The Borel-Kolmogorov paradox occurs in the case of a conditional expectation with respect to a set A of vanishing prior measure. The paper's thesis is that this can be analyzed as follows: A measure $q_{\mathcal{A}}$ on a sigma-algebra containing this set A is defined (e.g. by setting $q_{\mathcal{A}}(A) = 1$) and that transporting this measure on the original sigma-algebra $\mathcal{S}$ by setting $\psi(f) := \psi_{\mathcal{A}}(\mathcal{E}(f|\mathcal{A}))$ and then defining the posterior measure $\mu(B) := \psi(\chi_B)$ resolves or rather explains this phenomenon. The problem is that in the case $q_{\mathcal{A}}(A) = 1$ and A having prior measure 0, then the construction of ψ is not well-defined due too the measure $q_{\mathcal{A}}$ not being absolutely continuous with respect to the prior. This results in an ill-defined notion of ψ on the set of measure 0. This is seen most directly by looking at remark 4 and the following counterexample. Unfortunately, this technicality seems to weaken both the paper's analysis of the Borel-Kolmogorov paradox on the unit square (section 3) and the discussion of the full 3d paradox on the sphere (section 4 and 5). (Philipp Wacker, FAU Erlangen-Nürnberg, phkwacker@gmail.com)

1 A counterexample

Our underlying probability space is $([0,1], \mathcal{B}([0,1]), \lambda)$ with λ the Lebesgue measure. We first condition wrt to $\mathcal{A} := \{\emptyset, \{0\}, (0,1], [0,1]\}$.

The conditional expectation. By the measurability requirement, any version of conditional expectation is of the form $\mathcal{E}(f|\mathcal{A}) = \chi_{\{0\}} \cdot C_1 + \chi_{(0,1]} \cdot C_2$. The integration criterion yields $C_2 = \int_0^1 f(x)\mathrm{d}x$ but cannot resolve the value of C_1 (because $\{0\}$ is a 0-set of λ). This means that

$$\mathcal{E}(f|\mathcal{A}) = C \cdot \chi_{\{0\}} + \int_0^1 f(x)\mathrm{d}x \cdot \chi_{(0,1]}$$

for any $C \in \mathbb{R}$.

Choosing a measure $q_{\mathcal{A}}$ on $\mathcal{A}$. We set $q_{\mathcal{A}}(\{0\}) = \rho$ and $q_{\mathcal{A}}((0,1]) = 1-\rho$ for arbitrary but fixed $\rho \in [0,1]$. Now assume a $\mathcal{A}$-measurable function g, necessarily of the form $g = C_1 \cdot \chi_{\{0\}} + C_2 \cdot \chi_{(0,1]}$. Then the linear functional $\psi_{\mathcal{A}}$ applied to this is defined as

$$\psi_{\mathcal{A}}(g) = \rho \cdot C_1 + (1-\rho) \cdot C_2.$$

Extending the measure $q_{\mathcal{A}}$ to a measure q on $\mathcal{B}([0,1])$. We define the extension of the linear functional by $\psi(f) = \psi_{\mathcal{A}}(\mathcal{E}(f|\mathcal{A})$ for any f which is Borel-measurable. This is

$$\psi(f) = \rho \cdot C + (1-\rho) \cdot \int_0^1 f(x)\mathrm{d}x.$$

The extension of the measure is then defined as $q(B) = \psi(\chi_B)$.

When is ψ a proper extension of $\psi_{\mathcal{A}}$? We need to derive conditions such that $q(\{0\}) = q_{\mathcal{A}}(\{0\})$ and $q((0,1]) = q_{\mathcal{A}}((0,1])$. Concerning the first set, write $A = \{0\}$. Then

$$q(A) = \psi(\chi_A) = \psi_{\mathcal{A}}(\mathcal{E}(\chi_A|\mathcal{A})) = \rho \cdot C + (1-\rho) \cdot 0 = \rho \cdot C.$$

If this is supposed to be equal to $q_{\mathcal{A}}(A) = \rho$, we need $C = 1$.

Secondly, write $A^c = (0,1]$. Then (using $C = 1$)

$$q(A^c) = \psi(\chi_{A^c}) = \rho \cdot C + (1-\rho) = 1.$$

But this is only equal to $q_{\mathcal{A}}(A^c) = 1 - \rho$, if $\rho = 0$, i.e. if $\{0\}$ has $q_{\mathcal{A}}$-measure 0 and $q_{\mathcal{A}}$ is thus necessarily absolutely continuous with respect to $p = \lambda$.

Comparison to remark 4 in the paper The paper states in remark 4 that ψ will be an extension of $\psi_{\mathcal{A}}$ if $p(A) = 0$ (this is the case here) and $q_{\mathcal{A}}(A) = 1$. But if we enforce this condition (which means that $\rho = 1$), then from above we know that $q(A^c) \neq q_{\mathcal{A}}(A^c)$ and thus q is not an extension of $q_{\mathcal{A}}$.

