Bounds for expected loss in Bayesian decision theory with imprecise prior probabilities

Citation for published version (APA):

Document status and date:
Published: 01/01/1992

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights. 

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computing Science

Memorandum COSOR 92-40

Bounds for expected loss in Bayesian decision theory with imprecise prior probabilities

F.P.A. Coolen

Eindhoven, September 1992
The Netherlands
BOUND FOR EXPECTED LOSS IN BAYESIAN DECISION THEORY
WITH IMPRECISE PRIOR PROBABILITIES

F.P.A. Coolen
Eindhoven University of Technology
Department of Mathematics and Computing Science
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Abstract
Classical Bayesian inference uses the expected value of a loss function with
regard to a single prior distribution for a parameter to compare decisions,
and an optimal decision minimizes the expected loss. Recently interest has
grown in generalizations of this framework without specified priors, to
allow imprecise prior probabilities. Within the Bayesian context the most
promising method seems to be the intervals of measures method.
A major problem for the application of this method to decision problems
seems to be the amount of calculation required, since for each decision
there is no single value for expected loss, but a set of such values
corresponding to all possible prior distributions. In this report the
determination of lower and upper bounds for such a set of expected loss
values with regard to a single decision is discussed, and general results
are derived which show that the situation is less severe than would be
expected at first sight. A simple algorithm to determine these bounds is
described. The choice of a decision can be based on a comparison of the
bounds of the expected loss per decision.

Key words and phrases
Bayesian decision theory, imprecise probabilities, intervals of measures,
lower and upper bounds for expected loss.

1. Introduction

The Bayesian theory of statistical inference demands a prior distribution for some parameter. The prior represents information available before experimental data became available. Walley (1991) presents a clear and extensive discussion about the major drawbacks of describing a certain amount of information, or lack of information, by a single prior distribution. This leads to a generalization of the concept in which the single prior distribution is replaced by a set of prior distributions, using a generalized concept of probability, imprecise probabilities. Pericchi and Walley (1991) conclude that the intervals of measures method, introduced by DeRobertis and Hartigan (1981), is the most suitable form for a set of prior distributions in Bayesian inference.

The idea of imprecise probabilities goes back to Boole (1854), but little attention was paid to it until the contributions of Smith (1961), who proposed an axiom system for imprecise probabilities (Wolfenson and Fine, 1982) and Good (1962). Dempster (1968) proposed a framework of statistical inference that leads to unreasonable results in some cases (Walley, 1991; section 5.13). Walley (1991) provides a thorough mathematical foundation of a coherent theory of imprecise probabilities, where the interpretation of lower and upper probabilities is in terms of betting rates, strongly related to the ideas of de Finetti (1974).

In his book Walley pays little attention to application of imprecise probabilities to statistical decision theory, in which field the Bayesian concept is attractive (Lindley, 1973). Other interesting contributions to decision theory with imprecise probabilities are provided by Fishburn (1965), Gärdenfors and Sahlin (1982,1983), Levi (1982) and Wolfenson and Fine (1982). Smith (1961) also pays attention to decision making. Berger (1985) discusses a method of sensitivity analysis within statistical decision theory, called robust Bayesian inference. This method differs from the theory discussed by Walley (1991) in that Berger does not use imprecision in probabilities as a tool to report the amount of information on which the probabilities are based, the most important reason for introducing the generalization.

While these contributions develop the theory they do not provide an effective basis for use through a method of calculation. All these methods imply an exhaustive process of evaluating a loss function at each decision
and for all possible priors.

In this report it is shown that the amount of calculation is less than expected at first sight, when the set of prior probabilities is given through intervals of measures. The results make the practical application of imprecise prior probabilities to decision problems possible. Important additional aspects for such application, e.g. introduction of suitable statistical models and assessment of imprecise prior probabilities, are discussed by Coolen (1992a,b,1993).

In section 2 of this report Bayesian decision theory with imprecise prior probabilities is introduced and the restriction to intervals of measures is suggested. In section 3 results are given for possible practical application by reducing the required amount of calculation. An algorithm for general problems is described.
2. Bayesian decision theory with imprecise prior probabilities

Decision problems (Lindley, 1990) can, with a good deal of generality, be described as follows. Let \( X \) be the sample space of points \( x \) and \( \Theta \) the parameter space of points \( \theta \). These are connected by \( p(x|\theta) \), the probability density of \( X \) for a given \( \theta \) (with respect to some measure). Let \( D \) be the decision space of points \( d \), and \( L(d,\theta) \) the loss in selecting \( d \) when \( \theta \) obtains (we assume \( L(d,\theta) \in \mathbb{R} \) for all \( d \) and \( \theta \)). The Bayesian approach uses a prior probability density \( \pi(\theta) \) over parameter space \( \Theta \) and chooses as the optimum decision that \( d \) which minimizes the expected loss

\[
E_L(d,x) = \int_{\Theta} L(d,\theta) p(\theta|x) d\theta,
\]

where \( p(\theta|x) \) is the posterior probability density of \( \theta \), given \( x \), obtained by Bayes' theorem. If no data \( x \) are available, the decision can be based on the prior, in which case the optimum decision is that \( d \) which minimizes the expected loss

\[
E_L(d) = \int_{\Theta} L(d,\theta) \pi(\theta) d\theta.
\]

Without loss of generality we work with (2), since the posterior at one stage is just the prior for the next (Coolen, 1992a, 1993). Suppose that the probability density \( \pi(\theta) \) is not known precisely, but is only known to belong to a space \( \Pi \) of probability densities. This generalization is known in literature as imprecise (prior) probabilities (Walley, 1991). In contrast to a hierarchical Bayesian approach we do not define a probability density over \( \Pi \). Assuming the loss function \( L(d,\theta) \) is known, the expected loss for \( \pi \in \Pi \) and \( d \in D \) is

\[
E_L(d,\pi) = \int_{\Theta} L(d,\theta) \pi(\theta) d\theta.
\]

For each decision \( d \in D \) this leads to a set values of expected loss, denoted by

\[
\mathcal{EL}(d,\Pi) = \{E_L(d,\pi) \mid \pi \in \Pi\}.
\]

To choose an optimal decision an additional criterion is needed, by which these sets \( \mathcal{EL}(d,\Pi) \) can be compared for all \( d \in D \). In this direction, two useful characteristics of \( \mathcal{EL}(d,\Pi) \) are the lower expectation \( \underline{E_L}(d,\Pi) \) and the upper expectation \( \overline{E_L}(d,\Pi) \):

\[
\underline{E_L}(d,\Pi) = \inf \mathcal{EL}(d,\Pi) \quad \text{and} \quad \overline{E_L}(d,\Pi) = \sup \mathcal{EL}(d,\Pi).
\]
As remarked by Pericchi and Wa1ley (1991), the intervals of measures method of DeRobertis and Hartigan (1981) is especially useful when imprecise probabilities are used for parameters, as in the Bayesian framework. Hence, in this paper we restrict \( \Pi \) to the form

\[
\Pi = \{ \pi | \pi(\theta) = q(\theta)/C_q \quad \text{with} \quad \ell(\theta) \leq q(\theta) \leq u(\theta) \quad \text{for all} \quad \theta \in \Theta \}
\]

and

\[
C_q = \int q(\theta)d\theta,
\]

where \( \ell \) and \( u \) are given and \( 0 \leq \ell(\theta) \leq u(\theta) \) for all \( \theta \in \Theta \). Further, in this paper we restrict to continuous \( \ell \) and \( u \) with \( \int \ell(\theta)d\theta > 0 \) and \( \int u(\theta)d\theta < \infty \). If \( \ell(\theta) = u(\theta) \) for all \( \theta \in \Theta \) then \( \Pi \) contains exactly one \( \theta \) probability density and the standard situation without imprecision occurs.

