Paradoxes in probability theory

Mathijs Kolkhuis Tanke

De Leidsche Flesch

March 20, 2019

Table of contents

The problem

Monty Hall's game

- There are three doors called a, b or c. One has a car, the others have a goat.
- Player chooses a door. Without loss of generality choose a.
- Game master opens one of the remaining doors. That door has a goat.
- Player is asked to switch to the other unopened door.
- Player wins contents of the door he ultimately chooses.

What is the probability of the other door having a car?

Early 'solutions'

What is the probability of the other door having a car?

Answer is $\frac{1}{2}$:

- There are two doors when the player can switch.
- One has a car, the other not.
- The other door has $\frac{1}{2}$ chance of having a car.

Answer is $\frac{2}{3}$:

- Chance of doors b and c having a car is $\frac{2}{3}$.
- One of them is opened.
- Thus the other door has chance $\frac{2}{3}$ of having a car.

In fact, both are wrong. Probability ranges from $\frac{1}{2}$ to 1.

Early 'solutions'

What is the probability of the other door having a car?

Answer is $\frac{1}{2}$:

- There are two doors when the player can switch.
- One has a car, the other not.
- The other door has $\frac{1}{2}$ chance of having a car.

Answer is $\frac{2}{3}$:

- Chance of doors b and c having a car is $\frac{2}{3}$.
- One of them is opened.
- Thus the other door has chance $\frac{2}{3}$ of having a car.

In fact, both are wrong. Probability ranges from $\frac{1}{2}$ to 1.

Early 'solutions'

What is the probability of the other door having a car?

Answer is $\frac{1}{2}$:

- There are two doors when the player can switch.
- One has a car, the other not.
- The other door has $\frac{1}{2}$ chance of having a car.

Answer is $\frac{2}{3}$:

- Chance of doors b and c having a car is $\frac{2}{3}$.
- One of them is opened.
- Thus the other door has chance $\frac{2}{3}$ of having a car.

In fact, both are wrong. Probability ranges from $\frac{1}{2}$ to 1.

Probability spaces

Definition (σ-algebra)

Let Ω be a set and $\mathcal{P}(\Omega)$ its power set.
The set $\Sigma \subseteq \mathcal{P}(\Omega)$ is a *σ-algebra* iff

1. $\Omega \in \Sigma$,
2. Σ is closed under countable unions,
3. Σ is closed under countable intersections and
4. Σ is closed under complement.

Definition (Probability space)

Let Ω be a set and Σ a σ-algebra on Ω. Let $\mathbb{P}$ be a probability measure on Σ.
Then $(\Omega, \Sigma, \mathbb{P})$ is a probability space.

Probability spaces

Definition (σ-algebra)

Let Ω be a set and $\mathcal{P}(\Omega)$ its power set.
The set $\Sigma \subseteq \mathcal{P}(\Omega)$ is a *σ-algebra* iff

1. $\Omega \in \Sigma$,
2. Σ is closed under countable unions,
3. Σ is closed under countable intersections and
4. Σ is closed under complement.

Definition (Probability space)

Let Ω be a set and Σ a σ-algebra on Ω. Let $\mathbb{P}$ be a probability measure on Σ.
Then $(\Omega, \Sigma, \mathbb{P})$ is a probability space.

Conditional expectation

We need to extend $\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$.

Definition (Conditional expectation)

Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space.
Let $\Sigma' \subseteq \Sigma$ be a sub-σ-algebra of Σ.
The *expectation value of X conditioned on Σ'* is $Y = \mathbb{E}[X|\Sigma']$ s.t.

- Y is Σ'-measurable,
- $\mathbb{E}[|Y|]$ is finite and
- $\int_A Y d\mathbb{P} = \int_A X d\mathbb{P}$ for all $A \in \Sigma'$.

Definition (Conditional probability)

The *probability of A conditioned on Σ'* is $\mathbb{P}[A|\Sigma'] = \mathbb{E}[\mathbb{1}_A|\Sigma']$.

Conditional expectation

We need to extend $\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$.

Definition (Conditional expectation)

Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space.
Let $\Sigma' \subseteq \Sigma$ be a sub-σ-algebra of Σ.
The *expectation value of X conditioned on Σ'* is $Y = \mathbb{E}[X|\Sigma']$ s.t.

