EUROPEAN CENTRAL BANK

EUROSYSTEM

Working Paper Series

Simone Manganelli

Double conditioning: the hidden connection between Bayesian and classical statistics

No 2786 / February 2023

Abstract

Bayesian decisions are observationally identical to decisions with judgment. Decisions with judgment test whether a judgmental decision is optimal and, in case of rejection, move to the closest boundary of the confidence interval, for a given confidence level. The resulting decisions condition on sample realizations, which are used to construct the confidence interval itself. Bayesian decisions condition on sample realizations twice, with the tested hypothesis and with the choice of the confidence level. The second conditioning reveals that Bayesian decision makers have an *ex ante* confidence level equal to one, which is equivalent to assuming an uncertainty neutral behavior. Robust Bayesian decisions are characterized by an *ex ante* confidence level strictly lower than one and are therefore uncertainty averse.

Keywords: Statistical Decision Theory; Hypothesis Testing; Confidence Intervals; Ambiguity Aversion.

JEL Codes: C1; C11; C12; C13.

Non technical summary

A celebrated property of Bayesian decisions is that they condition on the data. A simple example reveals how Bayesian decisions can be obtained as a special case of *decisions with judgment*, which also condition on the data.

Decisions with judgment start from a judgmental decision and, for a given confidence level, choose the action associated with the closest boundary of the confidence interval of the gradient of the loss function. Bayesian statistics applies Bayes formula to combine prior and likelihood, constructing a posterior distribution which exploits non sample and sample information. Bayesian decisions are obtained by minimizing the expected loss, using the posterior distribution to compute the expectation. In the decision space, this is a convex combination of the judgmental and maximum likelihood decisions, where the judgmental decision corresponds to the no data decision, that is the decision which minimizes the expected loss using the prior distribution. There must therefore exist a confidence interval around the maximum likelihood decision, whose edge coincides with the Bayesian decision. By making explicit the judgmental decision and the choice of the confidence level in a classical setting, it is possible to understand the relationship between Bayesian and classical procedures.

The decision with judgment is at the boundary of the confidence interval. Beyond this boundary, the probability of committing Type I errors becomes greater than the given confidence level. The confidence level reflects the attitude of the decision maker towards statistical uncertainty. Bayesian decision makers are uncertainty neutral, as they reject the judgmental decision implicit in their priors with probability one. Robust Bayesian decision makers are uncertainty averse, as they reject the judgmental decision implicit in their set of priors with probability less than one. In both cases, the Bayesian approach conditions the choice of the confidence level on the data, as it is modified after the sample realization is observed. Suitable experiments may shed light on the degree of uncertainty aversion of decision makers and on how they would revise their confidence levels after the data is revealed.

1 Introduction

A celebrated property of Bayesian decisions is that they condition on the data (Berger (1985)). A simple example reveals how Bayesian decisions can be obtained as a special case of *decisions with judgment*, which also condition on the data. The example is kept deliberately simple, as it is sufficient to deliver the main insight of the paper: the Bayesian priors and posterior updating procedure represents an indirect way to introduce judgment in statistical decision making problems.

Decisions with judgment start from a judgmental decision and, for a given confidence level, test whether it is optimal, conditional on the observed sample realization. The chosen action is the one associated with the closest boundary of the confidence interval of the gradient of the loss function. There is a mapping between the Bayesian prior and posterior on the one hand and the judgmental decision and confidence level on the other hand, which makes the two decisions identical, for any sample realization. This mapping reveals that the confidence level of Bayesian decisions is sample dependent, implying that they condition on the data twice: in the hypothesis to be tested and in the choice of the confidence level. The double conditioning implicit in the Bayesian approach results in decision rules which reject the judgmental decision with probability one, but stop at the boundary of a confidence interval associated with a confidence level less than one.

This introduction first summarizes the decision with judgment of Manganelli (2021), second explains its link with Bayesian decisions, third illustrates the source of double conditioning in the Bayesian analysis, and fourth highlights the connections with the Knightian decision theory of Bewley (2011).

The key ingredient of the decision with judgment is a formal definition of judgment. Judgment is defined as a pair formed by a judgmental decision and a confidence level associated with it. It is used to set up the hypothesis to test whether the judgmental decision is optimal. The test for optimality checks whether the population and sample gradients of the loss function evaluated at the judgmental decision have opposite signs, thereby conditioning on the data. Rejection of the null hypothesis implies that marginal moves away from the judgmental decision in the direction indicated by the sample gradient do not increase the loss function in population with probability lower than the confidence level. The resulting decision with judgment is either the judgmental decision itself or is at the boundary of the confidence interval of the sample gradient of the loss function. The confidence level determines the probability of committing Type I errors.

The decision with judgment is admissible, a result that already hints at a deeper relationship with Bayesian decisions, since by the complete class theorem all admissible decisions are generalized Bayes rules. To understand the link, consider that the judgmental decision coincides with the Bayesian decision implied by the prior distribution. Since the confidence level determines the width of the confidence interval and the decision with judgment moves from the judgmental decision to the closest boundary of the confidence interval, the confidence level determines the deviation from the judgmental decision, resulting in a convex combination between the judgmental and the maximum likelihood decisions. Similarly, the Bayesian posterior updating shrinks from the prior decision to the decision implied by the likelihood of the data.

By exploiting this relationship, it is possible to obtain a mapping between the sample realization (*p-value*) and the confidence level which makes the Bayesian and the decision with judgment observationally identical. It is this mapping which reveals the source of double conditioning in the Bayesian analysis. The Bayesian decision conditions on the data not only in the choice of the hypothesis to be tested, but also in the choice of the confidence level. It is shown that Bayesian decisions always reject the null hypothesis of optimality with probability one, implying an *ex ante* confidence level equal to one. Conditional on the data, the *ex post* confidence level is revised to be strictly less than one and the resulting decision can be interpreted as that associated with the corresponding confidence interval.

