Nonparametric Predictive Utility Inference

B. Houlding[a,*], F.P.A. Coolen[b]

[a]*Discipline of Statistics, Trinity College Dublin, Ireland.*
[b]*Department of Mathematical Sciences, Durham University, UK.*

Abstract

We consider the natural combination of two strands of recent statistical research, *i.e.*, that of decision making with uncertain utility and that of Nonparametric Predictive Inference (NPI). In doing so we present the idea of Nonparametric Predictive Utility Inference (NPUI), which is suggested as a possible strategy for the problem of utility induction in cases of extremely vague prior information. An example of the use of NPUI within a motivating sequential decision problem is also considered for two extreme selection criteria, *i.e.*, a rule that is based on an attitude of extreme pessimism and a rule that is based on an attitude of extreme optimism.

Keywords: Decision Analysis, Nonparametrics, Predictive Inference, Exchangeability, Uncertain Utility.

1. Introduction

The Bayesian paradigm, *e.g.*, de Finetti (1974), coupled with the expected utility hypothesis of Bernoulli (1738), provides a transparent and attractive methodology for solving problems of decision making under uncertainty. In this approach preferences over a set of possible decisions are reconstructed by taking into account both the probability that each decision leads to a particular outcome, and the relative preference for obtaining that outcome as measured by its utility value. Furthermore, if the outcome that pertains from a particular decision depends on the value of an unknown random quantity, then the probabilities associated with the set of possible decision outcomes are typically assumed to be subject to an assigned prior parametric distribution. Learning then occurs following observation of data that has a probabilistic dependence with the unknown random quantity of interest, and the usual 'posterior is proportional to likelihood times prior' of Bayes' Theorem is employed.

However, implicit within this theory (and hence necessary for its application) is the assumption that the Decision Maker (DM) knows her preferences, meaning that she can assign an appropriate utility function (with domain the full set of all possible decision outcomes) for use within the problem. In application this is usually achieved by either assuming a fixed utility form, *e.g.*, a logarithmic utility function for monetary returns, or by selecting specific utility values for particular and relevant decision outcomes. As such, classical Bayesian subjective expected utility theory does not permit inherent uncertainty in preferences over decisions. It also does not allow the learning of utility and specifies that the DM will never be surprised by the utility of a realized outcome.

Nevertheless, not for all situations is the assumption of a known preference relation over outcomes deserved, and often a DM may instead need to learn her preferences through suitable experimentation. Indeed, a DM may consider it inappropriate to assign a particular and fixed utility value for any outcome that is novel or unfamiliar, choosing instead to only do so after direct experience or exposure. Such cases of utility uncertainty motivate so-called adaptive utility theory, *e.g.*, Cyert and DeGroot (1975), Houlding (2008), and Houlding and Coolen (2007, 2011), which generalizes the traditional utility concept by only

*Corresponding author, address: Rm 105, Discipline of Statistics, Lloyd's Institute, Trinity College Dublin, Dublin 2, Ireland. E-mail: brett.houlding@tcd.ie, tel: +353 (0) 1896 2933, fax: +353 (0) 1 677 0711.

requiring the utility function be known up to the value of some uncertain utility parameter. The principal idea of adaptive utility is then to treat the uncertain utility parameter in the same manner that unknown random quantities are typically treated in standard Bayesian statistical inference, *i.e.*, they are subjected to a parametric learning model in accordance with Bayes' Theorem. Yet, and despite adaptive utility theory explicitly permitting a DM to remain uncommitted to a presumed known and correct utility function, its previous use has required knowledge of a precise and meaningful prior distribution concerning true preferences, something that is unlikely to be considered either reasonable or justifiable when selecting from outcomes that include (initially) new and foreign possibilities.

Rather than assuming a precise prior distribution over an uncertain utility parameter, interest here is in the use of the Nonparametric Predictive Inference (NPI) technique of Coolen (1996, 2006), and Augustin and Coolen (2004), which is a low structure statistical technique arising naturally as a result of Hill's $A_{(n)}$ assumption, Hill (1968, 1988, 1993). Given a known ordered series of utility values that are considered subject to a post-data exchangeability assumption, NPUI proceeds by assigning equal mass to the probability that a new utility value falls within any of the intervals formed by the known ordered utility, leading to the quantification of such utility uncertainty through interval probability.

An outline of the remainder is as follows; Section 2 reviews the concept of uncertain utility and briefly describes how this can be taken into account through adaptive utility theory, whilst Section 3 reviews the statistical technique of NPI. In Section 4 the two strands of uncertain utility and nonparametric predictive inference are combined to formally introduce the NPUI model, with an illustrative example being given in Section 5. Finally, Section 6 concludes with a discussion of possible future directions.

2. Uncertain Utility

The assumption that a DM can accurately identify the utility value of any considered decision outcome is prevalent within the theory and application of Bayesian decision making under uncertainty. Often this will be through the use of an assumed model for the utility of outcomes from a continuous domain, *e.g.*, a logarithmic model for monetary returns. Alternatively, specific utility values may be considered appropriate following suitable introspection concerning the decision problem. In either case, the possibility that a DM may not *a priori* know their preference for a given outcome is often ignored.

In contrast, the theory of adaptive utility, as introduced by Cyert and DeGroot (1975) and further developed by Houlding (2008) and Houlding and Coolen (2007, 2011), generalises classical Bayesian decision theory, suggesting a methodology for decision selection even when the DM is unable to fully specify her preferences. In this setting the DM is permitted to be uncertain over her true preferences, with the appropriate utility function over decision outcomes only being known up to the value of some uncertain parameter θ. Such a parameter is referred to as the DM's *state of mind*, and it can be used to model uncertainty over any aspect of the DM's preferences, with interesting examples include vectors of unknown trade-off weights or unknown level of risk aversion. Notationally, such a dependence between the utility function $u(\cdot)$ and the state of mind θ was displayed via the inclusion of a conditioning argument, *e.g.*, $u(\cdot|\theta)$.