- Y is Σ'-measurable,
- $\mathbb{E}[|Y|]$ is finite and
- $\int_A Y d\mathbb{P} = \int_A X d\mathbb{P}$ for all $A \in \Sigma'$.

Definition (Conditional probability)

The *probability of A conditioned on Σ'* is $\mathbb{P}[A|\Sigma'] = \mathbb{E}[\mathbb{1}_A|\Sigma']$.

Conditional expectation

We need to extend $\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$.

Definition (Conditional expectation)

Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space.
Let $\Sigma' \subseteq \Sigma$ be a sub-σ-algebra of Σ.
The *expectation value of X conditioned on Σ'* is $Y = \mathbb{E}[X|\Sigma']$ s.t.

- Y is Σ'-measurable,
- $\mathbb{E}[|Y|]$ is finite and
- $\int_A Y d\mathbb{P} = \int_A X d\mathbb{P}$ for all $A \in \Sigma'$.

Definition (Conditional probability)

The *probability of A conditioned on Σ'* is $\mathbb{P}[A|\Sigma'] = \mathbb{E}[\mathbb{1}_A|\Sigma']$.

Back to Monty Hall

The situation:

- $\mathcal{X} = \{a, b, c\}$ is set of doors player can choose.
- Take $\Omega = \mathcal{X}$.
- σ-algebra is $\mathcal{P}(\Omega)$.
- Condition on smallest σ-algebra having $\{a, b\}$ and $\{a, c\}$.

However, $\sigma(\{a, b\}, \{a, c\}) = \mathcal{P}(\Omega)$ is original σ-algebra.
No information is gained using this conditioning!

Back to Monty Hall

The situation:

- $\mathcal{X} = \{a, b, c\}$ is set of doors player can choose.
- Take $\Omega = \mathcal{X}$.
- σ-algebra is $\mathcal{P}(\Omega)$.
- Condition on smallest σ-algebra having $\{a, b\}$ and $\{a, c\}$.

However, $\sigma(\{a, b\}, \{a, c\}) = \mathcal{P}(\Omega)$ is original σ-algebra.
No information is gained using this conditioning!

Extending the set of possibilities

The situation:

- $\mathcal{X} = \{a, b, c\}$ is set of doors player can choose.
- $\mathcal{Y} = \{b, c\}$ is set of doors Monty Hall can open.
- Take $\Omega = \mathcal{X} \times \mathcal{Y}$.
- σ-algebra is $\mathcal{P}(\Omega)$.
- X: r.v. on $\mathcal{X}$ player choosing a door.
- Y: r.v. on $\mathcal{Y}$ which door to open after player's choice.

Then

$$\mathbb{P}[X = a | Y = b] = \frac{\mathbb{P}[X = a, Y = b]}{\mathbb{P}[Y = b]} = \frac{\mathbb{P}[X = a, Y = b]}{\sum_{x \in \mathcal{X}} \mathbb{P}[Y = b | X = x] \mathbb{P}[X = x]}.$$

$\mathbb{P}[X = x]$ is unknown, thus the conditional probability cannot be calculated. Real answer has many variables.

Extending the set of possibilities

The situation:

- $\mathcal{X} = \{a, b, c\}$ is set of doors player can choose.
- $\mathcal{Y} = \{b, c\}$ is set of doors Monty Hall can open.
- Take $\Omega = \mathcal{X} \times \mathcal{Y}$.
- σ-algebra is $\mathcal{P}(\Omega)$.
- X: r.v. on $\mathcal{X}$ player choosing a door.
- Y: r.v. on $\mathcal{Y}$ which door to open after player's choice.

Then

$$\begin{aligned}\mathbb{P}[X = a|Y = b] &= \frac{\mathbb{P}[X = a, Y = b]}{\mathbb{P}[Y = b]} \\ &= \frac{\mathbb{P}[X = a, Y = b]}{\sum_{x \in \mathcal{X}} \mathbb{P}[Y = b|X = x]\mathbb{P}[X = x]}.\end{aligned}$$

$\mathbb{P}[X = x]$ is unknown, thus the conditional probability cannot be calculated. Real answer has many variables.

Example

Probability $\frac{1}{2}$

If

- the car is distributed uniformly between the doors,
- door a is first chosen,
- door a actually has the car and
- Monty Hall never opens door b in this case,

then probability of having the correct door is $\frac{1}{2}$.

The Borel-Kolmogorov paradox

The setting of the Borel-Kolmogorov paradox

- Take a sphere and distribute points uniformly.
- Draw a great circle through one point.
- Mask all distributed points.