This report is restricted to imprecise prior probabilities. Another possible generalization of the standard Bayesian decision theory is obtained by allowing imprecise loss functions. However, if it is only known that the loss function \( L(d, \theta) \) is between two functions, say \( \underline{L}(d, \theta) \leq L(d, \theta) \leq \overline{L}(d, \theta) \) for all \( d \in D \) and \( \theta \in \Theta \), then the analysis does not become much more difficult, as (3), (4) and (5) imply that the lower bound of the set of all possible values for the expected loss must be calculated using \( \underline{L}(d, \theta) \), and the upper bound using \( \overline{L}(d, \theta) \).
3. Bounds for expected loss

To choose an optimal decision the sets \( E(d, \Pi) \) for all \( d \in D \) must be compared. An attractive simplification is derived by replacing this problem by one of comparing intervals \([E(d, \Pi), \overline{E}(d, \Pi)]\). This reduces, for each \( d \in D \), the calculation of \( EL(d, \pi) \) for all \( \pi \in \Pi \) to the calculation of only two values, and seems to be reasonable as \( E(d, \Pi) \subset [E(d, \Pi), \overline{E}(d, \Pi)] \) whereas no smaller closed interval exists that contains \( E(d, \Pi) \). However, it is still not clear that this reduces the amount of calculation. To this point, theorem 3.1 is an important result for \( \Pi \) of the form (6), as it shows that \( E(d, \Pi) \) and \( \overline{E}(d, \Pi) \) are determined by a subspace of \( \Pi \) consisting of probability densities that depend on \( L(d, \Theta) \), \( l(\Theta) \) and \( u(\Theta) \), and that can be characterized by one parameter that belongs to a parameter space that is bounded if the loss function is bounded. To derive at theorem 3.1 several additional definitions and lemmas are needed.

For each decision \( d \in D \) we define

\[
L_{l,d} = \inf\{L(d, \Theta) \mid \Theta \in \Theta\} \quad \text{and} \quad L_{u,d} = \sup\{L(d, \Theta) \mid \Theta \in \Theta\}, \tag{7}
\]

which may be equal to \(-\infty \) or \(+\infty \), and a partition of the parameter space \( \Theta \) for \( \psi \in [L_{l,d}, L_{u,d}] \)

\[
\Theta_0(d, \psi) = \{\Theta \in \Theta \mid L(d, \Theta) > \psi\}, \quad \Theta_1(d, \psi) = \{\Theta \in \Theta \mid L(d, \Theta) = \psi\} \quad \text{and} \quad \Theta_2(d, \psi) = \{\Theta \in \Theta \mid L(d, \Theta) < \psi\}. \tag{8}
\]

From this point we restrict to situations where the loss function \( L(d, \Theta) \) is such that

\[
\int \Theta_0(d, \psi) = 0 \quad \text{for all} \quad \psi \in [L_{l,d}, L_{u,d}] . \tag{9}
\]

After theorem 3.1 we will discuss the situation without this restriction. Remark that, if \( \Theta = \mathbb{R} \), this restriction implies that there is no interval in \( \Theta \) on which \( L(d, \Theta) \) is constant.

For \( d \in D \) we define, for \( \psi \in [L_{l,d}, L_{u,d}] \):

\[
C_{u,d}(\psi) = \int_{\Theta_0(d, \psi)} u(\Theta)d\Theta + \int_{\Theta_1(d, \psi)} \ell(\Theta)d\Theta \quad \text{and} \quad \tag{10}
\]
We define two probability densities in $\Pi$ that depend through $u$ on the above partition of $\Theta$:

$$
\pi_u^{\ell}(\theta) = \begin{cases}
    \frac{u(\theta)}{C_{u,d}(u)} & \text{for } \theta \in \Theta_u(d,u) \\
    \ell(\theta) & \text{for } \theta \in \Theta_1(d,u)
\end{cases}
$$

and

$$
\pi_u^{\ell}(\theta) = \begin{cases}
    \ell(\theta) & \text{for } \theta \in \Theta_u(d,u) \\
    \frac{u(\theta)}{C_1,d(u)} & \text{for } \theta \in \Theta_1(d,u)
\end{cases}
$$

Next we define subspaces of $\Pi$ by considering the normalizing constants of original densities between $\ell$ and $u$ that determine the probability densities in $\Pi$. With

$$
C_\ell = \int_\Theta \ell(\theta) d\theta \quad \text{and} \quad C_u = \int_\Theta u(\theta) d\theta
$$

we define, for $C_\ell \leq C \leq C_u$:

$$
\Pi(C) = \{ \pi \in \Pi \mid \pi(\Theta) = q(\theta)/C \quad \text{with} \quad \ell(\theta) \leq q(\theta) \leq u(\theta) \quad \text{for all } \theta \in \Theta
\}
$$

Lemma 3.1

$$
\Pi = U_{\Theta} \Pi(C) = U_{\Theta} \Pi(C_{\ell},d(u))
$$

and

$$
\Pi = U_{\Theta} \Pi(C) = U_{\Theta} \Pi(C_1,d(u))
$$

Proof

The first equality is an obvious consequence of (15).

The second equality of (16) holds because restriction (9) and continuity of $\ell(\theta)$ and $u(\theta)$ imply that $C_{u,d}(u)$ is a continuous non-increasing function of $u$, with $C_{u,d}(L_1,d) = C_\ell$ and $C_{u,d}(L_u,d) = C_u$.

Analogously, the second equality of (17) holds because $C_{\ell,d}(u)$ is a
continuous non-decreasing function of \( \psi \), with \( C_{\ell,d}(l_d, d) = C_{\ell} \) and \( C_{u,d}(l_u, d) = C_u \). □

Lemma 3.2 is an important result for the proof of theorem 3.1.

**Lemma 3.2**

Let \( d \in D \), and let \( \psi \in [l_u, d, l_u, d] \). Then the following relation holds for all probability densities \( \pi \in \Pi(C_u, d(\psi)) \):

\[
EL(d, \pi) \leq EL(d, \pi^u_{d, \psi}).
\]

(18)

And for all \( \pi \in \Pi(C_{l,d}(\psi)) \):

\[
EL(d, \pi) \leq EL(d, \pi^l_{d, \psi}).
\]

(19)

**Proof**

We prove (18).

Let \( \pi \in \Pi(C_u, d(\psi)) \) and let \( q_\pi(\theta) \) be a corresponding function between \( \ell \) and \( u \), so \( \pi(\theta) = q_\pi(\theta)/C_u, d(\psi) \) with \( \ell(\theta) \leq q_\pi(\theta) \leq u(\theta) \) for all \( \theta \in \Theta \). Then

\[
EL(d, \pi) = \int_{\Theta} L(d, \theta) \pi(\theta) d\theta = \int_{\Theta} L(d, \theta) q_\pi(\theta) d\theta/C_u, d(\psi)
\]

\[
= \int_{\Theta_u(d, \psi)} L(d, \theta) q_\pi(\theta) d\theta/C_u, d(\psi) + \int_{\Theta_l(d, \psi)} L(d, \theta) q_\pi(\theta) d\theta/C_u, d(\psi)
\]

\[
\leq \int_{\Theta_u(d, \psi)} L(d, \theta) u(\theta) d\theta/C_u, d(\psi) + \int_{\Theta_l(d, \psi)} L(d, \theta) \ell(\theta) d\theta/C_u, d(\psi)
\]

\[
= \int_{\Theta} L(d, \theta) \pi^u_{d, \psi}(\theta) d\theta = EL(d, \pi^u_{d, \psi}).
\]

To prove the inequality herein we use the fact that

\[
C_u, d(\psi) = \int_{\Theta_u(d, \psi)} u(\theta) d\theta + \int_{\Theta_l(d, \psi)} \ell(\theta) d\theta = \int_{\Theta_u(d, \psi)} q_\pi(\theta) d\theta + \int_{\Theta_l(d, \psi)} q_\pi(\theta) d\theta,
\]

so

\[
\int_{\Theta_u(d, \psi)} [u(\theta) - q_\pi(\theta)] d\theta = \int_{\Theta_l(d, \psi)} [q_\pi(\theta) - \ell(\theta)] d\theta,
\]

which, in combination with (8), leads to
\[
\int L(d, \theta) u(\theta) d\theta + \int L(d, \theta) \ell(\theta) d\theta - \int L(d, \theta) q_\pi(\theta) d\theta - \int L(d, \theta) q_{\pi, 1}(\theta) d\theta
= \int L(d, \theta) [u(\theta) - q_\pi(\theta)] d\theta + \int L(d, \theta) [\ell(\theta) - q_{\pi, 1}(\theta)] d\theta
\]

Remark that restriction (9) is necessary.