Robust Bayesian decision are characterised by *ex ante* confidence level strictly less than one, so that they do not always reject the null hypothesis that the judgmental decision is optimal. These results are reminiscent of the insights of Bewley (2011), who develops a theory of decisions with an inertia assumption. His intuition is that there is a *status quo* decision, which is abandoned only if there is another decision strictly preferred to it. That is similar to the idea of starting with a judgmental decision and abandoning it only if there is sufficient statistical evidence against it. As shown in the last section of this paper, classical confidence intervals can be interpreted as a set of posterior means corresponding to a set of prior distributions, which are used to construct the robust Bayesian decision. The *ex ante* confidence level defines the degree of uncertainty aversion in the sense of Bewley (2011). When the *ex ante* confidence level is equal to one, the judgmental (or status quo) decision is always rejected and corresponds to an uncertainty neutral (or Bayesian) preference ordering. *Ex ante* confidence levels strictly less than one define an uncertainty averse decision maker, with lower values implying higher degrees of aversion.

The paper is structured as follows. Section 2 defines the Bayesian decision and the decision with judgment. Section 3 establishes the link between the two decisions, highlights the source of double conditioning and illustrates its implications. Section 4 extends the analysis to the case of robust Bayesian decisions. Section 5 concludes.

2 Statistical Decision Rules

This section derives the Bayesian decision and the decision with judgment, under the following simplified setting.

Definition 2.1 (Decision Environment). *The decision environment is characterized by:*

1. $X \sim N(\theta, 1)$, *for an unknown* $\theta \in \mathbb{R}$.
2. *The sample realization* $x \in \mathbb{R}$ *is observed.*
3. $a \in \mathbb{R}$ *denotes the action of the decision maker.*
4. *The decision maker minimizes the loss function* $L(\theta, a) = -a\theta + 0.5a^2$.

2.1 The Bayesian Decision

The Bayesian approach assumes that the decision maker uses subjective information in the form of a prior distribution over the unknown parameter θ.

Proposition 2.1 (**Bayesian decision**). *Consider the decision environment of Definition 2.1. Bayesian decision makers with prior distribution $N(0,1)$ over θ choose the action:*

$$\delta^{\pi^*}(x) = x/2 \tag{1}$$

Proof — See Appendix.

2.2 The Decision with Judgment

The decision rule incorporating judgment of Manganelli (2021) assumes that the decision maker starts from a judgmental decision $\tilde{a}$ and uses the empirical gradient of the loss function to test whether such decision is optimal, for a given confidence level α.

Judgment, which is formed by the pair $A = \{\tilde{a}, \alpha\}$, is a primitive to the decision problem, like the Bayesian priors. $\tilde{a}$ is the decision that would be taken without statistical analysis. It is analogous to the status quo decision in the theory of Bewley (2011). The confidence level α determines the amount of statistical evidence needed to abandon $\tilde{a}$.

Given the judgment, the statistician can formally test whether the gradient evaluated at the judgmental decision is equal to zero, a necessary and sufficient condition for optimality in the decision environment of Definition 2.1. If the null hypothesis is not rejected, the decision maker chooses the judgmental decision $\tilde{a}$. If the null hypothesis is rejected, the decision maker selects the action that sets the empirical gradient equal to the closest boundary of the confidence interval. The rationale is that actions closer to $\tilde{a}$ would be rejected, while actions further away would be wrongly rejected, for the given confidence level α.

Proposition 2.2 (**Decision with judgment**). *Consider the decision environment of Def-*

inition 2.1. Decision makers with judgment $A = \{0, \alpha\}$ choose the action:

$$\delta^A(x) = \hat{\lambda} x \tag{2}$$

where $\hat{\lambda} = \max\{0, \lambda^*\}$ with $\lambda^* = 1 + c_{\alpha/2}/|x|$ if $x \neq 0$, and $\hat{\lambda} \in [0, 1]$ otherwise.

Proof — See Appendix.

3 Connection Between Bayesian and Classical Decisions

The link between Bayesian decisions and decisions with judgment is given by the link between priors, posteriors and judgment $A \equiv \{\tilde{a}, \alpha\}$. To establish the formal connection, it is necessary to have a more general definition of confidence level.

Definition 3.1 (**Confidence level**). *The confidence level is defined by the pair $\{\bar{\alpha}, \alpha(x)\}$. The* ex ante *confidence level $\bar{\alpha}$ is a probability over the unit interval* $[0, 1]$. *The* ex post *confidence level $\alpha(x)$ is a mapping from $\tilde{\alpha}(x)$ to the unit interval:*

$$\alpha(x) : \tilde{\alpha}(x) \to [0, 1]$$

where $\tilde{\alpha}(x)$ denotes the p-value *associated with the sample realization x, under the null hypothesis that the judgmental decision $\tilde{a}$ is optimal.*

Decision makers enter the decision problem with an *ex ante* confidence level $\bar{\alpha}$. Upon observing the sample realization x (and therefore the *p-value*), the confidence level is revised according to the mapping $\alpha(x)$. Note that also this mapping may be decided *ex ante*, but its value is determined only *ex post*, after the realization x is observed. The choice of the confidence level is entirely subjective and may be elicited through suitable experiments. It will be clear in the next section that an *ex ante* confidence level $\bar{\alpha} = 1$ corresponds to an uncertainty neutral and $\bar{\alpha} < 1$ to an uncertainty averse decision maker, in the sense of Bewley (2011).