What is the distribution of the point being on the great circle?

Answer: depends on parametrizations.

The Borel-Kolmogorov paradox

The setting of the Borel-Kolmogorov paradox

- Take a sphere and distribute points uniformly.
- Draw a great circle through one point.
- Mask all distributed points.

What is the distribution of the point being on the great circle?
Answer: depends on parametrizations.

Bayesian probability

Let C be the great circle and $E \subset C$ a subset. Then

$$\mathbb{P}[E|C] = \frac{\mathbb{P}[E \cap C]}{\mathbb{P}[C]} = \frac{0}{0}.$$

The conditional probability is undefined!

Widening our view

Let $S = [0, 2\pi] \times [0, \pi]$ be the unit sphere with spherical coordinates.

The uniform probability measure on $A \subseteq S$ is given by

$$\mathbb{P}[A] = \frac{1}{4\pi} \iint_A \sin \psi d\phi d\psi.$$

Look at longitudes

The σ-algebra of longitudes is

$$\mathfrak{C} = \{[0, 2\pi] \times A | A \subseteq [0, \pi] \text{ is measurable}\}.$$

The conditional expectation on $\mathfrak{C}$ is

$$\mathbb{E}[X|\mathfrak{C}](\phi, \psi) = \frac{1}{2\pi} \int_0^{2\pi} X(\phi', \psi) d\phi'.$$

The conditional probability of $B = [\phi_1, \phi_2] \times \{\psi'\}$ on $[0, 2\pi] \times \{\psi'\}$ is

$$\bar{\mathbb{P}}[B] = \frac{\phi_2 - \phi_1}{2\pi},$$

thus the conditional probability stays uniform!

Look at longitudes

The σ-algebra of longitudes is

$$\mathfrak{C} = \{[0, 2\pi] \times A | A \subseteq [0, \pi] \text{ is measurable}\}.$$

The conditional expectation on $\mathfrak{C}$ is

$$\mathbb{E}[X|\mathfrak{C}](\phi, \psi) = \frac{1}{2\pi} \int_0^{2\pi} X(\phi', \psi) d\phi'.$$

The conditional probability of $B = [\phi_1, \phi_2] \times \{\psi'\}$ on $[0, 2\pi] \times \{\psi'\}$ is

$$\bar{\mathbb{P}}[B] = \frac{\phi_2 - \phi_1}{2\pi},$$

thus the conditional probability stays uniform!

Look at meridians

The σ-algebra of meridians is

$$\mathfrak{M} = \{A \times [0, \pi] | A \subseteq [0, 2\pi] \text{ is measurable}\}.$$

The conditional expectation on $\mathfrak{M}$ is

$$\mathbb{E}[X|\mathfrak{M}](\phi, \psi) = \frac{1}{2} \int_0^\pi X(\phi, \psi') \sin \psi' d\psi'.$$

The conditional probability of $B = \{\phi'\} \times [\psi_1, \psi_2]$ on $\{\phi'\} \times [0, \pi]$ is

$$\bar{\mathbb{P}}[B] = \frac{1}{2}(\cos \psi_1 - \cos \psi_2),$$

thus the conditional probability does not stay uniform!

Look at meridians

The σ-algebra of meridians is

$$\mathfrak{M} = \{A \times [0,\pi] | A \subseteq [0,2\pi] \text{ is measurable}\}.$$

The conditional expectation on $\mathfrak{M}$ is

$$\mathbb{E}[X|\mathfrak{M}](\phi,\psi) = \frac{1}{2}\int_0^\pi X(\phi,\psi')\sin\psi' d\psi'.$$

The conditional probability of $B = \{\phi'\} \times [\psi_1,\psi_2]$ on $\{\phi'\} \times [0,\pi]$ is

$$\bar{\mathbb{P}}[B] = \frac{1}{2}(\cos\psi_1 - \cos\psi_2),$$

thus the conditional probability does not stay uniform!

Recap

- Looking at longitudes the conditional distribution is uniform.
- Looking at meridians the conditional distribution is cosine.

Which one is right?

Theorem (Gyenis, Hofer-Szabó, Rédei (2017))

Consider probability spaces $(S, \mathfrak{C}, \mathbb{P})$ and $(S, \mathfrak{M}, \mathbb{P})$, then those spaces are not isomorphic.

