A REVERSE TO THE JEFFREYS–LINDLEY PARADOX∗

BY

WIEBE R. PESTMAN (LEUVEN), FRANCIS TUERLINCKX (LEUVEN),
AND WOLF VANPAEMEL (LEUVEN)

Abstract. In this paper the seminal Jeffreys–Lindley paradox is regarded
from a mathematical point of view. We show that in certain scenarios
the paradox may emerge in a reverse direction.

2010 AMS Mathematics Subject Classification: Primary: 62A01,
62C10; Secondary: 60A10.

Key words and phrases: Hypothesis testing, Bayes factor, p-value,
reverse, paradox, Lindley, Jeffreys.

1. INTRODUCTION

The Jeffreys–Lindley paradox (see [2], [6], [9]) describes a discordance
between frequentist and Bayesian hypothesis testing. When comparing a simple null
hypothesis against a diffuse alternative, it has been found that a given sample may
simultaneously lead to a frequentist rejection of the null hypothesis (because the
p-value is smaller than a critical alpha) and a Bayesian support for the null hypoth-
esis (because the value of the Bayes factor exceeds some critical threshold). The
‘paradox’ has been the subject of intensive debate in the statistical literature and
this debate is still ongoing ([6], [10], [11]).

The classical example used to illustrate the discordance involves Gaussian
populations with known variance and a point null hypothesis for the mean versus
a diffuse alternative hypothesis. In words, the argument usually goes as follows:
A value for the test statistic is chosen that gives a small but constant p-value with
increasing sample size (hence the test value scales with the sample size), thus leading
to a systematic rejection of the null hypothesis in frequentist statistics. Under the
same scenario, it can then be shown that, with increasing sample sizes, the Bayes
factor of the alternative hypothesis over the null goes to zero. Thus, asymptoti-
cally, the Bayesian will favour the null hypothesis. The same phenomenon can be

∗ The research leading to the results reported in this paper was supported in part by the Research
Fund of the KU Leuven (OT/11/032 and CREA/11/005).
observed with several non-Gaussian populations. In this paper we show, however, that there also exist scenarios in which the Jeffreys–Lindley paradox may appear in a reversed way. That is to say, we show that it may occur that the frequentist asymptotically maintains the null hypothesis whereas the Bayesian rejects it.

2. THE JEFFREYS–LINDLEY PARADOX REVERSED

We will construct an example where the Jeffreys–Lindley paradox appears in a reversed way. As a first step, we define an auxiliary function \( f : \mathbb{R} \rightarrow [0, +\infty) \) as follows:

\[
(2.1) \quad f(x) = \begin{cases} 
A \frac{\exp \left( - \frac{\log(1/|x|)}{\log(\log(1/|x|))} \right)}{|x|} & \text{if } 0 < |x| < \frac{1}{e^2}, \\
0 & \text{elsewhere}
\end{cases}
\]

The constant \( A \) in the above formula is chosen in such a way that the function \( f \) integrates to one. Thus \( f \) can be considered a probability density in the strict sense of the word. This density has the following properties:

- \( f \) is symmetric around the origin.
- \( f \) is piecewise continuous.
- \( f \) has compact support and thus has moments of all orders.
- Due to a singularity at \( x = 0 \), all convolution powers \( f^{*n} \) are unbounded.

A proof of the last property can be found in [4]. In the following, the variance of \( f \) will be denoted by \( \sigma^2 \). A family \( \{f(x \mid \theta) \mid \theta \in \mathbb{R}\} \) of densities is defined as

\[
(2.2) \quad f(x \mid \theta) = f(x - \theta).
\]

Suppose now that a population is given with a probability density \( f(x \mid \theta) \), where the parameter \( \theta \), presenting the population mean, is unknown. In the following two subsections we will compare the Bayesian and frequentist approach when testing the null hypothesis \( H_0 : \theta = 0 \) against the alternative hypothesis \( H_1 : \theta \neq 0 \). To this end, in both scenarios, a sample \( X_1, X_2, \ldots, X_n \) is drawn from the population, and the sample mean \( Y_n \), defined as

\[
(2.3) \quad Y_n = \frac{X_1 + X_2 + \ldots + X_n}{n},
\]

is chosen to be the test statistic.