The following proposition shows that for the Bayesian decision associated with a normal prior there is a corresponding choice of $\tilde{a}$ and $\{\bar{\alpha}, \alpha(x)\}$ which produces an observationally equivalent decision with judgment.

Proposition 3.1 (**Relationship between Bayesian Decisions and Decisions with Judgment**). *Consider the decision environment of Definition 2.1. The Bayesian decision with prior distribution $N(0,1)$ over θ is equal to the decision with judgment when the judgmental decision is $\tilde{a} = 0$, the* ex ante *confidence level is $\bar{\alpha} = 1$ and the* ex post *confidence level is:*

$$\alpha(x) = 2\Phi[\Phi^{-1}(\tilde{\alpha}(x)/2)/2] \tag{3}$$

where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution.

Proof — See Appendix.

Bayesian decisions with a standard normal prior are characterized by an *ex ante* confidence level $\bar{\alpha} = 1$ and an *ex post* confidence level $\alpha(x) < 1$ given in expression (3). This implies that the probability that Bayesian decision rules perform worse than $\tilde{a} = 0$, when it is optimal, cannot be bounded away from 1, or in other words Bayesian decisions have no control on Type I errors.

Decisions with judgment, which include Bayesian decisions as a special case, are admissible.

Proposition 3.2 (**Admissibility of Decisions with Judgment**). *Consider the judgment $A \equiv \{0, \{\bar{\alpha}, \alpha(x)\}\}$. The decision rule with judgment $\delta^A(X)$ in (2) is admissible.*

Proof — See Appendix.

3.1 Implications of Bayesian Decisions

Bayesian decisions condition on the sample realization x twice: first in the hypothesis testing procedure and second in the choice of the confidence level. This is the major insight from the previous analysis, that provides the connection between Bayesian and classical statistics.

Let us start from the hypothesis testing procedure. Interpreting the Bayesian decision as a decision with judgment, the Bayesian decision first tests whether the judgmental decision implicit in the prior, $\tilde{a} = 0$, is optimal given the *ex ante* confidence level $\bar{\alpha}$. Upon rejection, which happens with probability one since $\bar{\alpha} = 1$, it selects the action which sets the gradient of the loss function equal to the closest boundary of the $(1 - \alpha(x))$ *ex post* confidence interval. As explained in Manganelli (2021) and highlighted in the proof of Proposition 2.2, this procedure conditions on the sample realization.

Let us apply his reasoning to the example of this paper. Under the conditions of Proposition 3.1, the Bayesian decision is $\lambda^* x$, where λ^* is given by Proposition 2.2 and $\alpha(x)$ is given by Proposition 3.1. This is the closest decision to the judgmental decision $\tilde{a} = 0$ which is not rejected at the confidence level $\alpha(x)$. The null hypothesis being tested is:

$$H_0 : \nabla_\lambda L(\theta, \lambda^* x) = 0 \quad \Rightarrow \quad \theta = \lambda^* x$$

Notice that the hypothesis conditions on x, which is observed at the time of the decision. The test statistic is obtained by substituting θ with its maximum likelihood estimator, which in the current example is Y, a random variable with the same distribution as X.

The frequentist interpretation of hypothesis testing for Bayesian decision rules can be understood by performing the following thought experiment. To any sample realization x corresponds a confidence level $\alpha(x)$ which maps the Bayesian decision into the decision rule $\lambda^* x$ and is used to test the null hypothesis $H_0 : \theta = \lambda^* x$. Conditional on the observed realization x, one can imagine of drawing a sample $\{y^j\}_{j=1}^J$ from $Y \sim N(\lambda^* x, 1)$. The decision $\lambda^* x$ has the property that, as $J \to \infty$, it is wrongly rejected $\alpha(x)$ of the times when it is optimal:

$$
\begin{aligned}
\alpha(x) &= 2P_{\lambda^* x}(\nabla_\lambda L(Y, \lambda^* x) < \nabla_\lambda L(x, \lambda^* x)) \\
&= 2P_{\lambda^* x}(-x(Y - \lambda^* x) < -x^2 + \lambda^* x^2) \\
&= 2P_{\lambda^* x}(-x(Y - \lambda^* x) < c_{\alpha(x)/2}|x|) \\
&= P_{\lambda^* x}(-sgn(x)(Y - \lambda^* x) < c_{\alpha(x)}) \\
&= \lim_{J\to\infty} J^{-1} \sum_{j=1}^{J} I(-sgn(x)(y^j - \lambda^* x) < c_{\alpha(x)})
\end{aligned}
$$

where $I(\cdot)$ is the indicator function.

The second conditioning of the Bayesian decision is in the choice of $\alpha(x)$. As shown in Proposition 3.1, the null hypothesis that $\tilde{a} = 0$ is optimal is rejected with probability 1, implying an *ex ante* confidence level $\bar{\alpha} = 1$. After seeing the data, however, Bayesian decision makers revise their confidence level to $\alpha(x) < 1$ and the Bayesian decision is determined by the boundary of the $(1 - \alpha(x))$ confidence interval.