Proof.

- Let $f : S \to S$ any measurable isomorphism.
- Let $h_f : \mathfrak{C} \to \mathfrak{M}$ be a Boolean algebra preserving bijection.
- If C is longitude, $h_f(C)$ is a meridian.
- All meridians cross in m_0, set of north and south poles.
- Let c_0 such that $h_f(c_0) = m_0$. Note that c_0 are longitudes.
- Then

$$\emptyset = h_f(\emptyset) = h_f(C \cap c_0) = h_f(C) \cap h_f(c_0) = h_f(C) \cap m_0 = m_0.$$

$\square$

Theorem (Gyenis, Hofer-Szabó, Rédei (2017))

Consider probability spaces $(S, \mathfrak{C}, \mathbb{P})$ *and* $(S, \mathfrak{M}, \mathbb{P})$, *then those spaces are not isomorphic.*

Proof.

- Let $f : S \to S$ any measurable isomorphism.
- Let $h_f : \mathfrak{C} \to \mathfrak{M}$ be a Boolean algebra preserving bijection.
- If C is longitude, $h_f(C)$ is a meridian.
- All meridians cross in m_0, set of north and south poles.
- Let c_0 such that $h_f(c_0) = m_0$. Note that c_0 are longitudes.
- Then

$$\emptyset = h_f(\emptyset) = h_f(C \cap c_0) = h_f(C) \cap h_f(c_0) = h_f(C) \cap m_0 = m_0.$$

$\square$

Discussion

Our two probability spaces are vastly different! So what now?

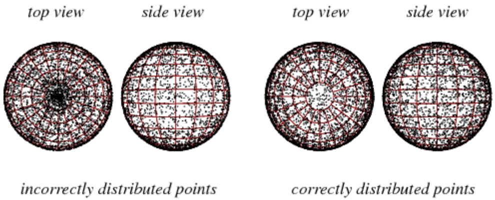

Figure: Simulation by Eric Weisstein

- Simulate the uniform distribution on the sphere by longitudes and meridians.
- If points are placed uniformly on longitudes or cosinely on meridians, the uniform distribution on the sphere is recovered!

Asking the right question

The question

'what is the conditional distribution on the zero set given I have points there'

is unanswerable.

The question for conditional probability must rather be:

'if I want to simulate the original distribution using this parametrization, what should my conditional probability be?'

The game

Two envelope game

- Game master picks a value x.
- Game master fills one envelope with x, other with $2x$.
- Game master gives an arbitrary envelope to the player.
- Player sees contents of his envelope.
- Player gets the option to switch to the other envelope.
- Should the player switch?

Call the player's envelope A, then

$$\mathbb{E}[B] = \frac{1}{2} \cdot \frac{x}{2} + \frac{1}{2} \cdot 2x = \frac{5}{4}x.$$

Player must always switch!

The game

Two envelope game

- Game master picks a value x.
- Game master fills one envelope with x, other with $2x$.
- Game master gives an arbitrary envelope to the player.
- Player sees contents of his envelope.
- Player gets the option to switch to the other envelope.
- Should the player switch?

Call the player's envelope A, then

$$\mathbb{E}[B] = \frac{1}{2} \cdot \frac{x}{2} + \frac{1}{2} \cdot 2x = \frac{5}{4}x.$$

Player must always switch!

Playing the game

Can I get two volunteers?

Probability space

Probability space:

- $\mathcal{X} = \{(x, 2x), (2x, x)\}$,
- $\Sigma = \mathcal{P}(\mathcal{X})$,
- $\mathbb{P}$ is uniform.

Then

$$\begin{aligned}\mathbb{E}[B] &= \mathbb{E}[B|A = x]\mathbb{P}[A = x] + \mathbb{E}[B|A = 2x]\mathbb{P}[A = 2x] \\ &= \frac{1}{2} \cdot 2x + \frac{1}{2}x = \frac{3}{2}x.\end{aligned}$$

Thus $\mathbb{E}[A] = \mathbb{E}[B] = \frac{3}{2}x$, switching does not help!
However, the player does not know x...

Probability space

Probability space:

- $\mathcal{X} = \{(x, 2x), (2x, x)\}$,
- $\Sigma = \mathcal{P}(\mathcal{X})$,
- $\mathbb{P}$ is uniform.