Instead of assuming that the utility function is fully known, adaptive utility theory makes use of a probabilistic specification concerning the uncertain state of mind θ. Bayesian updating can then occur once the DM receives additional information concerning her true preferences, and previous examples considered for such utility related information include noise corrupted observations of the true utility value, or the sign of the difference between the prior expectation of the utility value for a given outcome and the utility value that was actually received (*i.e.,* an indication of elation or disappointment). An adaptive utility function ${}_a u(\cdot)$ was then defined as the expectation of the possible utility values with respect to beliefs over the state of mind, *i.e.*, ${}_a u(\cdot) = E_\theta[u(\cdot|\theta)]$.

In essence, the use of adaptive utility theory is analogous to the use of a hierarchical prior within robust Bayesian analysis, *e.g.*, Berger (1993). A utility value, once scaled to fall within the interval $[0, 1]$, corresponds to a probability, with the utility of decision outcome o being that probability p which makes the DM indifferent between receiving o for sure, or selecting the decision which results in most preferred outcome o^* with probability p and least preferred outcome o_* otherwise, see DeGroot (1970). Adaptive utility theory allows the DM to be uncertain of the value of p that results in her indifference, instead considering a non-degenerate prior subjective probability distribution for its value.

That such a probability distribution concerning utility values may change following updating in light of additional information motivates the name adaptive utility theory, for if the probability distribution did alter, then so would the expected utility return, *i.e.*, the adaptive utility value. In other words, the adaptive utility value of an unfamiliar decision outcome will 'adapt' in light of additional information concerning preferences. This is in contrast to classical utility theory in which the utility values of all decision outcomes are considered known and fixed. Yet, explicitly permitting utility uncertainty does not alter the suggested decision selection within a one-off decision problem, as the strategy arising from the adaptive utility setting is the same as that which would arise from traditional theory if the adaptive utility values were assumed equal to the true utility values. However, in a sequential decision problem the acceptance that certain decision outcomes do not have known utility can alter the optimal selection strategy.

Notable alternative theories that similarly seek to incorporate uncertain preference within a decision making paradigm include the approaches of Farrow and Goldstein (2006) and Ben-Haim *et al.* (2009). Rather than expressing uncertainty over an unknown utility parameter through the use of a precise prior probability distribution, Farrow and Goldstein allow the DM to remain non-committed and to instead only provide an upper and lower bound for the true utility value (through the declaration of lower and upper bounds on trade-off parameters in a multi-attribute utility hierarchy). In this respect the approach of Farrow and Goldstein is similar to the NPUI approach presented here. However, differences result in that the use of the NPI statistical technique would appear to be a simpler approach that is data driven. In particular, the NPUI approach considered here does not require any explicit declaration of subjective judgments and/or expressions concerning the utility of a novel outcome other than that of a usually reasonable and objective post-data exchangeability assumption.

In contrast, the approach of Ben-Haim *et al.* is to instead specify an Info-Gap model of utility uncertainty, and is relevant for situations of severe uncertainty whereby, other than the specification of a best point-estimate guess, nothing else can be elicited. In particular, there is no probabilistic specification of how accurate such a best point-estimate guess may be. Instead a nested subset of possible horizons of uncertainty is specified and the decision selected which is deemed most robust in that it guarantees a specified minimum critical return for the largest horizon of uncertainty. Unlike the imprecise theory of Farrow and Goldstein, or the NPI method used here, the Info-Gap approach of Ben-Haim *et al.* is non-probabilistic, and can not quantify, or even given bounds on, the probability of outcome for any quantity that is subject to 'Info-Gapping'.

3. Nonparametric Predictive Inference

The implicit assumption underlying adaptive utility, *i.e.*, that there exists an available and precise prior distribution representing the DM's possible preferences, may be unreasonable in situations where there exists an opportunity of receiving unexperienced or novel outcomes. Instead it may be more appropriate to permit a less restricting belief model that is closer to resembling a state of ignorance, and one such possibility is that of Nonparametric Predictive Inference (NPI).

NPI is a low structure statistical technique that is predictive by nature and arises naturally as a result of Hill's $A_{(n)}$ assumption, which in turn only assumes exchangeability at the level of the pre-observation random variables associated with the data. Note that two random variables X and Y are exchangeable if and only if, for all values x and y, the relation $P(X = x, Y = y) = P(X = y, Y = x)$ holds (the concept generalizes straightforwardly for more than two random variables). As such, exchangeability is an assumption of symmetry that is particularly suitable in situations where draws are randomly taken from a population (which can be finite or infinite), hence making it an appropriate assumption in many situations. Indeed, the common statistical assumption of sampling independent and identically distributed random variables would clearly result in exchangeability, but so would sampling without replacement, which is not independent.

The assumption $A_{(n)}$ is as follows: Let real-valued $x_{(1)} < \cdots < x_{(n)}$ be the ordered statistics of data $x_1, \ldots, x_n$, and let X_i be the corresponding pre-data random quantities, then under $A_{(n)}$:

1. The observable random quantities $X_1, \dots, X_n$ are exchangeable. Note this aspect was not included in Hill (1968), but here interest lies in the more specific case in which it is true.
2. Ties have probability 0, so $x_i \neq x_j$ for all $i \neq j$, almost surely (generalization to include ties is straightforward, but requires more awkward notation).
3. Given data $x_1, \dots, x_n$, the probability that the next observation X_{n+1} falls in the open interval $I_j = (x_{(j-1)}, x_{(j)})$ is $1/(n+1)$ for $j = 1, \dots, n+1$, with the definition that $x_{(0)} = -\infty$ and $x_{(n+1)} = \infty$.

As such, $A_{(n)}$ is a nonparametric (distribution free) post-data assumption that is inductive by nature and which may be relevant when there is a lack of additional information further to the data itself. Its use may be considered a more measured or less presumptuous alternative for making inference over a future observation than the direct specification of conditional independencies and specific distributional forms, especially if these are difficult to justify due to the existence of only extremely vague prior information concerning the underlying distribution.