The proof of (19) is analogous.

Now theorem 3.1 can be proved. This is an essential result since it shows that for a given decision \(d \in D\), one need consider only probability densities of the form \(\pi^u_d(\theta)\) to determine \(E^u(d, \Pi)\), and probability densities of the form \(\pi^l_d(\theta)\) to determine \(E^l(d, \Pi)\).

We define

\[
\Pi^u_d = \{ \pi^u_d, \psi \in \Pi \mid \psi \in \{L_1, d, L_u, d'\} \}; \\
E^u(d, \Pi^u_d) = \{ EL(d, \pi) \mid \pi \in \Pi^u_d \};
\]

\[
E^u(d, \Pi^u_d) = \sup \{ \Pi^u_d \};
\]

\[
\Pi^l_d = \{ \pi^l_d, \psi \in \Pi \mid \psi \in \{L_1, d, L_u, d'\} \}; \\
E^l(d, \Pi^l_d) = \{ EL(d, \pi) \mid \pi \in \Pi^l_d \};
\]

\[
E^l(d, \Pi^l_d) = \inf \{ E^l(d, \pi) \mid \pi \in \Pi^l_d \};
\]

8
Theorem 3.1

Let \( d \in D \) and let the loss function \( L(d, \theta) \) satisfy restriction (9), then
\[
\mathcal{E}L(d, \Pi) = \mathcal{E}L(d, \Pi_d^u)
\] (26)

and
\[
\mathcal{E}L(d, \Pi) = \mathcal{E}L(d, \Pi_d^u).
\] (27)

Proof

We prove (26).

\[
\mathcal{E}L(d, \Pi) = \sup \{ EL(d, \pi) \mid \pi \in \Pi \} = \sup \{ \sup \{ EL(d, \pi) \mid \pi \in \Pi(C) \} \mid C \in C_d \} =
\]

\[
\sup \{ \sup \{ EL(d, \pi) \mid \pi \in \Pi(C_u, \tilde{d}(w)) \} \mid \tilde{w} \in [L_1, d' \tilde{L}_u, d] \} =
\]

\[
\sup \{ EL(d, \pi_d^u) \mid \tilde{w} \in [L_1, d' \tilde{L}_u, d] \} = \sup \mathcal{E}L(d, \Pi_d^u) = \mathcal{E}L(d, \Pi_d^u).
\]

The second and third equality are based on lemma 3.1, the fourth on lemma 3.2. The other equalities follow from (4), (5), (20), (21) and (22).

The proof of (27) is analogous. \( \square \)

In theorem 3.1 restriction (9) needs to be satisfied. Of course loss functions that do not satisfy this restriction can be of interest, so the above theory needs to be generalized. We first consider what happens if restriction (9) is not satisfied, so if
\[
\exists \tilde{w} \in [L_1, d' \tilde{L}_u, d] \text{ such that } \int d\theta > 0.
\] (28)

Then the second equalities of (16) and (17) do not hold, as \( C_{u, d}(w) \) and \( C_{1, d}(w) \) are no longer continuous functions of \( w \). In fact, if
\[
C_{u, d}(w) = \lim_{\tilde{w} \uparrow w} C_{u, d}(w) \text{ and } C_{u, d}(w^+) = \lim_{\tilde{w} \downarrow w} C_{u, d}(w)
\] (29)

and
\[
C_{1, d}(w^-) = \lim_{\tilde{w} \uparrow w} C_{1, d}(w) \text{ and } C_{1, d}(w^+) = \lim_{\tilde{w} \downarrow w} C_{1, d}(w)
\] (30)

exist, then
\[ C_{u,d}^{(-)} = \int u(\theta) d\theta + \int u(\theta) d\theta + \int \ell(\theta) d\theta; \]  
\[ C_{u,d}^{(+) = \int u(\theta) d\theta + \int \ell(\theta) d\theta + \int \ell(\theta) d\theta. \]  
\[ C_{l,d}^{(-)} = \int \ell(\theta) d\theta + \int \ell(\theta) d\theta + \int u(\theta) d\theta; \]  
\[ C_{l,d}^{(+) = \int \ell(\theta) d\theta + \int u(\theta) d\theta + \int u(\theta) d\theta. \]

If \( \int [u(\theta) - \ell(\theta)] d\theta > 0 \), then \( \pi \not\in \bigcup_{\delta=L_1,d} \Pi(C_{u,d}(\omega)) \) for \( \pi \in \Pi(C) \) if \( \Theta_L(d,\omega) \).

To solve this problem, we define (for \( \omega \in [L_1,d,L_{u,d}] \) and \( q \in [0,1] \))

\[ C_{u,d}^{(\omega,q)} = C_{u,d}(\omega) + \int_{\Theta_L(d,\omega)} [qu(\theta)+(1-q)\ell(\theta)] d\theta \]  
and \[ C_{l,d}^{(\omega,q)} = C_{l,d}(\omega) + \int_{\Theta_L(d,\omega)} [q\ell(\theta)+(1-q)u(\theta)] d\theta. \]

Of course, if \( \int d\theta = 0 \) then \( C_{u,d}^{(\omega,q)} = C_{u,d}(\omega) \) and \( C_{l,d}^{(\omega,q)} = C_{l,d}(\omega) \) for all \( q \in [0,1] \).
Further, we define

\[
\Pi(C_{u,d}(\psi,q)) = \{ \pi \in \Pi \mid \pi(\theta) = q(\theta)/C_q \text{ with } \ell(\theta) \leq q(\theta) \leq u(\theta) \}
\]

for all \( \theta \in \Theta \),

\[
C = \int_{\Theta} q(\theta) d\theta = C_{u,d}(\psi,q) \quad \text{and} \quad \int_{\Theta} q(\theta) d\theta = \int_{\Theta} [q\ell(\theta) + (1-q)u(\theta)] d\theta. \tag{37}
\]

and

\[
\Pi(C^+_{u,d}(\psi,q)) = \{ \pi \in \Pi \mid \pi(\theta) = q(\theta)/C_q \text{ with } \ell(\theta) \leq q(\theta) \leq u(\theta) \}
\]

for all \( \theta \in \Theta \),

\[
C = \int_{\Theta} q(\theta) d\theta = C^+_{u,d}(\psi,q) \quad \text{and} \quad \int_{\Theta} q(\theta) d\theta = \int_{\Theta} [q\ell(\theta) + (1-q)u(\theta)] d\theta. \tag{38}
\]

Lemma 3.3 is a generalized form of lemma 3.1.

**Lemma 3.3**

\[
\begin{align*}
\Pi(C_{u,d}^+) & = U \Pi(C) = U \Pi(C_{u,d}(\psi,q)) \quad (39) \\
\Pi(C_{u,d}^+) & = U \Pi(C) = U \Pi(C_{u,d}(\psi,q)) \quad (40)
\end{align*}
\]

Proof

The second equality of (39) follows from

\[
\{ C^+_{u,d}(\psi,q) \mid \psi \in [L_u,d'], q \in [0,1] \} = [C^+_u,C]
\]

where again continuity of \( \ell(\theta) \) and \( u(\theta) \) plays a role.