Then

$$\begin{aligned}\mathbb{E}[B] &= \mathbb{E}[B|A = x]\mathbb{P}[A = x] + \mathbb{E}[B|A = 2x]\mathbb{P}[A = 2x] \\ &= \frac{1}{2} \cdot 2x + \frac{1}{2}x = \frac{3}{2}x.\end{aligned}$$

Thus $\mathbb{E}[A] = \mathbb{E}[B] = \frac{3}{2}x$, switching does not help!
However, the player does not know x...

Prior on lowest value

Let X be the continuous r.v. for the lowest value x in the envelopes with density f.

- If $\mathbb{P}[B = 2a | A = a] = \mathbb{P}\left[B = \frac{a}{2} | A = a\right] = \frac{1}{2}$, then X is uniform on an unbounded set. This is not possible.
- Suppose $\mathbb{E}[X] < \infty$ to avoid the St. Petersburg paradox.

Switching strategies

Definition (Switching strategy)

Let $P\colon (0,\infty) \to [0,1]$ be arbitrary. Then P is a *switching strategy* if the player switches his envelope with probability $P(a)$ after observing a in his envelope.

Definition (Gain)

Using switching strategy P, the player's *gain* relative to never switching is

$$G = \frac{1}{2}\int_0^\infty x f(x)(P(x) - P(2x))dx.$$

Never switching has gain $G = 0$. If $G > 0$, then P is a better strategy than never switching. If $G < 0$, then P is a worse strategy than never switching.

Switching strategies

Definition (Switching strategy)

Let $P\colon (0,\infty) \to [0,1]$ be arbitrary. Then P is a *switching strategy* if the player switches his envelope with probability $P(a)$ after observing a in his envelope.

Definition (Gain)

Using switching strategy P, the player's *gain* relative to never switching is

$$G = \frac{1}{2}\int_0^\infty x f(x)(P(x) - P(2x))dx.$$

Never switching has gain $G = 0$. If $G > 0$, then P is a better strategy than never switching. If $G < 0$, then P is a worse strategy than never switching.

Switching strategies

Definition (Switching strategy)

Let $P\colon (0,\infty) \to [0,1]$ be arbitrary. Then P is a *switching strategy* if the player switches his envelope with probability $P(a)$ after observing a in his envelope.

Definition (Gain)

Using switching strategy P, the player's *gain* relative to never switching is

$$G = \frac{1}{2}\int_0^\infty x f(x)(P(x) - P(2x))dx.$$

Never switching has gain $G = 0$. If $G > 0$, then P is a better strategy than never switching. If $G < 0$, then P is a worse strategy than never switching.

Positive gains

Lemma

If P is decreasing and strictly decreasing on an interval I with $f(I) > 0$, then $G > 0$.

Example (Threshold switching)

Let $P(x) = \mathbb{1}_{[0,b]}(x)$, thus always switch if value lower than b is encountered. Then P is decreasing and has non-negative gains. If f is strictly decreasing, choose b such that $f(2b) = \frac{1}{4}f(b)$. This P is the best switching strategy possible.

Positive gains

Lemma

If P is decreasing and strictly decreasing on an interval I with $f(I) > 0$, then $G > 0$.

Example (Threshold switching)

Let $P(x) = \mathbb{1}_{[0,b]}(x)$, thus always switch if value lower than b is encountered. Then P is decreasing and has non-negative gains. If f is strictly decreasing, choose b such that $f(2b) = \frac{1}{4}f(b)$. This P is the best switching strategy possible.

Optimize threshold switching

Theorem (Egozcue and García (2015))

Let X be continuous with $\mathbb{E}[X] = \mu$ and $\mathbb{V}[X] = \sigma^2$. Let $m = \mu^2 + \sigma^2$. Let $P(x) = \mathbb{1}_{[0,b]}(x)$ be threshold switching.

1. *The gain has lower bound*

$$G(b) \geq \frac{3 + 2\sqrt{2}}{b} \left(\frac{3}{2} b\mu - \frac{1}{2} b^2 - m \right).$$

2. *If $\mu^2 > 8\sigma^2$, then $b^* = \sqrt{2m}$ is optimal with gain*

$$G(b^*) \geq (3 + 2\sqrt{2}) \left(\frac{3}{2}\mu - \sqrt{2m} \right) > 0.$$

Trade?

- Countdown from 3, tell if you want two switch.
- Switch when both players agree.
- After possibly switching reveal the contents of your envelopes.