Hence as a rule for statistical inference, NPI is 'frequentist' by nature, allowing predictions to be made over future events without the specification of a prior distribution. Indeed, as NPI is a data driven approach, it is a natural tool for consideration in situations where the specification of subjective beliefs would be difficult or inappropriate. For example, Coolen *et al.* (2011) considered NPI for inference in studying aspects of the size and composition of juries so as to determine their representativeness of the views of society.

Nevertheless, Hill (1988, 1993) demonstrated that $A_{(n)}$ does fully coincide with the general framework of Bayesian statistics and the use of a finitely additive prior, whilst Augustin and Coolen (2004) were able to relate $A_{(n)}$ and the resulting NPI to the theory of interval or imprecise probability. In the case of interval probability, $A_{(n)}$ was used to provide bounds on the probability of a future event (rather than the specification of a precise value), with the 'subjectivist' interpretation that such bounds represent lower and upper prices for a bet where the payout depends on whether or not the event of interest is seen to occur.

Further extensions include Coolen (1996) where NPI was considered for cases of two finitely exchangeable populations (where exchangeability is only assumed within populations and not between them), and Abellán *et al.* (2011) where algorithms were developed for obtaining maximum entropy for multinomial data. NPI has also been considered in relation to the concept of objective Bayesianism, *e.g.*, Coolen (2006). In particular, the assumption $A_{(n)}$ (and its variants) were shown to support the following two norms:

Empirical: Objective inferences should not disagree with empirical evidence.

Logical: If there is no information suggesting that one possible outcome is more likely than another, then this should be reflected by identical uncertainty quantifications for these outcomes.

For more information on NPI see `www.npi-statistics.com`

4. Nonparametric Predictive Utility Inference

Concerning the use of an exchangeability assumption for the pre-observation utility values of a collection of outcomes, it seems sensible to assume exchangeability at least at the level of collections of outcomes which are sensibly grouped, or which fall under the same taxonomic category, *e.g.*, cereal brands, or fruits *etc.* Indeed, exchangeability is simply a formalization of the notion that the future is predictable on the basis of past experience, and would be well-suited for any situation where the DM believes that their past experience of joy or dislike from experiencing certain outcomes grouped under some categorization, is likely to inform them of the possibility of similarly experiencing joy or dislike from a new and novel outcome that is also listed under the same categorization.

However, whilst $A_{(n)}$ concerns the prediction of a random variable with domain $\mathbb{R}$, utility values are instead bound to a finite interval $[a, b]$, and this is to prevent any outcome from being deemed infinitely better (worse) than an alternative, meaning a DM would always accept (reject) any strategy that had positive probability of that outcome, regardless of how small that probability would be. In addition, as a utility function is only unique up to a positive linear transformation, this finite interval is often considered to be the unit interval $[0, 1]$.

The resulting problem is that, if it is agreed that the utility scaling should be on the unit interval $[0, 1]$, how should the utility values for experienced outcomes be placed on that scale? If it is known that the collection of experienced outcomes includes within it the best and worst possible outcomes from the taxonomic collection which is deemed to have exchangeable utility values, then this is achieved by specifying those outcomes to have utilities 1 and 0, respectively. The remaining experienced outcomes are then scored through the previously described process of hypothetical comparisons between obtaining that outcome with certainty, or to engage in the lottery in which the best outcome is obtained with probability p and the worst outcome otherwise. Furthermore, as we are discussing the experienced outcomes we assume that the DM is able to achieve this.

Nevertheless, once we allow the possibility of novel outcomes with unknown utility, then it would appear restricting to argue that no future outcome could be considered better (worse) than all previously experienced outcomes. This is a general problem of modelling uncertain utility, and to the best of our knowledge, there is not yet any practical solution that does not rely on similarly unreasonable assumptions. Previous possibilities have been to include scaling subject to commensurability assumptions in which the same outcome is always deemed most preferable, *e.g.*, in Boutilier (2003) and Houlding and Coolen (2011), but this does not allow uncertainty as to which outcome is most preferred. As such, we here suggest the interpretation that the utilities of experienced outcomes have been placed on the $[0, 1]$ scale by considering 'hypothetical' best and worst possible outcomes that could exist within the exchangeable collection and by assigning those hypothetical outcomes to have utilities 1 and 0, respectively.

Under the above assumption, let $u_{(1)}, u_{(2)}, \ldots, u_{(n)}$, with $0 < u_{(i)} \leq u_{(j)} < 1$ for $i < j$, be the ordered values of the known utilities $u_1, \ldots, u_n$ representing preferences over the set of distinct decision outcomes $\mathcal{O}_n = \{o_1, o_2, \ldots, o_n\}$. Furthermore, denote by $\mathcal{U}_n = \{U_1, \ldots, U_n\}$ the set of pre-observation random quantities which represent the utilities of the elements within $\mathcal{O}_n$ before they are experienced, and suppose that the elements within $\mathcal{U}_n$ are considered exchangeable. Then given a new and novel outcome o_{new} whose utility value $U_{new} \in (0, 1)$ is unknown but considered exchangeable with the elements of $\mathcal{U}_n$, the NPUI model considered here states only the following:

$$\begin{aligned} P\Big(U_{new} \in [u_{(i)}, u_{(i+1)}]\Big) &= P\Big(U_{new} \in (0, u_{(1)}]\Big) \\ &= P\Big(U_{new} \in [u_{(n)}, 1)\Big) \\ &= \frac{1}{n+1} \end{aligned} \tag{1}$$

In contrast to Hill's $A_{(n)}$ assumption, which concerns observations on the entire real line, the NPUI model discussed here refers to observations that are restricted to the open interval $(0, 1)$. To make clear such a distinction we will denote the assumptions underlying NPUI by $UA_{(n)}$, where the inclusion of the 'U' can be understood as indicating both that we are concerned with utility inference and that we are restricted to the unit interval.