The proof of (40) is analogous. \( \Box \)
Definitions (12) and (13) must also be replaced in this generalized situation. Here we use the fact that we know that \( L(d, \theta) = \omega \) for all \( \theta \in \Theta_\omega (d, \omega) \), so to determine \( \mathcal{F}_\omega (d, \Pi) \) and \( \mathcal{F}_\omega (d, \Pi) \) we do not need to specify probability densities analogous to definitions (12) and (13) for all \( \theta \in \Theta \).

We define, for all \( q \in [0, 1] \)

\[
L^{*}_{d, \omega, q, \theta} = \begin{cases} \frac{u(\theta)}{C_{u, d}(\omega, q)} & \text{for } \theta \in \Theta_\omega (d, \omega) \\ \frac{\ell(\theta)}{C_{u, d}(\omega, q)} & \text{for } \theta \in \Theta (d, \omega) \end{cases}
\]

(42)

whereas we do not need to specify \( L^{*}_{d, \omega, q, \theta} \) exactly for \( \theta \in \Theta_\omega (d, \omega) \), but these values need to be such that

\[
\int_{\Theta_\omega (d, \omega)} q(\theta) d\theta = \int_{\Theta (d, \omega)} \left[ q u(\theta) + (1-q) \ell(\theta) \right] d\theta.
\]

(43)

Analogously, we define, for all \( q \in [0, 1] \)

\[
L^{*}_{d, \omega, q, \theta} = \begin{cases} \frac{\ell(\theta)}{C_{u, d}(\omega, q)} & \text{for } \theta \in \Theta_\omega (d, \omega) \\ \frac{u(\theta)}{C_{u, d}(\omega, q)} & \text{for } \theta \in \Theta (d, \omega) \end{cases}
\]

(44)

and the values of \( L^{*}_{d, \omega, q, \theta} \) for \( \theta \in \Theta_\omega (d, \omega) \) need to be such that

\[
\int_{\Theta_\omega (d, \omega)} q(\theta) d\theta = \int_{\Theta (d, \omega)} \left[ q \ell(\theta) + (1-q) u(\theta) \right] d\theta.
\]

(45)

For both (43) and (45) it is possible to derive such \( q(\theta) \) for each \( q \in [0, 1] \); for \( q \) we could even choose continuous functions as \( \ell \) and \( u \) are continuous. However, we do not need to specify these \( q \).

Now a generalized form of lemma 3.2 is given.
Lemma 3.4

Let $d \in D$, $u \in [L^d_U, L^d_U]$, and $q \in [0,1]$. Then the following relation holds for all probability densities $\pi \in \Pi(C^*, d(U,q))$:

$$EL(d, \pi) \leq EL(d, \pi_{d, u, q}^*)$$  \hspace{1cm} (46)

And for all $\pi \in \Pi(C_{d}^*, d(U,q))$:

$$EL(d, \pi) \geq EL(d, \pi_{d, u, q}^*)$$  \hspace{1cm} (47)

Proof

We prove (46).

Let $\pi \in \Pi(C^*, d(U,q))$ and let $q^*_\pi(\theta)$ be a corresponding function between $l$ and $u$, so $\pi(\theta) = q^*_\pi(\theta) / C^*_{u, d}(\psi, q)$ with $\ell(\theta) \leq q^*_\pi(\theta) \leq u(\theta)$ for all $\theta \in \Theta$, and also with $\int \int q^*_\pi(\theta) d\theta = \int [q(\theta) + (1-q) \ell(\theta)] d\theta$. Then

$$EL(d, \pi) = \int L(d, \theta) \pi(\theta) d\theta = \int L(d, \theta) q^*_\pi(\theta) d\theta / C^*_{u, d}(\psi, q)$$

$$= \int \int L(d, \theta) q^*_\pi(\theta) d\theta / C^*_{u, d}(\psi, q) + \int \int L(d, \theta) q^*_\pi(\theta) d\theta / C^*_{u, d}(\psi, q)$$

$$\Theta_u(d, \psi) \Theta_u(d, \psi)$$

$$= \int \int L(d, \theta) q^*_\pi(\theta) d\theta / C^*_{u, d}(\psi, q) + \int \int L(d, \theta) q^*_\pi(\theta) d\theta / C^*_{u, d}(\psi, q)$$

$$\Theta_u(d, \psi) \Theta_u(d, \psi)$$

$$= \int \int q(\theta) + (1-q) \ell(\theta) d\theta / C^*_{u, d}(\psi, q)$$

The inequality is proved analogously to the proof of lemma 3.2.

The proof of (47) is analogous.
Now a generalized form of theorem 3.1 can be proved. We define

\[ \Pi^*_d = \{ \pi_d, \psi, q \in \Pi \mid \psi \in [L_{d'}, L_{u, d}], q \in [0, 1] \} ; \]  
\[ S(d, \Pi^*_d) = \{ E(d, \pi) \mid \pi \in \Pi^*_d \} ; \]  
\[ \overline{E}(d, \Pi^*_d) = \sup \ S(d, \Pi^*_d) ; \]  
\[ \overline{E}(d, \Pi^*_d) = \sup \ S(d, \Pi^*_d) ; \]

\[ \Pi^*_d = \{ \pi_d, \psi, q \in \Pi \mid \psi \in [L_{d'}, L_{u, d}], q \in [0, 1] \} ; \]  
\[ S(d, \Pi^*_d) = \{ E(d, \pi) \mid \pi \in \Pi^*_d \} \quad \text{and} \]
\[ \overline{E}(d, \Pi^*_d) = \inf \ S(d, \Pi^*_d) . \]

**Theorem 3.2**

Let \( d \in D \), then

\[ \overline{E}(d, \Pi) = \overline{E}(d, \Pi^*_d) , \]  
and
\[ \overline{E}(d, \Pi) = \overline{E}(d, \Pi^*_d) . \]

**Proof**

We prove (54).

\[ \overline{E}(d, \Pi) = \sup \{ E(d, \pi) \mid \pi \in \Pi \} = \sup \{ \sup \{ E(d, \pi) \mid \pi \in \Pi \} \mid C \in [C_d, C_u] \} = \sup \{ \sup \{ E(d, \pi) \mid \pi \in \Pi \} \mid \psi \in [L_{d'}, L_{u, d}], q \in [0, 1] \} = \sup \{ E(d, \pi^*_d, \psi, q) \mid \psi \in [L_{d'}, L_{u, d}], q \in [0, 1] \} = \sup \overline{E}(d, \Pi^*_d) = \overline{E}(d, \Pi^*_d) . \]

The second and third equality are based on lemma 3.3, the fourth on lemma 3.4. The other equalities follow from (4), (5), (48), (49) and (50).

The proof of (55) is analogous.
The above theory suggests that, to determine $g_2(d,T_1)$ and $g_2(d,T_2)$, we need to calculate $u_\ast EL(d,l_l d', l_l d')$ and $1_\ast EL(d,l_l d', l_l d')$ for all $q \in [0,1]$. However, (see remark after (36)), if $\int \theta d\theta = 0$ then $\Theta(q,d,u)$ and $\Theta(q,d,u)$ are equal. Next we show that also for $\psi \in [L_1, d', L_1, u, d]$ with $\int \theta d\theta > 0$ it is not necessary to calculate $EL(d,\pi_{d', \psi} u, q)$ and $EL(d,\pi_{d', \psi} l, q)$ for all $q \in [0,1]$. This will lead to theorem 3.3, a more useful result than theorem 3.2.