The reason why things break down here is the following: On the one hand, from above,

$$\mathcal{E}(\chi_{(0,1]}|\mathcal{A}) = \chi_{(0,1]} + C \cdot \chi_{\{0\}}.$$

On the other hand, $\chi_{(0,1]}$ is $\mathcal{A}$-measurable, i.e. is unchanged by conditional expectation and

$$\mathcal{E}(\chi_{(0,1]}|\mathcal{A}) = \chi_{(0,1]}.$$

This makes sense, because conditional expectation is only define up to p-zero-sets (of which $\{0\}$ is one. But this ambiguity now makes a huge difference because $q_{\mathcal{A}}$ poses a non-zero probability here:

$$q((0,1]) = \psi(\chi_{(0,1]}) = C\rho + 1 - \rho$$

and as every value of C is equally valid, there is no canonical extension of $q_{\mathcal{A}}$. Even if we single out a version by setting $C = 0$ (i.e. such that $q((0,1]) = q_{\mathcal{A}}((0,1]) = 1 - \rho$), then the other set makes problems:

$$q(\{0\}) = \psi(\chi_{\{0\}}) = (1-\rho) \cdot \int_0^1 \chi_{\{0\}}(x)\mathrm{d}x = 0$$

which is a violation of the requirement $q(\{0\}) = q_{\mathcal{A}}(\{0\})$.

If we set $C = 1$, then $q(\{0\}) = q_{\mathcal{A}}(\{0\})$ but $1 = q((0,1]) \neq q_{\mathcal{A}}((0,1]) = 1 - \rho$. Hence we can never choose a consistent version of conditional expectation (which is not pointwise defined anyway) such that we can extend $q_{\mathcal{A}}$ to q. This is because the conditional expectation has an arbitrary value on a set ultimately due to the fact that $q_{\mathcal{A}}$ is not absolutely continuous wrt p.

We can generalize this example:

Lemma 1. *Let $(X, \mathcal{S}, p)$ be a probability space and $\mathcal{A}$ be a sub-σ-field of $\mathcal{S}$ and $q_{\mathcal{A}}$ a probability measure on $(X, \mathcal{A})$. Define the $\|\cdot\|_1$-continuous linear functional $\psi_{\mathcal{A}}$ defined via $q_{\mathcal{A}}$, i.e.*

$$\psi_{\mathcal{A}}(g) = \int_X g\mathrm{d}q_{\mathcal{A}}.$$

Assume that $q_{\mathcal{A}}$ is not absolutely continuous wrt p. Then there is no consistent extension of $q_{\mathcal{A}}$ to all of q via

$$q(B) = \psi(\chi_B) \doteq \psi_{\mathcal{A}}(\mathcal{E}(\chi_B|\mathcal{A})).$$

Proof. Consider a set $A \in \mathcal{A}$ such that $q_{\mathcal{A}}(A) > 0$ but $p(A) = p_{\mathcal{A}}(A) = 0$. Then for consistency we need $q(A) = q_{\mathcal{A}}(A)$, hence we compute

$$q(A) = \psi(\chi_A) = \psi_{\mathcal{A}}(\mathcal{E}(\chi_A|\mathcal{A}))$$

Now χ_A is $\mathcal{A}$-measurable, and thus $\mathcal{E}(\chi_A|\mathcal{A}) = \chi_A$. But $\mathcal{E}(\cdot|\mathcal{A})$ is only defined up to sets of p-measure 0, hence $\mathcal{E}(\chi_A|\mathcal{A}) = C \cdot \chi_A$ are valid versions for all values of C

$$= \psi_{\mathcal{A}}(C \cdot \chi_A) = C \cdot q_{\mathcal{A}}(A) \stackrel{!}{=} q(A)$$

(where the last equality is necessary for consistency of the two measures), and thus we need to choose $C = 1$.

On the other hand,

$$q(A^c) = \psi(\chi_{A^c}) = \psi_{\mathcal{A}}(\mathcal{E}(\chi_{A^c}|\mathcal{A}))$$

(again, A^c is $\mathcal{A}$-measurable but we need to account for arbitrariness in A)

$$= \psi_{\mathcal{A}}(C \cdot \chi_A + \chi_{A^c}) = C \cdot q_{\mathcal{A}}(A) + q_{\mathcal{A}}(A^c) \stackrel{!}{=} q(A^c)$$

and thus we need to choose $C = 0$. Hence even if we could "nail down" the conditional expectation on the set A by setting the constant C (which we cannot), there is no consistent way of doing so. The deeper reason for this problem here is that $\mathcal{E}(\cdot|\mathcal{A})$ is defined only uniquely up to measures of set 0 with respect to the prior measure p, while $\psi_{\mathcal{A}}$ is the functional defined by $q_{\mathcal{A}}$ and thus puts positive mass on the set A, hence the outer function $\psi_{\mathcal{A}}$ is very sensitive with respect to the set to which the inner function $\mathcal{E}(\cdot|\mathcal{A})$ is "agnostic". The technicality could thus be summarized as a "plug incompatibility". $\square$