Now, denoting the lower and upper bounds for the expected utility of outcome o_{new} as $\underline{E}[U_{new}]$ and $\overline{E}[U_{new}]$, respectively, the NPUI model and Equation (1) leads to the following:

$$\underline{E}[U_{new}] = \frac{1}{n+1}\sum_{i=1}^{n} u_{(i)} = \frac{1}{n+1}\sum_{i=1}^{n} u_i \tag{2}$$

$$\begin{aligned} \overline{E}[U_{new}] &= \frac{1}{n+1}\Big(1 + \sum_{i=1}^{n} u_{(i)}\Big) = \frac{1}{n+1}\Big(1 + \sum_{i=1}^{n} u_i\Big) \\ &= \frac{1}{n+1} + \underline{E}[U_{new}] \end{aligned} \tag{3}$$

A straightforward result arising from Equations (2) and (3) is that the difference between the upper and lower expected utility bounds for the new outcome o_{new}, to be denoted as $\Delta\Big(E[U_{new}]\Big)$, is the following:

$$\Delta\Big(E[U_{new}]\Big) \quad = \quad \overline{E}[U_{new}] - \underline{E}[U_{new}] = \frac{1}{n+1} \tag{4}$$

Equation (4) is an intuitive result that shows how the NPUI model decreases the range of possible values for the expected utility of a new and novel outcome as the number of known utility values for alternative outcomes is increased, or in other words, as past experience is increased. At the extreme of having observed no alternative outcomes whose utility could be considered exchangeable with the utility of the new outcome, the model simply states that the expected utility for the new outcome can take any value in the range $(0, 1)$, whilst in the limiting case of having an infinite number of known and exchangeable alternatives, the bounds for the expected utility value for the new outcome coincide. Of course that is not to say the actual utility for the new outcome becomes known, only that the expected utility value is then identified.

The NPUI model of Equation (1) offers a transparent and robust method for representing utility uncertainty that does not require a presumed known and possibly arbitrary prior distribution. In situations where most of the known utility values are relatively small, the NPUI model gives low valued expected utility bounds for a novel outcome, suggesting that it may not be appropriate to experiment in a pairwise choice against outcomes with known high utility. Similarly, in situations where most of the known utilities are relatively high, with only a few outcomes known to have small utility, the NPUI model gives large valued expected utility bounds.

In the case of a sequential decision problem in which there is more than one novel outcome available for selection, the results of Augustin and Coolen (2004) for updating within the NPI setting become applicable. Let o_{new_2} be a second novel outcome whose unknown utility U_{new_2} is assumed exchangeable with the set $\mathcal{U}_n \bigcup \{U_{new}\}$, and denote $(0, u_{(1)}]$ as I_1, $[u_{(i)}, u_{(i+1)}]$ as I_{i+1}, and $[u_{(n)}, 1)$ as I_{n+1}. Then if it is discovered that U_{new} has value u_{new} which falls in the interval I_j, this interval is split into the two distinct intervals with either lower or upper boundary u_{new}. Denoting these additional intervals as I_{lj} and I_{uj}, the assumption $UA_{(n)}$ is replaced by the stronger assumption $UA_{(n+1)}$ stating that U_{new_2} is equally likely, with probability $1/(n + 2)$, to fall in any of $I_1, \ldots, I_{j-1}, I_{lj}, I_{uj}, I_{j+1}, \ldots, I_{n+1}$.

A result of such updating is the following:

$$\begin{aligned}
\underline{E}[U_{new_2}|u_{new}] &= \frac{1}{n+2}\Big(\sum_{i=1}^{n} u_i + u_{new}\Big) \\
&= \frac{n+1}{n+2}\underline{E}[U_{new}] + \frac{u_{new}}{n+2}
\end{aligned} \tag{5}$$

$$\begin{aligned}
\overline{E}[U_{new_2}|u_{new}] &= \frac{1}{n+2}\Big(1 + \sum_{i=1}^{n} u_i + u_{new}\Big) \\
&= \frac{n+1}{n+2}\overline{E}[U_{new}] + \frac{u_{new}}{n+2} \\
&= \frac{1}{n+2} + \frac{n+1}{n+2}\underline{E}[U_{new}] + \frac{u_{new}}{n+2} \\
&= \frac{1}{n+2} + \underline{E}[U_{new_2}|u_{new}]
\end{aligned} \tag{6}$$

Furthermore,

$$\begin{aligned}
\Delta\Big(E[U_{new_2}|u_{new}]\Big) &= \overline{E}[U_{new_2}|u_{new}] - \underline{E}[U_{new_2}|u_{new}] \\
&= \frac{n+1}{n+2}\underline{E}[U_{new}] + \frac{u_{new}}{n+2} - \frac{n+1}{n+2}\overline{E}[U_{new}] - \frac{u_{new}}{n+2} \\
&= \frac{n+1}{n+2}\Delta\Big(E[U_{new}]\Big) = \frac{1}{n+2}
\end{aligned} \tag{7}$$

Equations (5) and (6) show how the updated expected utility bounds are both a weighted sum of their value before knowledge of u_{new} and of the actual value u_{new} is seen to take. Such updating rules appear intuitive. For example, if $u_{new} \in (0, \underline{E}[U_{new}])$, then both the lower and upper expected utility bounds are reduced as a result of the updating. Alternatively, if $u_{new} \in (\overline{E}[U_{new}], 1)$, then the result of updating is to increase both the lower and upper expected utility bounds, whilst in the situation of $u_{new} \in (\underline{E}[U_{new}], \overline{E}[U_{new}])$, updating leads to an increase in the lower expected utility bound and a decrease in the upper expected utility bound. Finally, it should be noted that the weights in Equations (5) and (6) indicate that new observations have diminishing effect on the updated expected utility bounds as the number of known utilities increases.