**Lemma 3.5**

Let $d \in D$, and let $\psi \in [L_1, d', L_1, u, d]$. Then, for all $q \in [0,1]$

$$EL(d,\pi_{d', \psi} u, q) \leq \sup \{ EL(d,\pi_{d', \psi} u, q) \mid q \in [0,1] \}$$

and

$$EL(d,\pi_{d', \psi} l, q) \geq \inf \{ EL(d,\pi_{d', \psi} l, q) \mid q \in [0,1] \}.$$

**Proof**

We prove (56). Let $q \in [0,1]$, then

$$EL(d,\pi_{d', \psi} u, q) = \int L(d,\theta)\pi_{d', \psi} u, q(\theta)d\theta = \frac{A + \psi h(q)}{B + h(q)},$$

with

$$A = \int \Theta(u,d,u)\theta d\theta + \int \Theta(l,d,u)\theta d\theta,$n

$$B = \int u(\theta)d\theta + \int l(\theta)d\theta,$n

and

$$h(q) = \int \Theta(q,d,\theta)\theta d\theta.$$

If $\psi > A/B$ then the function $\frac{A + \psi h(q)}{B + h(q)}$ is an increasing function of $h(q)$, and since $h(q)$ is a non-decreasing function of $q$ it follows that

$$\frac{A + \psi h(q)}{B + h(q)} \leq \frac{A + \psi h(1)}{B + h(1)} \text{ for all } q \in [0,1].$$

(58)
If $\psi < A/B$ then the function $\frac{A + \psi h(q)}{B + h(q)}$ is a decreasing function of $h(q)$ so $\frac{A + \psi h(q)}{B + h(q)} \leq \frac{A + \psi h(0)}{B + h(0)}$ for all $q \in [0,1]$. \hfill (59)

Finally, if $\psi = A/B$ then the function $\frac{A + \psi h(q)}{B + h(q)}$ is a constant function of $h(q)$.

The proof of (57) is analogous. \hfill $\square$

To derive a result that reduces the amount of calculation compared to theorem 3.2, we define

\[ \Pi^\ast_d \ni \{ \pi_d^\ast, \psi, q \in \Pi \mid \psi \in [L, d, L', u, d], q \in \{0,1\} \}; \quad (60) \]

\[ C_L(d, \Pi^\ast_d) = \{ E_L(d, \pi) \mid \pi \in \Pi^\ast_d \}; \quad (61) \]

\[ C_L(d, \Pi^\ast_d) = \sup \mathcal{C}_L(d, \Pi^\ast_d); \quad (62) \]

\[ \Pi^\dagger_d \ni \{ \pi_d^\dagger, \psi, q \in \Pi \mid \psi \in [L, d, L', u, d], q \in \{0,1\} \}; \quad (63) \]

\[ E_L(d, \Pi^\dagger_d) = \{ E_L(d, \pi) \mid \pi \in \Pi^\dagger_d \} \quad \text{and} \quad (64) \]

\[ E_L(d, \Pi^\dagger_d) = \inf \mathcal{C}_L(d, \Pi^\dagger_d). \quad (65) \]

**Theorem 3.3**

Let $d \in D$, then

\[ C_L(d, \Pi) = C_L(d, \Pi^\ast_d), \quad (66) \]

and

\[ C_L(d, \Pi) = C_L(d, \Pi^\dagger_d). \quad (67) \]

**Proof**

Equalities (66) and (67) are based on lemma 3.5 and theorem 3.2 together with definitions (60)-(65). \hfill $\square$
For each $d \in D$, the problem of determining $\mathcal{E}_X(d, \Pi)$ and $\mathcal{E}_X(d, \Pi_d)$ can, as a result of theorem 3.3, be replaced by the problem of determining $\mathcal{E}_X(d, \Pi_d^{u,**})$ and $\mathcal{E}_X(d, \Pi_d^{l,**})$. To solve these optimization problems we must consider $EL(d, \pi_d^{u,*}, \psi, \varphi)$ and $EL(d, \pi_d^{l,*}, \psi, \varphi)$ for all $\psi \in [L_d, d', L_d, d]$ and $\varphi \in \{0,1\}$.

We next show that these optimization problems are much simpler than they seem to be, because of the form of $EL(d, \pi_d^{u,*}, \psi, \varphi)$ and $EL(d, \pi_d^{l,*}, \psi, \varphi)$ as functions of $\psi$.

Lemma 3.6 is an important result on which the proof of the remarkable theorem 3.4 is based. We introduce a new notation:

\[ \Theta_{u,s}(d, \psi) = \{ \theta \in \Theta | L(d, \theta) \geq u \} = \Theta_u \cup \Theta_s(d, \psi) \quad \text{and} \]
\[ \Theta_{l,s}(d, \psi) = \{ \theta \in \Theta | L(d, \theta) \leq u \} = \Theta_l \cup \Theta_s(d, \psi). \]  

**Lemma 3.6**

Let $d \in D$ and $\psi \in [L_d, d', L_d, d]$. For $EL(d, \pi_d^{u,*}, \psi, \varphi)$ the following relations hold:

If $\min\{EL(d, \pi_d^{u,*}, \psi, \varphi) | \varphi \in \{0,1\} \} \geq \psi$, then for all $\psi \in [L_d, d', \psi]$:

\[ \max\{EL(d, \pi_d^{u,*}, \psi, \varphi) | \varphi \in \{0,1\} \} \leq \max\{EL(d, \pi_d^{u,*}, \psi, \varphi) | \varphi \in \{0,1\} \}. \]  

If $\max\{EL(d, \pi_d^{u,*}, \psi, \varphi) | \varphi \in \{0,1\} \} \leq \psi$, then for all $\psi \in (\psi, L_d, \psi)$:

\[ \max\{EL(d, \pi_d^{u,*}, \psi, \varphi) | \varphi \in \{0,1\} \} \leq \max\{EL(d, \pi_d^{u,*}, \psi, \varphi) | \varphi \in \{0,1\} \}. \]

And, for $EL(d, \pi_d^{l,*}, \psi, \varphi)$ analogous relations are:

If $\min\{EL(d, \pi_d^{l,*}, \psi, \varphi) | \varphi \in \{0,1\} \} \geq \psi$, then for all $\psi \in [L_d, d', \psi]$:

\[ \min\{EL(d, \pi_d^{l,*}, \psi, \varphi) | \varphi \in \{0,1\} \} \leq \min\{EL(d, \pi_d^{l,*}, \psi, \varphi) | \varphi \in \{0,1\} \}. \]  

If $\max\{EL(d, \pi_d^{l,*}, \psi, \varphi) | \varphi \in \{0,1\} \} \leq \psi$, then for all $\psi \in (\psi, L_d, \psi)$:

\[ \min\{EL(d, \pi_d^{l,*}, \psi, \varphi) | \varphi \in \{0,1\} \} \leq \min\{EL(d, \pi_d^{l,*}, \psi, \varphi) | \varphi \in \{0,1\} \}. \]
Proof
We prove (69) and (70).
For the proof of (69), let $w, w' \in [L_U, L_{d', L_{d}}]$ with $w < w'$. We need to compare $EL(d, \pi_{d, w}, q_w)$ with $EL(d, \pi_{d, w'}, q_{w'})$ for all four possible combinations of $q_w \in \{0, 1\}$ and $q_{w'} \in \{0, 1\}$. Let $q_w = q_{w'} = 0$, then the fact that the following relations hold:

$$
\Theta_{u(U)}(d, w) \leq \Theta_u(d, w), \quad \Theta_{l(w)}(d, w) \leq \Theta_{l}(d, w) \quad \text{and}
$$

$$
\Theta_u(d, w) \backslash \Theta_{u}(d, w) = \{ \theta \in \Theta_u(d, w) : \theta \notin \Theta_{u}(d, w) \} = \Theta_{u}(d, w) \backslash \Theta_{l(w)}(d, w),
$$

leads to

$$
EL(d, \pi_{d, w, 0}, 0) \leq EL(d, \pi_{d, w', 0}) \quad \implies \quad
$$

$$
\int \frac{L(d, \theta) u(\theta) d\theta + \int L(d, \theta) \ell(\theta) d\theta}{\Theta_u(d, w) \backslash \Theta_{u}(d, w)} \leq \frac{EL(d, \pi_{d, w', 0})}{\Theta_u(d, w) \backslash \Theta_{u}(d, w)} \quad \implies \quad
$$

$$
\int \frac{L(d, \theta) u(\theta) d\theta + \int L(d, \theta) \ell(\theta) d\theta + \int L(d, \theta) [u(\theta) - \ell(\theta)] d\theta}{\Theta_u(d, w) \backslash \Theta_{u}(d, w)} \leq \frac{EL(d, \pi_{d, w', 0})}{\Theta_u(d, w) \backslash \Theta_{u}(d, w)} \quad \implies \quad
$$

$$
\int \frac{L(d, \theta) u(\theta) d\theta + \int L(d, \theta) \ell(\theta) d\theta + \int \Theta_u(d, w) \backslash \Theta_{u}(d, w)}{\Theta_u(d, w) \backslash \Theta_{u}(d, w)} \leq \frac{EL(d, \pi_{d, w', 0})}{\Theta_u(d, w) \backslash \Theta_{u}(d, w)} \quad \implies \quad
$$

From the fact that $L(d, \theta) \leq w$ for all $\theta \in \Theta_u(d, w) \backslash \Theta_{u}(d, w)$, we conclude that (74) holds if $EL(d, \pi_{d, w', 0}) \geq w$.

