i) = \mathbb{P}(W) \text{ as } l \to \infty.$$  

This means that as long we simulate the process, increasing the number of contestants, the empirical probabilities get closer the theoretical ones. However, we need an additional tool to find a bound for the optimal number of iterations. Then, by using the central limit theorem (CLT) and denoting $\bar{X}_l = \frac{1}{l} \sum_{i=1}^{l} X_i$, we assume that $\bar{X}_l \sim N\left(\mathbb{P}(W), \frac{\mathbb{P}(W)(1-\mathbb{P}(W))}{l}\right)$ for $l$ large enough. In other words $\bar{X}_l$ is distributed asymptotically normal with mean $\mathbb{P}(W)$ and variance $\frac{\mathbb{P}(W)(1-\mathbb{P}(W))}{l}$. From the above we obtain the value of $l_0$, such that for all $l > l_0$:

$$\mathbb{P}\left(\left|\bar{X}_l - \mathbb{P}(W)\right| \geq \varepsilon\right) < \delta.$$  

Then, by straightforward calculations we obtain

$$\mathbb{P}\left(\bar{X}_l - \mathbb{P}(W) < \varepsilon\right) > 1 - \frac{\delta}{2} \implies \mathbb{P}\left(Z < \frac{\varepsilon \sqrt{l}}{\sqrt{\mathbb{P}(W)(1-\mathbb{P}(W))}}\right) > 1 - \frac{\delta}{2},$$

where $Z \sim N(0, 1)$. Denote $x = 1 - \frac{\delta}{2}$, then find $z_x = \frac{\varepsilon \sqrt{l}}{\sqrt{\mathbb{P}(W)(1-\mathbb{P}(W))}}$. Therefore,

$$l_0^{CL} = \frac{z_x^2 \mathbb{P}(W)(1-\mathbb{P}(W))}{\varepsilon^2}. \quad (3.1)$$

We remark that the approximation obtained by the CLT for the probability in (2.21) is worse if the number of doors $n$ is larger. Then, as an alternative, we use the Chebyshev inequality. Then

$$l_0^{Ch} = \frac{\mathbb{P}(W)(1-\mathbb{P}(W))}{\delta \varepsilon^2}. \quad (3.2)$$
Figure 5: Generalized Monty Hall problem in Section 2.3, several number of steps. Second column the bounds by using CLT, last one by using Chebishev inequality. First row with 4 doors, second with 5 doors, third with 8 doors and the last with 15 doors.
In each MH problem we use $\varepsilon = \delta = 0.01$ to obtain the corresponding critical $l_0$. For instance, in (2.10) represented by Figure 3, we consider three scenarios with 200, 1000 and 20000 contestants. The last one was obtained using $l_{CL}^0$ in (3.1).

The second problem in (2.13) is represented by Figure 4, last column considers the optimal number of iterations using CTL inequality (see (3.1)). In particular we consider two cases with 5 and 10 doors.

Finally, in (2.21) represented by Figure 5, the second and third scenarios consider the optimal number of iterations using $l_{CL}^0$ and $l_{Ch}^0$, respectively. For this MH version, we consider four cases with 4, 5, 8 and 15 doors.

### 4 Comments and concluding remarks

Some authors have considered the assumption that MH chooses randomly the door with the car (at the beginning). However, this is not relevant. Define

$$A_i : \text{MH puts the car behind door } i,$$

and let $P(A_i) = \alpha_i$ where $1 \leq i \leq n$ and $\sum_i \alpha_i = 1$. We are able to prove

$$P(E) = \sum_{i=1}^{n} P(E|A_i)P(A_i) = \sum_{i=1}^{n} \frac{1}{n} \alpha_i = \frac{1}{n}$$

Then Assumptions 2.1 are enough.

In addition, note that in the case of Monty Hall problem in Section 2.1, we obtain $P(W) = \frac{1}{3}(1+p)$ and in particular, if we assume $p = \frac{1}{2}$, then $P(W) = \frac{1}{2}$. That is, the solution proposed by the detractors of Vos Savant. In this sense, we claim that the intuition that leads to such solution is the implicit assumption that the contestant does not have a preference to change its initially selected door. In any case, the original question was posed as: Is it better to change or not? Then, Vos Savant solution is absolutely correct.

We remark that for the models in Sections 2.2 and 2.3, the probabilities of winning respectively (2.13) and (2.21), are linear functions of $p$. However, as $n$ goes to infinity, the slopes of these curves change in different ways. Particularly, (2.13) converges to $p$, then its slope goes to 1, and (2.21) converges to 0 with slope going to 0, see Figures 4-5.

In a similar sense, we also remark that there exists a critical $n$ such that (2.21) is strictly smaller than 1/3, that is the probability (2.10) when $p = 0$. In this case, the critical value is $n = 4$. Then, for all $n > 4$ the game in Section 2.3 is more advantageous for Monty Hall than the original one (Section 2.1), see Figure 5.

Finally, we highlight the relationship between the original Monty Hall problem and the well-known 3 prisoners paradox (see [3, 6]). Essentially, the probabilities involved in both problems are the same, where the arguments to calculate them are similar.