In a sequential decision context, myopic decision making is avoided by taking into account the influence that early decision choices will have upon later decisions. From the view point of the first decision in the sequence, the actual value of currently uncertain utilities will be unknown. However, by the time of a later decision epoch t, some novel outcomes may have been experienced and the uncertainty concerning their utilities removed. If indeed novel outcomes had been experienced by epoch t, then the expected utility bounds of residual novel outcomes will not be governed by Equations (2) and (3), but instead by Equations (5) and (6). Unfortunately, at the planning stage before any decision is made, all novel outcomes will have unknown utilities; hence to consider the possibility that some of these will be known in the future, the following is required:

$$\underline{E}[U_{new_2} | U_{new} \in I_j] = \frac{1}{n+2}\Big(\sum_{i=1}^{n} u_i + \inf(I_j)\Big) \tag{8}$$

$$\overline{E}[U_{new_2} | U_{new} \in I_j] = \frac{1}{n+2}\Big(1 + \sum_{i=1}^{n} u_i + \sup(I_j)\Big) \tag{9}$$

The conditioning arguments of Equations (8) and (9) reflect the fact that initially o_{new} has uncertain utility U_{new}, but that when o_{new} is obtained, U_{new} will be observed to fall in one of the intervals I_1 to I_{n+1}. The actual value that U_{new} will take within any particular interval I_j will of course influence future decisions, but the NPUI model makes no specification on this prior to its observation (as we wish to avoid making any *a priori* distributional assumptions for which there is no apparent justification), instead only specifying $P(U_{new} \in I_j)$.

An internal consistency result that arises straightforwardly from Equations (8) and (9) is the following:

$$\underline{E}[U_{new_2}] = \sum_{j=1}^{n+1} \underline{E}[U_{new_2} | U_{new} \in I_j] P(U_{new} \in I_j) \tag{10}$$

$$\overline{E}[U_{new_2}] = \sum_{j=1}^{n+1} \overline{E}[U_{new_2} | U_{new} \in I_j] P(U_{new} \in I_j) \tag{11}$$

5. Example

We now illustrate how the NPUI concept may be incorporated within a sequential decision problem, and so as to highlight the differences that may occur as a result of the choice rule employed, we consider the use of a rule that is based on an attitude of Extreme Pessimism (EP), and a rule that is based on an attitude Extreme Optimism (EO):

EP: Under extreme pessimism the DM will always select the outcome or sequential decision path whose lower expected utility bound is greatest. Furthermore, if in a sequential problem a strategy has to be decided for making decisions based on future knowledge of the interval to which a currently uncertain utility belongs to, then it will always be assumed that the uncertain utility will equal the infimum of the interval under consideration.

EO: Under extreme optimism the DM will always select the outcome or sequential decision path whose upper expected utility bound is greatest. Furthermore, if in a sequential problem a strategy has to be decided for making decisions based on future knowledge of the interval to which a currently uncertain utility belongs to, then it will always be assumed that the uncertain utility will equal the supremum of the interval under consideration.

The EP and EO choice rules can be identified with the usual Maximin and Maximax decision selection criteria, respectively, see for example Wald (1950). Nevertheless, there does exist a slight difference in that both EP and EO are sequential in nature and make additional assumptions concerning the treatment of decisions that are to be made following the future observation of an initially uncertain utility. Alternative sequential choice rules that are based on the characteristics of criteria such as opportunity loss (minimax regret) Savage (1954), the Hurwicz criterion, or satisficing *etc.*, will not be considered here. Kmietowicz and Pearman (1981) offer a review of such alternative non-probabilistic selection criteria.

Now suppose a DM is facing a situation in which she must sequentially select a fruit for consumption from amongst three available options. Only one of the three available options has previously been experienced by the DM, with both alternatives assumed unfamiliar. In addition, the DM's experience of fruit consumption is such that she has knowledge over her preferences for a collection of five different fruits with ordered labellings $f_1, \ldots, f_5$ (ordered so that fruit f_i has utility value $u_{(i)}$), and where the known utility values $u_{(1)}, \ldots, u_{(5)}$ are 0.3, 0.35, 0.4, 0.5, and 0.7, respectively. Figure 1 shows the DM's previous utility experience on a $[0, 1]$ utility line.

The two alternative and unexperienced fruits will be labeled f_{new} and f_{new_2}, and will be assumed to have exchangeable subjective utility (meaning the labellings are unimportant and that either can be taken if a decision is made to select one of them).

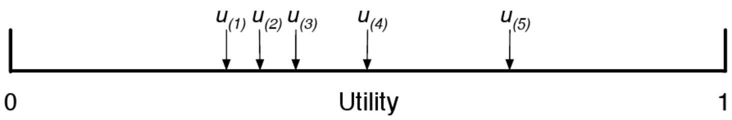

Figure 1: Assumed previous utility experience

Consider first the solution to this example if it were treated as a one-off solitary decision problem, rather than as a sequential problem. In such a setting, and using Equations (2) and (3), the lower expected utility bound for either of f_{new} or f_{new_2} is found to be $\underline{E}[U_{f_{new}}] = 3/8 = 0.375$, whilst the upper expected utility bound is $\overline{E}[U_{f_{new}}] = 13/24 \approx 0.542$. As such, under the EP choice rule the DM should only try one of the new fruits if the known alternative option is f_1, whilst the EO choice rule requires selection of an unfamiliar fruit unless the known alternative option is f_5. That is to say, the result of the NPUI model is that only f_1 is known to have a utility less than the lower expected utility bound of f_{new}, whilst only f_5 is known to have a utility greater than the upper expected utility bound of f_{new}.

Now suppose the situation is formally treated as a sequential problem with a planning horizon of three decision epochs. Furthermore, suppose that the DM's objective is to maximise the overall sum of utility returns generated by fruit consumption. A decision tree of this setting is provided in Figure 2. In such a sequential problem the DM must consider her possible updated expected utility bounds in light of the possibility of experiencing one of the new fruits at either the first or second decision epoch, and the way that this is performed under the NPUI model is provided for by Equation (8) in the case of the lower expected utility bound (the EP choice rule), and by Equation (9) in the case of the upper expected utility bound (the EO choice rule).