The proof is similar for the other three possible combinations of $q_w$ and $q_{w'}$. 
for which one has to replace the partitions of $\Theta$ but for which analogous
relations as (74) appear. This completes the proof of (69), as
\[
\max\{EL(d, \pi^{u,*}_d, w, q) \mid q \in \{0, 1\}\} \leq \min\{EL(d, \pi^{u,*}_d, w, q) \mid q \in \{0, 1\}\}
\]
\[
\leq \max\{EL(d, \pi^{u,*}_d, w, q) \mid q \in \{0, 1\}\}.
\]
(75)

The proof of (70) is quite similar to that of (69). Let $\psi, \omega \in [L_1, L_1, L_u, 0]$ with $\omega < \psi$ and let $\psi, \omega = 0$, then analogous relations as (73) are used to prove that:
\[
EL(d, \pi^{u,*}_{d, \omega, 0}) \leq EL(d, \pi^{u,*}_{d, \psi, 0})
\]
\[
\int \left[ L(d, \theta) u(\theta) d\theta + \int \left[ L(d, \theta) \ell(\theta) d\theta - \int L(d, \theta) [u(\theta) - \ell(\theta)] d\theta
\right]
\right]
\]
\[
\Theta_u(d, \psi) \Theta_{L_1}(d, \psi) \Theta_u(d, \psi) \Theta_u(d, \psi)
\]
\[
\Theta_u(d, \psi) \Theta_{L_1}(d, \psi) \Theta_u(d, \psi) \Theta_u(d, \psi)
\]
\[
\int [u(\theta) - \ell(\theta)] d\theta
\]
\[
\Theta_u(d, \psi) \Theta_u(d, \psi)
\]
\[
\leq EL(d, \pi^{u,*}_{d, \psi, 0}).
\]
(76)

From the fact that $\omega < L(d, \theta)$ for all $\theta \in \Theta_u(d, \psi) \Theta_u(d, \psi)$, we conclude that (76) certainly holds if $EL(d, \pi^{u,*}_{d, \psi, 0}) \leq \omega$.

Again the proof is analogous for the other three possible combinations of $q_\psi$ and $q_\omega$, for which one has to replace the partitions of $\Theta$ but for which analogous relations as (76) appear. This completes the proof of (70), again by use of (75).

The proofs of relations (71) and (72) are analogous. \qed
Theorem 3.4 is a useful result for the optimization process necessary to calculate \( \overline{gF}(d,\Pi) \) and \( \underbar{gF}(d,\Pi) \), as it gives us as well a strong tool to delete a part of \( \Theta \) while searching for these optima (after one calculation) as a sufficient condition for a calculated value to be such an optimum.

We define

\[
\Pi^{u,*} = \{ \pi^{u,*} \in \Pi \mid \omega \in [u, L_u, d], q \in \{0,1\} \};
\]  
\( (77) \)

\[
\overline{EL}(d, \Pi^{u,*}) = \{ EL(d, \pi) \mid \pi \in \Pi^{u,*} \};
\]  
\( (78) \)

\[
\overline{EL}(d, \Pi^{u,*}) = \sup \overline{EL}(d, \Pi^{u,*});
\]  
\( (79) \)

\[
\Pi^{u,*} = \{ \pi^{u,*} \in \Pi \mid \omega \in [L_1, d', \omega], q \in \{0,1\} \};
\]  
\( (80) \)

\[
\overline{EL}(d, \Pi^{u,*}) = \{ EL(d, \pi) \mid \pi \in \Pi^{u,*} \};
\]  
\( (81) \)

\[
\overline{EL}(d, \Pi^{u,*}) = \sup \overline{EL}(d, \Pi^{u,*});
\]  
\( (82) \)

\[
\Pi^{1,*} = \{ \pi^{1,*} \in \Pi \mid \omega \in [L_1, d', \omega], q \in \{0,1\} \};
\]  
\( (83) \)

\[
\overline{EL}(d, \Pi^{1,*}) = \{ EL(d, \pi) \mid \pi \in \Pi^{1,*} \};
\]  
\( (84) \)

\[
\overline{EL}(d, \Pi^{1,*}) = \inf \overline{EL}(d, \Pi^{1,*});
\]  
\( (85) \)

\[
\Pi^{1,*} = \{ \pi^{1,*} \in \Pi \mid \omega \in [L_1, d', \omega], q \in \{0,1\} \};
\]  
\( (86) \)

\[
\overline{EL}(d, \Pi^{1,*}) = \{ EL(d, \pi) \mid \pi \in \Pi^{1,*} \};
\]  
\( (87) \)

\[
\overline{EL}(d, \Pi^{1,*}) = \inf \overline{EL}(d, \Pi^{1,*});
\]  
\( (88) \)
Theorem 3.4

Let \( \theta \in D \) and \( \psi \in [L_1, L_u, L_d^*] \).

If \( \min \{ EL(d, \pi_{d, \psi}^*, q) \mid q \in [0,1] \} \geq \psi \), then
\[
\mathcal{E}(d, \Pi) = \mathcal{E}(d, \Pi_{d, \psi}^*).
\] (89)

If \( \max \{ EL(d, \pi_{d, \psi}^*, q) \mid q \in [0,1] \} \leq \psi \), then
\[
\mathcal{E}(d, \Pi) = \mathcal{E}(d, \Pi_{d, \psi}^*).
\] (90)

If \( \min \{ EL(d, \pi_{d, \psi}^j, q) \mid q \in [0,1] \} \geq \psi \), then
\[
\mathcal{E}(d, \Pi) = \mathcal{E}(d, \Pi_{d, \psi}^j).
\] (91)

If \( \max \{ EL(d, \pi_{d, \psi}^j, q) \mid q \in [0,1] \} \leq \psi \), then
\[
\mathcal{E}(d, \Pi) = \mathcal{E}(d, \Pi_{d, \psi}^j).
\] (92)

If \( \min \{ EL(d, \pi_{d, \psi}^*, q) \mid q \in [0,1] \} = \max \{ EL(d, \pi_{d, \psi}^*, q) \mid q \in [0,1] \} = \psi \), then
\[
\mathcal{E}(d, \Pi) = \psi.
\] (93)

If \( \min \{ EL(d, \pi_{d, \psi}^j, q) \mid q \in [0,1] \} = \max \{ EL(d, \pi_{d, \psi}^j, q) \mid q \in [0,1] \} = \psi \), then
\[
\mathcal{E}(d, \Pi) = \psi.
\] (94)

Proof

Relations (89)-(92) are based on theorem 3.3, lemma 3.6 and definitions (77)-(88). Relation (93) is a result of combination of (89), (90) and lemma 3.6, and (94) of combination of (91), (92) and lemma 3.6.