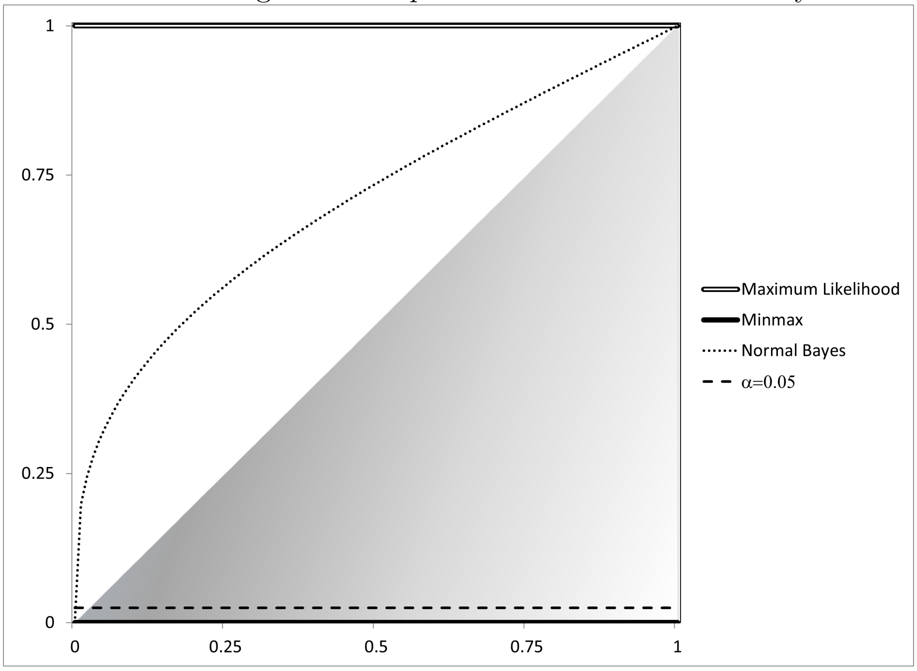
Figure 1 gives a graphical representation in the space $(\alpha(x), \tilde{\alpha}(x))$ of the *ex post* confidence levels associated with Bayesian decisions and with decisions with constant confidence level. The mapping for the Bayesian decisions is given by equation (3).

All statistical decision rules have a confidence level mapping which falls between the two extreme decision behaviors: the minmax decision (which corresponds to the decision with $\alpha(x) = 0$, where no data is taken into consideration) and the maximum likelihood decision (which corresponds to the decision with $\alpha(x) = 1$, where no judgment is taken into consideration).

The shaded area represents the combination of points for which the *p-value* $\tilde{\alpha}(x)$ associated with the judgmental decision $\tilde{a}$ is greater than the chosen confidence level $\alpha(x)$. For all these points the decision taken is the judgmental decision itself, because in this case the null hypothesis that $\tilde{a}$ is optimal is not rejected.

When the *ex post* confidence level associated with the decision with judgment is not sample dependent (in the example of Figure 1 reported as the flat line with $\alpha(x) = 0.05$),

Figure 1: *Ex post* confidence levels for Bayesian decisions



Note: The horizontal axis reports the *p-value* $\tilde{\alpha}(x)$ of the gradient evaluated at the observed realization x and at the judgmental decision $\tilde{a} = 0$. The vertical axis is the chosen confidence level $\alpha(x)$. The shaded area represents the points where the null hypothesis $H_0 : \theta = \tilde{a}$ is not rejected, as in this area the *p-value* is greater than $\alpha(x)$. The figure plots the mapping corresponding to four alternative decision rules: Maximum likelihood (obtained by setting $\alpha(x) = 1$), the minmax (obtained by setting $\alpha(x) = 0$), the decision with judgment with constant $\alpha(x) = 0.05$, and the Gaussian Bayesian decisions are based on priors with zero mean and unit variance (obtained by setting $\alpha(x)$ as given by equation (3)).

by the probability integral transform theorem, the *ex ante* probability of observing *p-values* smaller than the confidence level under the null hypothesis $H_0 : \theta = \tilde{a}$ is equal to the *ex post* confidence level:

$$\bar{\alpha} = P_{\tilde{a}}(\tilde{\alpha}(X) < \bar{\alpha}) = 0.05$$

The line associated with the Bayesian decision with a normal prior, instead, is always higher than the *p-value*, implying that the null hypothesis $H_0 : \theta = \tilde{a}$ is always rejected, as shown in Proposition 3.1:

$$\bar{\alpha} = P_{\tilde{a}}(\tilde{\alpha}(X) < \alpha(X)) = 1$$

The Bayesian approach reveals the possibility for decision makers to revise their confidence level after seeing the data. Such possibility could be tested by suitable experiments, with questions aimed at eliciting *ex ante* and *ex post* confidence levels. Two implications of the normal Bayesian priors seem dubious from a real world decision making perspective. First, they impose a confidence level equal to one on the initial judgmental decision associated with the prior. Decision makers have often strong views about the actions to be taken and substantial empirical evidence is required before abandoning those views. Second, Figure 1 reveals that the lower the *p-value* the lower the confidence level: Bayesian decision makers endowed with a normal prior increase their confidence in the judgmental decision (by choosing a smaller confidence level) when confronted with sample realizations which are widely inconsistent with it.

4 Connection Between Robust Bayesian and Classical Decisions

The previous discussion has highlighted an important feature of Bayesian decisions, namely that they reject the judgmental decision implicit in the prior with probability one. This issue has been addressed by the literature on ambiguity aversion, by considering classes of priors, instead of a single prior (see Gilboa and Marinacci (2013) for a review). Gilboa and Schmeidler (1989) have shown that an ambiguity averse decision maker characterized by a set of priors Γ minimizes the expected loss using the worst possible prior from the set Γ. Epstein and Schneider (2003) have axiomatized the decision in an intertemporal context, leading to a prior-by-prior Bayesian updating as the updating rule for such sets of priors. An application with an analogous updating mechanism is provided by Giacomini and Kitagawa (2021) in the context of set-identified vector autoregressive models. Bewley (2011) and his Knightian decision theory with a *status quo* assumption is the contribution which is closest to the theory of decision with judgment. He argues that classical confidence regions can be interpreted as the set of posterior means corresponding to a set of priors that define the ambiguity aversion of decision makers. This paper shows that decisions with judgment can be interpreted as Bayesian decisions when the *ex ante* confidence level is equal to one and as robust Bayesian decisions when the *ex ante* confidence level is strictly less than one.