The solution to such a sequential decision problem can be determined given a choice rule for comparing uncertain utilities with known utilities, but once this has been identified, *e.g.*, by EP or EO, the usual 'roll-back' algorithm for solving a decision tree can be implemented. Furthermore, the number of calculations

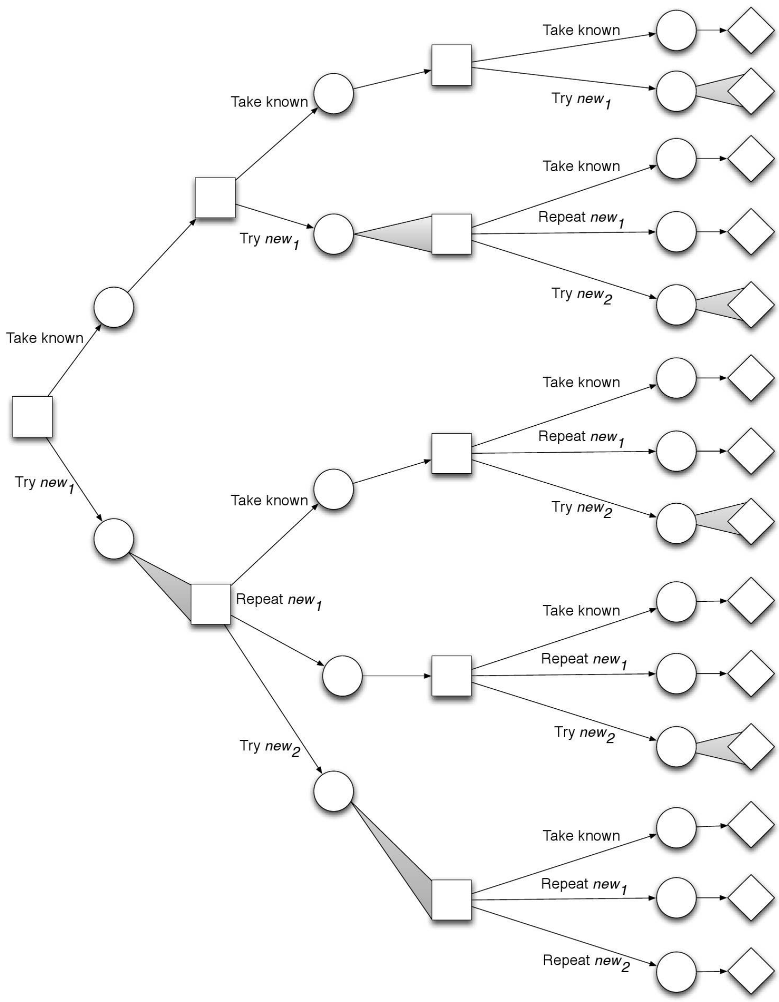

Figure 2: Decision tree for sequential selection problem with gray shadowing indicating a range of possible outcomes for the preceding choice node

required in performing this task can be reduced given the understanding that it would be irrational (never of positive benefit in terms of increased utility) to delay the selection of an uncertain utility outcome over a known utility outcome if such a selection was indeed to be made within the planning horizon. This is because we assume the utility for known outcomes to be fixed over the duration of the planning horizon, and are not altered by the sequence in which outcomes are experienced. As such, there would be a non-negative decrease to the resulting value of the information gained from selecting an uncertain utility outcome if it is selected later in the decision sequence, for if a novel outcome was found to have utility greater than those outcomes already experienced, its continued selection would provide additional utility in each subsequent epoch within the planning horizon. For the decision tree in Figure 2, this means that the following paths can all be ignored:

- Take known $\rightarrow$ Take known $\rightarrow$ Try new_1
- Take known $\rightarrow$ Try new_1 $\rightarrow \cdots$
- Try new_1 $\rightarrow$ Take known $\rightarrow$ Try new_2

- Try new_1 → Repeat new_1 → Try new_2

Two additional paths can also be ignored due to the irrationality in their selection:

- Try new_1 → Take known → Repeat new_1
- Try new_1 → Repeat new_1 → Take known

Figure 3 displays the reduced decision tree which only includes paths that are not *a priori* irrational. The solution to this decision tree is summarised in Table 1, which considers all possibilities for the available known fruit, with the actual known fruit available being indicated by the first column. The second column lists the lower bound of the expected utility for the optimal decision strategy under the EP choice rule, whilst the third column lists the upper bound of the expected utility for that decision strategy. The fourth and fifth columns list the lower and upper bounds for the expected utility of the optimal decision strategy under the EO choice rule, respectively, whilst finally columns six and seven indicate whether or not the optimal decision strategy under either the EP or EO choice rules includes the initial selection of an unfamiliar option. Note that when it is found optimal to not select a novel option, the lower and upper expected utility bounds for the choice rule under consideration will coincide, and this is because in each period the DM will select the experienced outcome that has known utility.