For a special loss function \( L(d, \theta) \) corollary 3.1 is a useful result of theorem 3.4.
Corollary 3.1

Let \( d \in D \) and let \( L(d, \theta) \) be continuous, bounded (so \( L_1, d^\gamma > \omega \) and \( L_u, d^\omega \)) and such that restriction (9) holds. As assumed before, let \( l(\theta) \) and \( u(\theta) \) also be continuous.

Then there exist \( \psi_0, \psi_1 \in [L_1, d', L_u, d] \) such that

\[
\int_{d}^{u} \frac{L(d, \theta)}{\psi_0} \, d\theta = \int_{d}^{u} \frac{1}{\psi_1} \, d\theta,
\]

which states that \( \psi_0 \) and \( \psi_1 \) are such that

\[
\int_{d}^{u} \frac{L(d, \theta)}{\psi_0} \, d\theta = \int_{d}^{u} \frac{1}{\psi_1} \, d\theta.
\]

Proof

We prove (95).

Let \( \psi_0 \in [L_1, d', L_u, d] \), then, as restriction (9) holds, theorem 3.4 states that a sufficient condition for \( \int_{d}^{u} \frac{L(d, \theta)}{\psi_0} \, d\theta = \int_{d}^{u} \frac{1}{\psi_1} \, d\theta \) is \( E(d, u, \psi_0) = \psi_0 \).

For \( \psi \in [L_1, d', L_u, d] \):

\[
E(d, u, \psi) - \psi = 0 \iff \int_{d}^{u} [\frac{1}{\psi} - \frac{L(d, \theta)}{\psi_0}] \, d\theta = 0.
\]

In (97) \( k_1(\psi) \) is a continuous non-increasing function of \( \psi \), with \( k_1(L_1, d) = \int_{d}^{u} [\frac{1}{\psi} - \frac{L(d, \theta)}{\psi_0}] \, d\theta \leq 0 \) and \( k_1(L_u, d) = 0 \), and \( k_2(\psi) \) is a continuous non-decreasing function of \( \psi \), with \( k_2(L_1, d) = 0 \) and \( k_2(L_u, d) = \int_{d}^{u} \frac{L(d, \theta)}{\psi_0} \, d\theta \geq 0 \). This implies that there is at least one \( \psi_0 \in [L_1, d', L_u, d] \) for which \( k_1(\psi_0) = k_2(\psi_0) \), which proves (95).

The proof of (96) is analogous.

The theory in this section allows us to describe a simple algorithm to calculate \( E(d, \psi) \) and \( E(d, \psi) \) for a decision \( d \in D \), if \( \psi \) is of the form (6). Hereto, we assume that \( D, \theta \) and \( L(d, \theta) \) are known \( (d \in D, \theta \in \Theta) \) and that \( \psi \) is of the form (6) with continuous \( l(\theta) \) and \( u(\theta) \) given, such that \( \int_{d}^{u} l(\theta) \, d\theta > 0 \) and \( \int_{d}^{u} u(\theta) \, d\theta < \omega \).
If $\Theta$ is bounded then it is possible to define a partition $\{\Theta_1, \ldots, \Theta_n\}$ of $\Theta$ such that all $\Theta_i$ have equal measure (so $\int d\theta = \int d\theta$ for all $i, j \in \{1, \ldots, n\}$). The integral over $\Theta_i$ for a function of interest is approximated by the value of that function in one point $\theta_i \in \Theta_i$ multiplied by $\int d\theta$. If $n$ is large this approximation will be reasonable. If $\Theta$ is unbounded, we must define bounds such that the integral over the bounded subspace is a good approximation of the integral over $\Theta$. This is a standard method of numerical integration.

Quite informally an algorithm for the determination of $\mathcal{E}(d, \Pi)$ for a decision $d \in D$ can be described by the following steps (given $d$):

1. Determine $L_1, d$ and $L_u, d$ (in many cases approximations can be used).

2. Take a $u \in [L_1, d', L_u, d']$. Calculate $\int d\theta \pi^{u, *}_{d, u, q}(\theta) d\theta$, if necessary numerically by using a partition $\{\Theta_i | i = 1, \ldots, n\}$ as described above, calculating $L(d, \theta, i)$ for some $\theta_i \in \Theta_i$ for all $i$, multiplying this by $u(\theta_i)$ if $L(d, \theta, i) > u$ or by $\ell(\theta_i)$ if $L(d, \theta, i) < u$, and approximating $\mathcal{E}(d, \pi^{u, *}_{d, u, q})$ by the sum of these terms divided by the sum of the $u(\theta_i)$ and $\ell(\theta_i)$ used herein. If $L(d, \theta, i) = u$, then we have to perform two calculations, one using $u(\theta_i)$ and one using $\ell(\theta_i)$, and afterwards take the largest of the corresponding results.

3. Compare the (approximate) value of $\int d\theta \pi^{u, *}_{d, u, q}(\theta) d\theta$ to $u$ and use the results of theorem 3.4 to determine whether these steps must be repeated using a new value $\omega(\theta, L_u, d')$ or $\omega(\theta, L_1, d')$, or the calculated value is the maximum we are looking for (or a good approximation to it). If the maximum is not yet found, repeat steps 2 and 3 until the maximum is found (or an approximation).

During this calculation of $\mathcal{E}(d, \Pi)$ we can also perform the necessary steps to calculate $\mathcal{E}(d, \Pi)$.

To end this section, we present a simple example of the above theory.
Example 3.1

Let, for a given decision $d \in D$, the loss function be equal to $L(d, \theta) = \theta (1-\theta)$, with $\theta \in [0,1]$. We want to determine, for this decision $d$, $E_2(d, \Pi)$ and $E_1(d, \Pi)$, with $\Pi$ given according to (6), with $l(\theta) = 1$ and $u(\theta) = 2$ for all $\theta \in [0,1]$. As $L(d, \theta)$ is not constant on an interval of positive length, this loss function satisfies restriction (9) and theorem 3.1 holds. Therefore, we can restrict to probability densities of the form (12) to determine $E_2(d, \Pi)$ and of the form (13) to determine $E_1(d, \Pi)$.

For this example, we have $L_1, d = 0, L_1, d = 1/4$ and, for $\omega \in [0,1/4]$, $\Theta_1(d, \omega) = (1-\sqrt{1-4\omega})/2, (1+\sqrt{1-4\omega})/2)$, $\Theta_1(d, \omega) = [0, (1-\sqrt{1-4\omega})/2) \cup ((1+\sqrt{1-4\omega})/2, 1]$ and $\Theta_1(d, \omega) = ((1-\sqrt{1-4\omega})/2, (1+\sqrt{1-4\omega})/2)$, from which it is clear that indeed restriction (9) holds.

Further, $C_1(d, \omega) = 1+\sqrt{1-4\omega}$ and $C_1(d, \omega) = 2-\sqrt{1-4\omega}$. Using theorem 3.1 we can determine $E_2(d, \Pi)$ by maximization of $E_L(d, \Pi, d) = \frac{1 + (1+2\omega)\sqrt{1-4\omega}}{6(1+\sqrt{1-4\omega})}$, leading to $E_2(d, \Pi) = 0.1875$ (for $\omega = 0.1875$) and $E_2(d, \Pi)$ by minimization of $E_L(d, \Pi, d) = \frac{2 - (1+2\omega)\sqrt{1-4\omega}}{6(2-\sqrt{1-4\omega})}$, leading to $E_2(d, \Pi) = 0.1435$ (for $\omega = 0.1435$).

The fact that the optima are equal to the values of $\omega$ for which these optima are adopted is in agreement with theorem 3.4.
References


Boole, G. (1854) An investigation of the laws of thought on which are founded the mathematical theories of logic and probabilities (Macmillan, London; reprinted by Dover (1958)).