### 5 Appendix

Finally, we include the pseudo-codes. For that, we need to define

$$\mathcal{P} = \{0.05 \cdot k : 0 \leq k \leq 20\}.$$ (5.1)
Algorithm 1 Monty Hall $n$ doors, original version

1: fix the number of contestants $l$
2: for all $p \in \mathcal{P}$ in (5.1) do
3:   for all $i = 1, \ldots, l$ do
4:     fix the door with the car (for instance 1)
5:     let $W = 0$ (initial number of winnings)
6:     choose $X \in \{1, \ldots, n\}$ uniformly (initial selected door)
7:     if $X = 1$ then
8:       choose $Y \in \{2, \ldots, n\}$ uniformly
9:     else
10:       choose $Y = 1$
11:     end if
12:     let $(X, Y)$ closed doors
13:     sample $C \sim Bern(p)$
14:     if $C = 0$ then
15:       $Z = X$ (final selected door)
16:     else
17:       $Z = Y$
18:     end if
19:     if $Z = 1$ then
20:       $W = W + 1$
21:     else
22:       $W = W$
23:     end if
24:   end for
25: end for
26: plot values of arrays $(p, W/l)$
Then, the Algorithm 1 is used for models in Sections 2.1 and 2.2. The Algorithm 2 is related to the model in Section 2.3.

**Algorithm 2** Monty Hall $n$ doors, opening one

1: fix the number of contestants $l$
2: for all $p \in P$ in (5.1) do
3: for all $i = 1, \ldots, l$ do
4: fix the door with the car (for instance 1)
5: let $W = 0$ (initial number of winnings)
6: choose $X \in \{1, \ldots, n\}$ uniformly (initial selected door)
7: if $X = 1$ then
8: choose $Y \in \{2, \ldots, n\}$ uniformly
9: else
10: choose $Y \in \{2, \ldots, X - 1, X + 1, \ldots, n\}$ uniformly
11: end if
12: let $\mathcal{R} = \{1, \ldots, n\} \setminus \{X, Y\}$ (remaining closed doors)
13: sample $C \sim \text{Bern}(p)$
14: if $C = 0$ then
15: $Z = X$ (final selected door)
16: else
17: choose $Z \in \mathcal{R}$ uniformly
18: end if
19: if $Z = 1$ then
20: $W = W + 1$
21: else
22: $W = W$
23: end if
24: end for
25: end for
26: plot values of arrays $(p, W_l)$

**Acknowledgements**

CCC was partially supported by DIUBB 173408 2/I of the Universidad del Bío-Bío. MGN is supported by Fondo Especial DIUBB 1901083-RS.

**References**

[1] Bartoszynski, R., Niewiadomska-Bugaj, M. (1996) *Probability and Statistical Inference*. John Wiley & Sons, New York.

[2] Bertsekas, D.P., Tsitsiklis, J.N. (2002). *Introduction to Probability, vol. 1*. Athena Scientific Belmont, MA.
[3] Carlton, M. *Pedigrees, Prizes, and Prisoners: The Misuse of Conditional Probability*. Journal of Statistics Education. 13(2). (2005).

[4] Chun, Young H. *Game Show Problem*. OR/MS Today. 18(3): 9. (1991).

[5] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications, Vol. 1*. John Wiley, New York.

[6] Gardner, M. *Mathematical Games: Problems involving questions of probability and ambiguity*. Scientific American 201(4): 174-182 (1959).

[7] Gill, R. *Monty Hall problem*. International Encyclopedia of Statistical Science. Springer. pp. 858-863 (2011).

[8] Gill, R. *The Monty Hall Problem is not a probability puzzle (it’s a challenge in mathematical modelling)*. Stat Neerl. 65(1): 58-71 (2011).

[9] Gillman, L. *The Car and the Goats*. American Mathematical Monthly. 99: 3-7 (1992).

[10] González-Navarrete, M. *La importancia de los supuestos en la creación y resolución de problemas del área de probabilidad*. Revista digital matemática, educación e internet 17(1) (2016).

[11] Grinstead, C. M., Snell, J. L. (1997). *Introduction to Probability*. Chapman & Hall/CRC Texts in Statistical Science.

[12] Hogbin, M., Nijdam, W. *Letter to the editor*. Am. Stat. 64: 193 (2010).

[13] Lucas, S., Rosenhouse, J., Schepler, A. *The Monty Hall Problem, Reconsidered*. Mathematics Magazine. 82(5). (2009).

[14] Morgan, J.P., Chaganty, N.R., Dahiya, R.C., Doviak, M.J. *Lets Make a Deal: The Players Dilemma*. Am. Stat. 45(4): 284-287 (1991).

[15] Morgan, J.P., Chaganty, N.R., Dahiya, R.C., Doviak, M.J. *Rejoinder to Seymanns comment on Lets make a deal: the player’s dilemma*. Am. Stat. 45(4): 289 (1991).

[16] Morgan, J.P., Chaganty, N.R., Dahiya, R.C., Doviak, M.J. *Response to Hogbin and Nijdams letter*. Am. Stat. 64: 193-194 (2010).

[17] Rincon, L. (2014). *Introducción a la probabilidad*. Las Prensas de Ciencias, Fac. de Ciencias, UNAM.

[18] Rosenhouse, J. (2009). *The Monty Hall Problem*. Oxford University Press, New York.

[19] Ross, S. M. (2010). *A First Course in Probability*. (8th Edition). Prentice Hall.

[20] Selvin, S. *A problem in probability*. Am. Stat. 29: 67 (1975).

[21] Selvin, S. *On the Monty Hall problem*. Am. Stat. 29: 134 (1975).

[22] Seyman, R.G. *Comment on Lets make a deal: the players dilemma*. Am. Stat. 45(4): 287-288 (1991).

[23] Vos Savant, M. *Ask Marilyn*. Parade Magazine, 9 September 1990, p. 15.

[24] Vos Savant, M. *Ask Marilyn*. Parade Magazine, 2 December 1990, p. 25.

[25] Vos Savant, M. *Marilyn vos Savants reply*. Am. Stat. 45: 347-348 (1991).