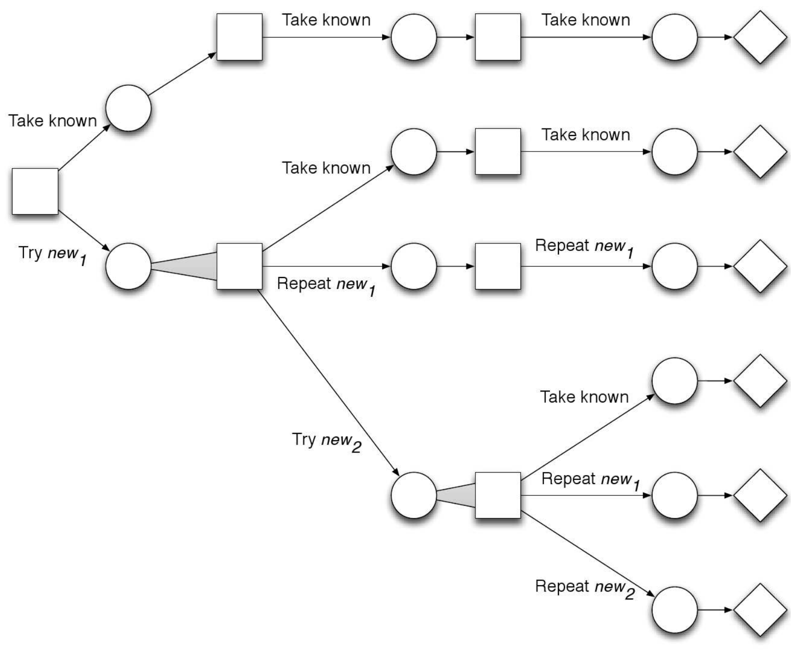

Figure 3: Reduced Decision tree for the selection problem with irrational paths omitted

At this point a note should be made concerning how the upper bound was found for the EP choice rule, and similarly, how the lower bound was found for the EO choice rule. Whilst the EP (EO) choice rule is specifically designed to find the decision strategy that has maximum lower (upper) expected utility bound, it does not take into consideration the upper (lower) expected utility bound of that decision strategy. To determine the upper (lower) bound for the EP (EO) choice rule, the decision strategy suggested remains fixed, but we consider the resulting utility from the selection of a novel outcome to be the supremum (infimum) of the utility interval I_j that it is assumed to fall in.

In addition, there are occasions when equality exists between the utility value for the known option and either the lower or upper utility bound for a novel option, yet the actual choice that is made will affect the

Bounds on Expected Utility Sum for Optimal Decision Strategy over 3 Periods						
	Pessimistic		Optimistic		Select a Novel Option	
Available	Lower	Upper	Lower	Upper	Pessimistic	Optimistic
f_1	1.298	1.817	1.298	1.817	Yes	Yes
f_2	1.305	1.819	1.305	1.819	Yes	Yes
f_3	1.323	1.785	1.319	1.826	Yes	Yes
f_4	1.500	1.500	1.367	1.855	No	Yes
f_5	2.100	2.100	2.100	2.100	No	No

Table 1: Expected utility bounds for sequential selection problem and optimal selection strategy

upper expected utility bound under the EP choice rule and the lower expected utility bound for the EO choice rule. To account for such problems an additional rule was included stating that an uncertain utility option is selected in cases of indifference under the EP decision strategy, whilst a known utility option is selected in cases of indifference for the EO strategy. This would appear instinctive, as under the EP choice rule the lower expected utility bound for the novel option is calculated on a worst case scenario, and so if indifference only exists in this worst case scenario, the upper bound should be calculated by assuming selection of the novel outcome. Similarly, the upper bound used in the EO choice rule is based on a best case situation, and so it would appear most appropriate to select a known option in the case of indifference.

As is to be expected, the results given in Table 1 show that the lower bound of the expected utility corresponding to the EO choice rule is never greater than the lower bound for the EP choice rule, whilst the upper bound for the EP choice rule is never greater than the upper bound for the EO choice rule. Indeed, there can be no other choice rule that will generate a larger lower expected utility bound than that resulting from the EP choice rule, nor can there exist an alternative choice rule that will generate a larger upper expected utility bound than that resulting from the EO choice rule. However, when the considered known available fruit was f_1, f_2 or f_5, both EP and EO resulted in the same expected utility bounds. For f_1 and f_2 this occurs because the available known option has a utility value that is small enough to make experimentation the preferred strategy under both EP and EO, and so both choice rules give rise to the same optimal decision strategy. The EO choice rule is of course more readily disposed to experimentation in any case, but even the EP choice rule is supportive of experimenting when the alternatives are f_1 or f_2.

In the case of available alternative f_5, both the EP and EO strategies again coincide, but now by suggesting experimentation be avoided and that the DM should instead repeatedly select the available known option. Furthermore, as there is no uncertainty in the utility of option f_5, both the lower and upper expected utility bounds for the EP and EO choice rules also coincide. However, under the EO choice rule, the expected value of information gained following the initial experimentation of an unfamiliar outcome will increase as the length of the planning horizon is also increased, and this is because if the unfamiliar outcome were found to have greater utility, then it would be re-selected in every future epoch, see for example Raiffa and Schlaifer (1961) and Houlding and Coolen (2011). Hence, if the planning horizon of the strategy was extended beyond three decision epochs, then there would exist a horizon for which the EO choice rule would suggest initial experimentation. In contrast though, the EP choice rule will never suggest experimentation against the selection of the best known outcome regardless of the length of the planning horizon, and that is because even if the unfamiliar outcome is found to be better, the EP choice rule will assume that its utility will equal the infimum of the interval $[u_{(n)}, 1)$.

The EP and EO choice rules do disagree, however, when either fruit f_3 or f_4 is considered as the available known option. In the case of f_4 this is because, under a planning horizon of three periods, the EP choice rule suggests avoiding experimentation, whilst this duration is indeed considered sufficient for initial experimentation under the trial seeking EO choice rule. Finally, both EP and EO suggest initial experimentation against an initial selection of f_3, but differences arise in the expected utility bounds because the EP and EO choice rules suggest alternate strategies depending on the resulting observation from the initial experimentation, *i.e.*, the EO choice rule would suggest continued experimentation for some observations that would lead to the EP choice rule reverting to known outcomes.

6. Concluding Remarks

There has been limited discussion on the idea that preferences over decision outcomes may be uncertain, even though such scenarios have empirical support in life, *e.g.*, when a new brand becomes available for purchase. Here we introduced the possibility of using NPUI as a normative decision theory for such situations, and have discussed its virtues for when a DM faces a decision problem that includes the possibility of obtaining novel or unfamiliar outcomes. The inclusion of the NPI statistical technique appears natural for such situations, and is a simple inferential procedure that avoids making any unjustifiable distributional or conditional independence assumptions.