In the simple example discussed in this paper, the ambiguity averse framework can be characterized as follows.

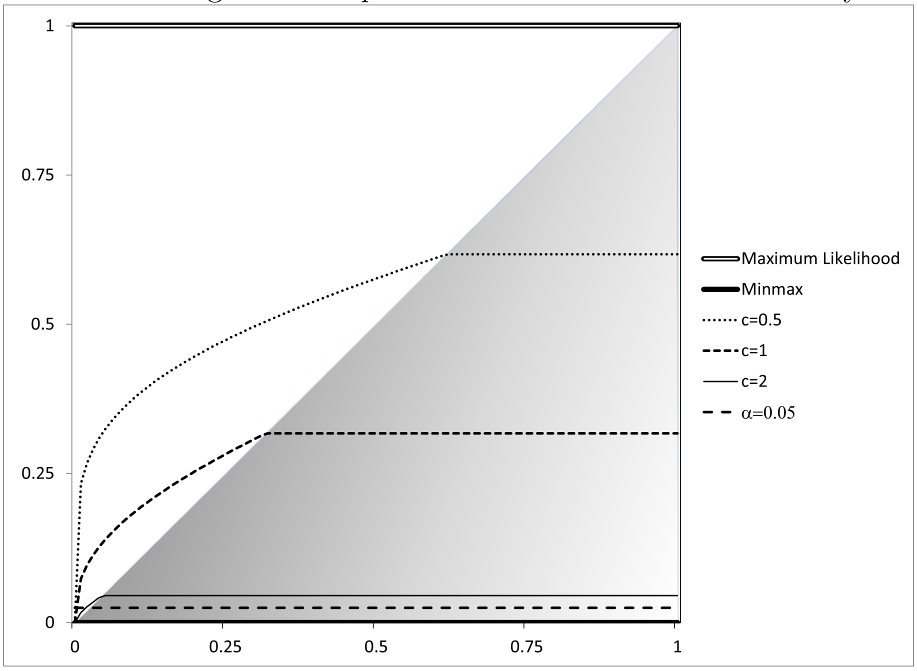
Proposition 4.1 (**Ambiguity Averse Decisions**). *Consider the decision environment of Definition 2.1. An ambiguity averse decision maker with the set of priors* $\Gamma = \{\mu : \mu \text{ is } N(\pi, c^{-1}), \quad \pi, c \in \mathbb{R}, \quad c > 0, \quad -1 \leq \pi \leq 1\}$ *over* θ *chooses the action:*

$$\delta^{\Gamma}(x) = \begin{cases} (1+c)^{-1}(x+c) & \text{if } x < -c \\ 0 & \text{if } -c \leq x \leq c \\ (1+c)^{-1}(x-c) & \text{if } x > c \end{cases} \tag{4}$$

This action is equal to a decision with judgment when the judgmental decision is $\tilde{a} = 0$*, the* ex ante *confidence level is* $\bar{\alpha} = 2\phi(-c)$ *and the* ex post *confidence level is:*

$$\alpha(x) = 2\phi\left(-c(1+c)^{-1}(1-\phi^{-1}(\tilde{\alpha}(x)/2))\right) \tag{5}$$

Figure 2: *Ex post* confidence levels for robust Bayesian decisions



Note: The horizontal axis reports the *p-value* $\tilde{\alpha}(x)$ of the gradient evaluated at the observed realization x and at the judgmental decision $\tilde{a} = 0$. The vertical axis is the *ex post* confidence level $\alpha(x)$. The shaded area represents the points where the null hypothesis $H_0 : \theta = \tilde{a}$ is not rejected, as in this area the *p-value* is greater than $\alpha(x)$. The figure plots the confidence level mappings corresponding to an ambiguity averse decision maker, who chooses from the set of priors $\Gamma = \{\mu : \mu \text{ is } N(\pi, c^{-1}), -1 \leq \pi \leq 1\}$, for different levels of precision c. For comparison, the figure also report the three decisions with constant α from Figure 1.

where $\Phi(\cdot)$ *denotes the cdf of the standard normal distribution.*

Proof — See Appendix.

The *ex post* confidence level associated with decision (4) is reported in Figure 2 for $c = \{0.5, 1, 2\}$. If the sample realization falls within the interval $(-c, c)$, ambiguity averse decision makers retain their judgmental decision $\tilde{a} = 0$. This is like the situation when, in the decision with judgment, the test fails to reject the null hypothesis, that is, when the test statistic $-x + \tilde{a}$ falls within the frequentist confidence interval. In such a case, the decision with judgment coincides with the judgmental decision $\tilde{a}$. This is consistent with the result of Bewley (2011) that classical confidence regions correspond the set of prior means of an uncertainty averse decision maker.

The null hypothesis $H_0 : \theta = \tilde{a}$ in the case of the ambiguity averse decision rule of

Figure 2 is rejected with *ex ante* probability less than one:

$$\bar{\alpha} = P_{\tilde{a}}(\tilde{\alpha}(X) < \alpha(X)) < 1$$

The exact probability depends on the coefficient c, as shown in Proposition 4.1. It can be read on the horizontal axis of Figure 2, at the point where the confidence level mappings cross the diagonal of the unit square.