2.1. The Bayesian approach. In a Bayesian framework of hypothesis testing we need to specify priors for the null and the alternative hypothesis. In our example the null prior is chosen to be the Dirac measure \( \delta_0 \) in \( \theta = 0 \). In order to reverse the Jeffreys–Lindley paradox, the alternative prior \( p \) is constructed as follows. First we define a sequence \( I_1, I_2, I_3, \ldots \) of subsets of \( \mathbb{R} \) by

\[
I_n = \{ \theta \in \mathbb{R} \mid f^{*n}(\sqrt{n} - n\theta) \geq n^{2n} f^{*n}(\sqrt{n}) \}.
\]
A reverse to the Jeffreys–Lindley paradox

Now, on the one hand, the functions $f^{*n}$ being unbounded and piecewise continuous, the subsets $I_n$ must have non-empty interior, and therefore must be of strictly positive Lebesgue measure. On the other hand, the $f^{*n}$ being integrable, the subsets $I_n$ must be of bounded Lebesgue measure. Altogether one may talk about the uniform distribution on $I_n$ and about its density $p_n$. In terms of these $p_n$, we define the prior $p$ as

$$p(\theta) = \sum_{k=1}^{\infty} 2^{-k} p_k(\theta).$$

From Lebesgue’s convergence theorems (see, for example, [3], [7], [8]) it follows that $p$ is a well-defined density with total probability mass equal to one. With the likelihood given by (2.2) the hypotheses to be tested are

$H_0$: prior is $\delta_0$ against $H_1$: prior is $p$.

Bayesians will base their decision on the Bayes factor, which is in this scenario, for an arbitrary outcome $y$ of $Y_n$, given by

$$BF_n(y) = \frac{\int_{-\infty}^{\infty} f_{Y_n}(y \mid \theta) p(\theta) \, d\theta}{\int_{-\infty}^{\infty} f_{Y_n}(y \mid 0) \, d\theta}.$$

In the above formula, $f_{Y_n}(\bullet \mid \theta)$ stands for the probability density of the variable $Y_n$, given $\theta$. If this Bayes factor exceeds some prescribed threshold, then the Bayesian rejects the hypothesis $H_0$. Suppose now that we observe the outcome

$$\hat{y}_n = \frac{\sigma}{\sqrt{n}}$$

for the sample mean $Y_n$. The value of the density $f_{Y_n}$ in $y_n$ may be expressed (see [3], [5]) in terms of a convolution power of $f$ as

$$f_{Y_n}(y_n \mid \theta) = n f^{*n}(\sigma \sqrt{n} - n\theta).$$

Using this in (2.5), one arrives at

$$BF_n(\hat{y}_n) = \int_{-\infty}^{\infty} f_{Y_n}(\hat{y}_n \mid \theta) p(\theta) \, d\theta = \int_{-\infty}^{\infty} \frac{f^{*n}(\sigma \sqrt{n} - n\theta)}{f^{*n}(\sigma \sqrt{n})} p(\theta) \, d\theta$$

$$= \sum_{k=1}^{\infty} \int_{I_k} \frac{f^{*n}(\sigma \sqrt{n} - n\theta)}{f^{*n}(\sigma \sqrt{n})} 2^{-k} p_k(\theta) \, d\theta$$

$$= \int_{I_n} \frac{f^{*n}(\sigma \sqrt{n} - n\theta)}{f^{*n}(\sigma \sqrt{n})} 2^{-n} p_n(\theta) \, d\theta$$

$$= \int_{I_n} n^2 2^{-n} p_n(\theta) \, d\theta = n.$$
It thus appears that

\[(2.6) \quad \lim_{n \to \infty} BF_n(y_n) = \infty.\]

Thus, asymptotically, the Bayesian will reject \(H_0\) when confronted with outcomes \(y_n\) of the sample mean \(Y_n\).