<table>
<thead>
<tr>
<th>Number</th>
<th>Month</th>
<th>Author</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>92-01</td>
<td>January</td>
<td>F.W. Steutel</td>
<td>On the addition of log-convex functions and sequences</td>
</tr>
<tr>
<td>92-02</td>
<td>January</td>
<td>P. v.d. Laan</td>
<td>Selection constants for Uniform populations</td>
</tr>
<tr>
<td>92-04</td>
<td>February</td>
<td>H.J.C. Huijberts H. Nijmeijer</td>
<td>Strong dynamic input-output decoupling: from linearity to nonlinearity</td>
</tr>
<tr>
<td>92-05</td>
<td>March</td>
<td>S.J.L. v. Eijndhoven J.M. Soethoudt</td>
<td>Introduction to a behavioral approach of continuous-time systems</td>
</tr>
<tr>
<td>92-06</td>
<td>April</td>
<td>P.J. Zwietering E.H.L. Aarts J. Wessels</td>
<td>The minimal number of layers of a perceptron that sorts</td>
</tr>
<tr>
<td>92-07</td>
<td>April</td>
<td>F.P.A. Coolen</td>
<td>Maximum Imprecision Related to Intervals of Measures and Bayesian Inference with Conjugate Imprecise Prior Densities</td>
</tr>
<tr>
<td>92-08</td>
<td>May</td>
<td>I.J.B.F. Adan J. Wessels W.H.M. Zijm</td>
<td>A Note on “The effect of varying routing probability in two parallel queues with dynamic routing under a threshold-type scheduling”</td>
</tr>
<tr>
<td>92-10</td>
<td>May</td>
<td>P. v.d. Laan</td>
<td>Subset Selection: Robustness and Imprecise Selection</td>
</tr>
<tr>
<td>92-11</td>
<td>May</td>
<td>R.J.M. Vaessens E.H.L. Aarts J.K. Lenstra</td>
<td>A Local Search Template (Extended Abstract)</td>
</tr>
<tr>
<td>92-12</td>
<td>May</td>
<td>F.P.A. Coolen</td>
<td>Elicitation of Expert Knowledge and Assessment of Imprecise Prior Densities for Lifetime Distributions</td>
</tr>
<tr>
<td>92-13</td>
<td>May</td>
<td>M.A. Peters A.A. Stoorvogel</td>
<td>Mixed $H_2/H_\infty$ Control in a Stochastic Framework</td>
</tr>
<tr>
<td>Number</td>
<td>Month</td>
<td>Author</td>
<td>Title</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>92-14</td>
<td>June</td>
<td>P.J. Zwietering, E.H.L. Aarts, J. Wessels</td>
<td>The construction of minimal multi-layered perceptrons: a case study for sorting</td>
</tr>
<tr>
<td>92-15</td>
<td>June</td>
<td>P. van der Laan</td>
<td>Experiments: Design, Parametric and Nonparametric Analysis, and Selection</td>
</tr>
<tr>
<td>92-16</td>
<td>June</td>
<td>J.J.A.M. Brands, F.W. Steutel, R.J.G. Wilms</td>
<td>On the number of maxima in a discrete sample</td>
</tr>
<tr>
<td>92-17</td>
<td>June</td>
<td>S.J.L. v. Eijndhoven, J.M. Soethoudt</td>
<td>Introduction to a behavioral approach of continuous-time systems part II</td>
</tr>
<tr>
<td>92-18</td>
<td>June</td>
<td>J.A. Hoogeveen, H. Oosterhout, S.L. van der Velde</td>
<td>New lower and upper bounds for scheduling around a small common due date</td>
</tr>
<tr>
<td>92-19</td>
<td>June</td>
<td>F.P.A. Coolen</td>
<td>On Bernoulli Experiments with Imprecise Prior Probabilities</td>
</tr>
<tr>
<td>92-20</td>
<td>June</td>
<td>J.A. Hoogeveen, S.L. van de Velde</td>
<td>Minimizing Total Inventory Cost on a Single Machine in Just-in-Time Manufacturing</td>
</tr>
<tr>
<td>92-21</td>
<td>June</td>
<td>J.A. Hoogeveen, S.L. van der Velde</td>
<td>Polynomial-time algorithms for single-machine bicriteria scheduling</td>
</tr>
<tr>
<td>92-22</td>
<td>June</td>
<td>P. van der Laan</td>
<td>The best variety or an almost best one? A comparison of subset selection procedures</td>
</tr>
<tr>
<td>92-23</td>
<td>June</td>
<td>T.J.A. Storcken, P.H.M. Ruys</td>
<td>Extensions of choice behaviour</td>
</tr>
<tr>
<td>92-25</td>
<td>July</td>
<td>P.J. Zwietering, E.H.L. Aarts, J. Wessels</td>
<td>Exact Classification With Two-Layered Perceptrons</td>
</tr>
<tr>
<td>92-26</td>
<td>July</td>
<td>M.W.P. Savelsbergh</td>
<td>Preprocessing and Probing Techniques for Mixed Integer Programming Problems</td>
</tr>
<tr>
<td>Number</td>
<td>Month</td>
<td>Author</td>
<td>Title</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>---------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>92-27</td>
<td>July</td>
<td>I.J.B.F. Adan, W.A. van de Waarenburg, J. Wessels</td>
<td>Analysing $E_k</td>
</tr>
<tr>
<td>92-28</td>
<td>July</td>
<td>O.J. Boxma, G.J. van Houtum</td>
<td>The compensation approach applied to a $2 \times 2$ switch</td>
</tr>
<tr>
<td>92-29</td>
<td>July</td>
<td>E.H.L. Aarts, P.J.M. van Laarhoven, J.K. Lenstra, N.L.J. Ulder</td>
<td>Job Shop Scheduling by Local Search</td>
</tr>
<tr>
<td>92-30</td>
<td>August</td>
<td>G.A.P. Kindervater, M.W.P. Savelsbergh</td>
<td>Local Search in Physical Distribution Management</td>
</tr>
<tr>
<td>92-31</td>
<td>August</td>
<td>M. Makowski, M.W.P. Savelsbergh</td>
<td>MP-DIT Mathematical Program data Interchange Tool</td>
</tr>
<tr>
<td>92-32</td>
<td>August</td>
<td>J.A. Hoogeveen, S.L. van de Velde, B. Veltman</td>
<td>Complexity of scheduling multiprocessor tasks with prespecified processor allocations</td>
</tr>
<tr>
<td>92-33</td>
<td>August</td>
<td>O.J. Boxma, J.A.C. Resing</td>
<td>Tandem queues with deterministic service times</td>
</tr>
<tr>
<td>92-34</td>
<td>September</td>
<td>J.H.J. Einmahl</td>
<td>A Bahadur-Kiefer theorem beyond the largest observation</td>
</tr>
<tr>
<td>92-35</td>
<td>September</td>
<td>F.P.A. Coolen</td>
<td>On non-informativeness in a classical Bayesian inference problem</td>
</tr>
<tr>
<td>92-36</td>
<td>September</td>
<td>M.A. Peters</td>
<td>A Mixed $H_2/H_\infty$ Function for a Discrete Time System</td>
</tr>
<tr>
<td>92-37</td>
<td>September</td>
<td>I.J.B.F. Adan, J. Wessels</td>
<td>Product forms as a solution base for queueing systems</td>
</tr>
<tr>
<td>92-38</td>
<td>September</td>
<td>L.C.G.J.M. Habets</td>
<td>A Reachability Test for Systems over Polynomial Rings using Gröbner Bases</td>
</tr>
<tr>
<td>92-39</td>
<td>September</td>
<td>G.J. van Houtum, I.J.B.F. Adan, J. Wessels, W.H.M. Zijm</td>
<td>The compensation approach for three or more dimensional random walks</td>
</tr>
<tr>
<td>Number</td>
<td>Month</td>
<td>Author</td>
<td>Title</td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>----------------</td>
<td>----------------------------------------------------------------</td>
</tr>
<tr>
<td>92-40</td>
<td>September</td>
<td>F.P.A. Coolen</td>
<td>Bounds for expected loss in Bayesian decision theory with imprecise prior probabilities</td>
</tr>
</tbody>
</table>