Since $\alpha(x) < \bar{\alpha}$ for $\tilde{\alpha}(x) < \bar{\alpha}$, the judgmental decision $\tilde{a}$ is wrongly rejected when correct with probability $\bar{\alpha}$, but in case of rejection the ambiguity averse decision stops at the boundary of the $(1 - \alpha(x))$ confidence interval. Notice, however, that also in this case the *ex post* confidence level $\alpha(x)$ used to construct the decision is sample dependent and eventually decreasing with lower *p-values*, so it shares one of the dubious implications discussed in the context of the normal Bayesian decision.

5 Conclusion

Bayesian statistics applies Bayes formula to combine prior and likelihood, constructing a posterior distribution which exploits non sample and sample information. Bayesian decisions are obtained by minimizing the expected loss, using the posterior distribution to compute the expectation. In the decision space, this is a convex combination of the judgmental and maximum likelihood decisions, where the judgmental decision corresponds to the no data decision, that is the decision which minimizes the expected loss using the prior distribution. There must therefore exist a confidence interval around the maximum likelihood decision, whose edge coincides with the Bayesian decision. By making explicit the judgmental decision and the choice of the confidence level in a classical setting, it is possible to understand the relationship between Bayesian and classical procedures.

The decision with judgment is at the boundary of the confidence interval. Beyond this boundary, the probability of committing Type I errors becomes greater than the given confidence level. The confidence level reflects the attitude of the decision maker towards statistical uncertainty. Bayesian decision makers are uncertainty neutral, as they reject the judgmental decision implicit in their priors with probability one. Robust Bayesian decision makers are uncertainty averse, as they reject the judgmental decision implicit in their set of priors with probability less than one. In both cases, the Bayesian approach conditions the choice of the confidence level on the data, as it is modified after the sample realization is observed. Suitable experiments may shed light on the degree of uncertainty aversion of decision makers and on how they would revise their confidence levels after the data is revealed.

Appendix — Proofs

Proof of Proposition 2.1 (Bayesian Decision) — Bayesian decision makers minimize the expected loss function, using the posterior distribution of θ to compute the expectation. The decision solves the minimization problem:

$$\begin{aligned}\min_a E^{\pi^*}[L(\theta, a)] &= -aE^{\pi^*}(\theta) + 0.5a^2 \\ &= -ax/2 + 0.5a^2\end{aligned}$$

where π^* denotes the posterior distribution of θ after observing x, which has mean $x/2$. Solving for the first order conditions, gives the result. $\square$

Proof of Proposition 2.2 (Decision with judgment) — Let $a(\lambda; x) \equiv \lambda x$, for $\lambda \in [0, 1]$, the action that shrinks from the judgmental decision $\tilde{a} = 0$ towards the maximum likelihood decision x.

If $x = 0$, the judgmental decision coincides with the maximum likelihood estimate, and the optimal action is $\delta^A(x) = 0$, which trivially holds for any λ. If $x \neq 0$, the gradient of the loss function with respect to λ is:

$$\begin{aligned} Z(\theta, \lambda; x) &\equiv \nabla_\lambda L(\theta, a(\lambda; x)) \\ &= -x\theta + \lambda x^2 \end{aligned}$$

The maximum likelihood estimator of the gradient is:

$$\begin{aligned} Z(\hat{\theta}(Y), \lambda; x) &= Z(Y, \lambda; x) \\ &= -xY + \lambda x^2 \end{aligned}$$

where Y has the same distribution as X.

Notice that since $Z(x, \lambda; x) < 0$ for $\lambda < 1$, the hypothesis to be tested is:

$$H_0 : Z(\theta, \lambda; x) \geq 0 \qquad vs \qquad H_1 : Z(\theta, \lambda; x) < 0$$

Since the null hypothesis is rejected if and only if $H_0 : \theta = \lambda x$ is rejected, the test statistic under the null hypothesis satisfies the following property:

$$\begin{aligned} P_{\lambda x}(Z(Y, \lambda; x) < Z(x, \lambda; x) | Z(Y, \lambda; x) < 0) = \\ &= 2P_{\lambda x}(-x(Y - \lambda x) < -x(x - \lambda x)) \\ &= 2P_0(-x\mathcal{X} < -x(x - \lambda x)) \\ &= I(x > 0) 2P_0(\mathcal{X} > |x|(1 - \lambda)) + I(x < 0) 2P_0(\mathcal{X} < -|x|(1 - \lambda)) \\ &= I(x \neq 0) 2P_0(\mathcal{X} < -|x|(1 - \lambda)) \end{aligned}$$

where $\mathcal{X} \sim N(0, 1)$.

The null hypothesis is rejected if $P_0(\mathcal{X} < -|x|(1 - \lambda)) < \alpha/2$. If for $\lambda = 0$ it is not rejected, the judgmental decision is retained as optimal decision. In case of rejection, the optimal decision is the one at the boundary of the confidence level:

$$P_0(\mathcal{X} < -|x|(1-\lambda^*)) = \alpha/2$$

resulting in $\lambda^* = 1 + c_{\alpha/2}/|x|$.

In fact, any $\lambda \neq \lambda^*$ would not be compatible with the given judgement, because the null would be rejected for $\lambda < \lambda^*$ and would be wrongly rejected for $\lambda > \lambda^*$. $\square$

Proof of Proposition 3.1 (Relationship between Bayesian Decisions and Decisions with Judgment) — The judgmental decision associated with the prior is obtained by minimizing $E^\pi[L(\theta, a)]$ with respect to a, where π denotes the standard normal prior distribution and the superscript that the expectation is taken with respect to this prior:

$$\tilde{a} = \arg\min_a[-aE^\pi(\theta) + 0.5a^2]$$

Since under the standard normal $E^\pi(\theta) = 0$, the action that minimizes the expected loss is $\tilde{a} = 0$.