2.2. The frequentist approach. The frequentist approach is based on the \(p\)-value. If this \(p\)-value is below some prescribed threshold (typically 0.05), then the frequentist will reject the hypothesis \(H_0\). The \(p\)-value is computed starting from the null-hypothesized value for \(\theta\), in our scenario the value \(\theta = 0\). Contrary to the Bayesian approach, the form of the alternative hypothesis \(H_1\) does not play an essential role in the frequentist decision procedure. When testing in a two-sided way, the \(p\)-value in our particular scenario, given the outcome \(y_n = \sigma/\sqrt{n}\) for the sample mean \(Y_n\), would be determined as

\[(2.7) \quad PV_n = 2 \times \{1 - F_{Y_n}(y_n \mid H_0)\},\]

where \(F_{Y_n}(\bullet \mid H_0)\) stands for the cumulative distribution function of \(Y_n\) under \(H_0\).

In order to evaluate the \(p\)-value asymptotically, we define the variable \(Z_n\) as

\[Z_n = \frac{Y_n}{\sigma/\sqrt{n}}.\]

Note that \(Z_n\) is under \(H_0\) precisely the standardization of \(Y_n\). Given the outcome \(y_n\) for the sample mean \(Y_n\) one may rewrite (2.7) as follows in terms of the \(Z_n\):

\[PV_n = 2 \times \{1 - F_{Z_n}(1 \mid H_0)\}.\]

The second moment of the population being finite, the classical central limit theorem may be applied. Thus, denoting the cumulative distribution function of the standard Gaussian distribution by \(\Phi\), the asymptotic \(p\)-value turns out to be

\[(2.8) \quad \lim_{n \to \infty} PV_n = 2 \times \{1 - \Phi(1)\} = 0.32.\]

When observing a \(p\)-value of this size the frequentist will generally decide not to reject \(H_0\). Hence, asymptotically, the frequentist will maintain \(H_0\) when confronted with an outcome \(y_n\) of the sample mean \(Y_n\).

3. CLOSING REMARKS

We have constructed an example in which the Bayesian (basing his decision on (2.6)) will asymptotically reject \(H_0\) whereas the frequentist (basing his decision on (2.8)) will asymptotically not reject this hypothesis. It thus appears that, in the absence of sufficient regularity of likelihood or prior, the Jeffreys–Lindley paradox
may manifest itself in a reversed way. It should be noted that, rather than the specific function $f$ defined by (2.1), any persistently unbounded probability density (see [4]) will lead to a reversion of the paradox.

In the likelihood defined by (2.2), the parameter $\theta$ presents the population mean. For this reason, at first sight, it may seem natural to use its empirical counterpart, the sample mean $Y_n$ defined by (2.3), as the test statistic. However, in the proposed scenario, the sample mean fails to be a sufficient estimator for the parameter $\theta$. Future work needs to be done to construct an example in which the paradox is reversed through a sufficient test statistic.

REFERENCES

[1] H. Jeffreys, *The Theory of Probability*, Oxford University Press, Oxford 1961.
[2] D. V. Lindley, *A statistical paradox*, Biometrika 44 (1957), pp. 187–192.
[3] W. R. Pestman, *Mathematical Statistics*, second edition, De Gruyter Verlag, Berlin 2009.
[4] W. R. Pestman, F. Tuerlinckx, and W. Vanpaemel, *Persistently unbounded probability densities*, Statist. Probab. Lett. 118 (2016), pp. 135–138.
[5] A. Rényi, *Foundations of Probability*, Dover Publications, Inc., New York 2007.
[6] C. P. Robert, *On the Jeffreys–Lindley paradox*, Philos. Sci. 81 (2014), pp. 216–232.
[7] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York 1987.
[8] L. Schwartz, *Mathematics for the Physical Sciences*, Dover Publications, New York 2008.
[9] G. Shafer, *Lindley’s paradox*, J. Amer. Statist. Assoc. 77 (378) (1982), pp. 325–334.
[10] A. Spanos, *Who should be afraid of the Jeffreys–Lindley paradox?*, Philos. Sci. 80 (2013), pp. 73–93.
[11] J. Sprenger, *Testing a precise null hypothesis: The case of Lindley’s paradox*, Philos. Sci. 80 (2013), pp. 733–744.