The two decisions are equal if the null hypothesis that $\tilde{a}$ is optimal is rejected and when $\delta^{\pi^*}(x) = \lambda^* x$, that is when $x/2 = (1 + c_{\alpha(x)/2}/|x|)x$. Solving for $\alpha(x)$ gives $\alpha(x) = 2\Phi(-|x|/2)$. Since the *p-value* associated with the judgmental decision $\tilde{a} = 0$ is $\tilde{\alpha}(x) = 2\Phi(-|x|)$, solving for $-|x|$ and substituting in the expression for $\alpha(x)$ gives the result.

Finally, to show that the *ex ante* confidence level is $\bar{\alpha} = 1$, note that the null hypothesis that $\tilde{a}$ is optimal is rejected when the p-value is lower than the chosen confidence level. Using the previous result:

$$\begin{aligned}
P_{\tilde{a}}[\tilde{\alpha}(X) < \alpha(X)] &= P_{\tilde{a}}[\tilde{\alpha}(X) < 2\Phi(\Phi^{-1}(\tilde{\alpha}(X)/2)/2)] \\
&= P_{\tilde{a}}[\Phi^{-1}(\tilde{\alpha}(X)/2) < \Phi^{-1}(\tilde{\alpha}(X)/2)/2] \\
&= P_{\tilde{a}}[1 > 1/2] \\
&= 1
\end{aligned}$$

where the inequality changes sign because $\Phi^{-1}(\tilde{\alpha}(X)/2)$ is negative. $\square$

Proof of Proposition 3.2 (Admissibility of Decisions with Judgment) — To prove admissibility, one has to recognize that there are two random variables involved in the statistical decision problem. First, conditional on the sample realization x, there is the random variable that determines whether the null hypothesis of optimality of a given decision is accepted or rejected. Second, there is the random variable that determines the sample realization x.

For given x, it is possible to test the following hypotheses for any $\lambda \in [0, 1]$:

$$H_0 : Z(\theta, \lambda; x) \geq 0 \qquad H_1 : Z(\theta, \lambda; x) < 0$$

where $Z(\theta, \lambda; x) \equiv \nabla_\lambda L(\theta, a(\lambda; x)) = -x(\theta - \lambda x)$. This is equivalent to the following hypotheses:

$$\begin{array}{lll} \text{if } x < 0 & H_0 : \theta \geq \lambda x & H_1 : \theta < \lambda x \\ \text{if } x > 0 & H_0 : \theta \leq \lambda x & H_1 : \theta > \lambda x \end{array}$$

To construct the test function, substitute θ with the maximum likelihood estimator Y, where to avoid confusion between the potential realization of X and the observed one, x, I have replaced X with the identically distributed random variable $Y \sim N(\theta, 1)$. The test statistic is:

$$\begin{aligned} Z(Y, \lambda; x) &= -x(Y - \lambda x) \\ &= I(x < 0)|x|(Y - \lambda x) - I(x > 0)|x|(Y - \lambda x) \end{aligned}$$

If $x = 0$, the maximum likelihood estimate coincides with the judgmental decision, which is therefore not rejected and retained as decision. In the other cases, the decision rule is given by the following test function:[1]

$$\text{if } x < 0 \qquad \Psi^-(y,\lambda;x) = \begin{cases} 0 & \text{if } y-\lambda x > c_{\alpha/2} \\ \gamma & \text{if } y-\lambda x = c_{\alpha/2} \\ 1 & \text{if } y-\lambda x < c_{\alpha/2} \end{cases}$$

$$\text{if } x > 0 \qquad \Psi^+(y,\lambda;x) = \begin{cases} 0 & \text{if } y-\lambda x < -c_{\alpha/2} \\ \gamma & \text{if } y-\lambda x = -c_{\alpha/2} \\ 1 & \text{if } y-\lambda x > -c_{\alpha/2} \end{cases}$$

since H_0 is rejected if and only if $H_0 : \theta = \lambda x$ is rejected, which implies $Y - \lambda x \sim N(0,1)$.

Adopting the positive linear transformation $L^* = \varepsilon^{-1} L$ for $\varepsilon > 0$, denoting with $(\lambda + \varepsilon) + x$ the action taken in case of rejection, the risk function of a test Ψ^* is:

$$\begin{aligned} R(\theta,\Psi^*;x) &= \int \varepsilon^{-1} \left[L(\theta,\lambda x)(1-\Psi^*(y,\lambda;x)) + L(\theta,(\lambda+\varepsilon)x)\Psi^*(y,\lambda;x) \right] dF(y) \\ &= \varepsilon^{-1} L(\theta,\lambda x) + \int \varepsilon^{-1} \left[L(\theta,(\lambda+\varepsilon)x) - L(\theta,\lambda x) \right] \Psi^*(y,\lambda;x) dF(y) \end{aligned}$$

Following the reasoning of Section 8.3 of Berger (1985) and considering only the case when $x < 0$ (the other case can be proven in a similar way):

$$\begin{aligned} R(\theta,\Psi^*;x) - R(\theta,\Psi^-;x) &\sim \\ &\sim \int Z(\theta,\lambda x)[\Psi^*(y,\lambda;x) - \Psi^-(y,\lambda;x)] dF(y) \quad \text{for sufficiently small } \varepsilon \\ &= |x|(\theta-\lambda x) E_\theta[\Psi^*(Y,\lambda;x) - \Psi^-(Y,\lambda;x)] \end{aligned}$$

Since Ψ^- is a uniformly most powerful test, $E_\theta[\Psi^*(Y,\lambda;x) - \Psi^-(Y,\lambda;x)] \leq 0$ for $\theta < \lambda x$. By symmetry, $1-\Psi^-$ is a uniformly most powerful test of size $1-\alpha$ for testing $H_0 : \theta \leq \lambda x$ versus $H_1 : \theta > \lambda x$, and $E_\theta[(1-\Psi^*(Y,\lambda;x)) - (1-\Psi^-(Y,\lambda;x))] \leq 0$ for $\theta > \lambda x$. This

[1] Note that the *ex ante* confidence level $\bar{\alpha}$ is used for $\lambda = 0$ and the *ex post* confidence level $\alpha(x)$ for $\lambda > 0$. For simplicity, I use the notation α to denote both cases.

implies that $(\theta - \lambda x)E_\theta[\Psi^*(Y,\lambda;x) - \Psi^-(Y,\lambda;x)] \geq 0$ and that $R(\theta,\Psi^*;x) - R(\theta,\Psi;x) \geq 0$ for all θ. The test Ψ^- is therefore admissible, for any x and λ.

Suppose now that the decision rule $\delta^A(x)$ is not admissible. Then it exists another decision rule $\delta^*(x)$ such that (again considering only the case $x < 0$ for notation simplicity):

$$
\begin{aligned}
R(\theta,\delta^*) - R(\theta,\delta^A) &= \int \left[R(\theta,\Psi^*;x) - R(\theta,\Psi^-;x)\right] dF(x) \\
&\sim \int\int \left[|x|(\theta - \hat{\lambda}x)[\Psi^*(y,\hat{\lambda};x) - \Psi^-(y,\hat{\lambda};x)]\right] dF(y)dF(x) \\
&= \int\int \left[|x|(\theta - \hat{\lambda}x)[\Psi^*(y,\hat{\lambda};x) - \Psi^-(y,\hat{\lambda};x)]\right] dF(x)dF(y) \\
&= \int \left[|\bar{x}|(\theta - \hat{\lambda}\bar{x})[\Psi^*(y,\hat{\lambda};\bar{x}) - \Psi^-(y,\hat{\lambda};\bar{x})]\right] dF(y) \\
&\leq 0
\end{aligned}
$$

for some finite $\bar{x} \in \mathbb{R}$, with strict inequality for at least one θ. But this contradicts the fact that Ψ^- is admissible. Note that any alternative decision rule has to select $\hat{\lambda}$ to be consistent with the given judgment. Note also that the dependence of $\hat{\lambda}$ and α on x has been left implicit and does not affect the reasoning. □

Proof of Proposition 4.1 (Ambiguity Averse Decisions) — An ambiguity averse decision maker chooses the action:

$$\arg\min_a \max_{\mu\in\Gamma} \int L(\theta,a)dF^{\mu^*}(\theta|x)$$

where $F^{\mu^*}(\theta|x)$ denotes the posterior updating of a given μ distribution in Γ.

For a prior $N(\pi, c^{-1})$, the posterior mean of θ is $(1+c)^{-1}(x+c\pi)$. Since the loss function is linear in π:

$$
\begin{aligned}
&\max_{\mu\in\Gamma} \int (-a\theta + 0.5a^2)dF^{\mu^*}(\theta|x) = \\
&= -a[I(a<0) - I(a>0) + \pi I(a=0)](1+c)^{-1}c - a(1+c)^{-1}x + 0.5a^2
\end{aligned}
$$

Solving for the first order conditions with respect to a:

$$a = \frac{x + c[I(a < 0) - I(a > 0) + \pi I(a = 0)]}{1 + c}$$

The result is obtained by noting that $a < 0$ if $x + c < 0$, $a > 0$ if $x - c > 0$, and $a = 0$ if $-c < x < c$ and $-1 \leq \pi = -x/c \leq 1$.

The derivation of the judgmental decision $\tilde{a}$ is similar, but now the prior distribution has expected value equal to π.

For the derivation of the confidence levels, first notice that the ambiguity averse decision and the decision with judgment can be rewritten as:

$$\delta^{\Gamma}(x) = I(|x| > c)(x - sgn(x)c)/(1 + c)$$
$$\delta^{A}(x) = I(|x| > -c_{\bar{\alpha}/2})(x + sgn(x)c_{\alpha(x)/2})$$

The two decisions are equal when $c = -c_{\bar{\alpha}/2}$ and $(x - sgn(x)c)/(1 + c) = x + sgn(x)c_{\alpha(x)/2}$. Inverting the first condition gives the *ex ante* confidence level $\bar{\alpha}$. Solving the second condition for $\alpha(x)$ gives the *ex post* confidence level. □

References

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis.* New York: Springer-Verlag.

Bewley, T. (2011). Knightian decision theory and econometric inferences. *Journal of Economic Theory 146*, 1134–1147.

Epstein, L. and M. Schneider (2003). Recursive multiple-priors. *Journal of Economic Theory 113*(1), 1–31.

Giacomini, R. and T. Kitagawa (2021). Robust Bayesian inference for set-identified models. *Econometrica 89*(4), 1519–1556.

Gilboa, I. and M. Marinacci (2013). Ambiguity and the Bayesian paradigm. In D. Acemoglu, M. Arellano, and E. Dekel (Eds.), *Advances in Economics and Econometrics: Theory and Applications*. New York, Cambridge University Press.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

Manganelli, S. (2021). Statistical decision functions with judgment. *ECB Working Paper Series 2512*, 1–23.

Acknowledgements

The views expressed in this paper are my own and do not necessarily reflect those of the ECB.

Simone Manganelli

European Central Bank, Frankfurt am Main, Germany; email: simone.manganelli@ecb.int

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

PDF ISBN 978-92-899-5982-7 ISSN 1725-2806 doi:10.2866/67554 QB-AR-23-023-EN-N