The decision selection criteria of EP and EO offer bounds on the expected utility resulting from a decision strategy, and were specifically designed for use within a sequential problem. Either could be employed to determine a specific strategy that is selected before the first decision epoch, and both have their merits in either maximizing the minimum expected utility stream, or maximizing the maximum possible expected utility stream. In selecting one over the alternative, a DM will indicate a level of trial aversion, with EP indicating extreme trial aversion and a preference for the known, whilst EO indicates extreme trial seeking behaviour and a preference for the unknown. The theory of trial aversion was introduced by Houlding and Coolen (2011) and is an analogous, but orthogonal, concept to the usual risk aversion of Pratt (1964) and Arrow (1971), but only arising in theories of decision making under uncertain preferences.

Finally, a further area that could be addressed is the possibility of there being two (or more) distinct sets of outcome types, with exchangeability of outcome utility only being assumed applicable within outcome type. Following Coolen (1996), given observation sets $u_{(1)}, u_{(2)}, \ldots, u_{(n)}$ and $v_{(1)}, v_{(2)}, \ldots, v_{(m)}$, the lower and upper probabilities of the event that $v_{m+1} \geq u_{n+1}$, are as follows:

$$\underline{P}(v_{m+1} \geq u_{n+1}) = \frac{\sum_{i,j} I_{\{v_{(i)} \geq u_{(j)}\}}}{(m+1)(n+1)} \quad (12)$$

$$\overline{P}(v_{m+1} \geq u_{n+1}) = \frac{n+m+1+\sum_{i,j} I_{\{v_{(i)} \geq u_{(j)}\}}}{(m+1)(n+1)} \quad (13)$$

In terms of utility inference, this extension of NPI is relevant when there is one group of outcomes that are considered to have exchangeable utility values (say fruits) and another distinct group of outcomes that also have their own exchangeable utility values (say vegetables), and the decision is a selection from amongst the union of the two. Unfortunately, it would appear to be first necessary to determine how differing outcome classes can be placed on a common utility scaling, for the suggestion used here of a 'hypothetical' best and worst outcome from within a single exchangeable class is no longer sufficient. The obvious extension would be to elicit relative preference between a hypothetical best outcome from one exchangeable class and the hypothetical best outcome from a second (or third *etc.*) exchangeable class, and this returns to the situation of commensurability discussed in Houlding and Coolen (2011), yet such an assumption would be difficult to implement in practice.

7. Acknowledgments

The authors wish to express their grateful thanks to the supportive comments and suggestions provided by Yakov Ben-Haim. B. Houlding is funded by the STATICA project, a Principle Investigator program of Science Foundation Ireland, Grant Number 08/IN.1/I1879.

References

Abellán, J., Baker, R.M., Coolen, & F.P.A., (2011). Maximising Entropy on the Nonparametric Predictive Inference Model for Multinomial Data. *European Journal of Operational Research, 212*, 112-122.

Arrow, K.J., (1971). *Essays in the Theory of Risk-Bearing*. Markham, Chicago.

Augustin, T., & Coolen F.P.A., (2004). Nonparametric Predictive Inference and Interval Probability. *Journal of Statistical Planning and Inference, 124*, 251-272.

Ben-Haim, Y., Dasco, C.C., Carrasco, J., & Rajan, N., (2009). Heterogeneous Uncertainties in Cholesterol Management. *International Journal of Approximate Reasoning, 50*, 1046-1065.

Berger, J.O., (1993). *Statistical Decision Theory and Bayesian Analysis.* Springer.

Bernoulli, D., (1738). Exposition of a New Theory on the Measurement of Risk (*Translation*). *Econometrica, 22*, 22-36.

Boutilier, C., (2003). On the Foundations of Expected Expected Utility. *18th International Joint Conference on AI*, 712-717.

Coolen, F.P.A., (1996). Comparing Two Populations Based on Low Stochastic Structure Assumptions. *Statistics & Probability Letters, 29*, 297-305.

Coolen, F.P.A., (2006). On Nonparametric Predictive Inference and Objective Bayesianism. *Journal of Logic, Language and Information, 15*, 21-47.

Coolen, F.P.A., Houlding, B., & Parkinson, G., (2011). Considerations on Jury Size and Composition using Lower Probabilities. *Journal of Statistical Planning and Inference, 141*, 382-391.

Cyert, R.M., & DeGroot, M.H., (1975). Adaptive Utility. In Day, R.H., Groves, T., (Eds.), *Adaptive Economic Models* (pp. 223-246). Academic Press, New York.

de Finetti, B., (1974). *Theory of Probability* (volume 1). John Wiley and Sons.

DeGroot, M.H., (1970). *Optimal Statistical Decisions.* McGraw-Hill.

Farrow, M., & Goldstein, M., (2006). Trade-Off Sensitive Experimental Design: A Multicriterion, Decision Theoretic, Bayes Linear Approach. *Journal of Statistical Planning and Inference, 136*, 498-526.

Hill, B.M., (1968). Posterior Distribution of Percentiles: Bayes' Theorem for Sampling from a Population. *Journal of the American Statistical Association, 63*, 677-691.

Hill, B.M., (1988). De Finetti's Theorem, Induction, and $A_{(n)}$ or Bayesian Nonparametric Predictive Inference. In Bernardo, J.M. *et al.* (Eds.), *Bayesian Statistics 3* (pp. 211-241). Oxford University Press.

Hill, B.M., (1993). Parametric Models for $A_{(n)}$: Splitting Processes and Mixtures. *Journal of the Royal Statistical Society B, 55*, 423-433.

Houlding, B., Coolen, & F.P.A., (2007). Sequential Adaptive Utility Decision Making for System Failure Correction. *Journal of Risk and Reliability, Proceedings of the IMech E Part O, 221*, 285-295.

Houlding, B., (2008). *Sequential Decision Making with Adaptive Utility.* PhD Thesis, Department of Mathematical Sciences, Durham University.

Houlding, B., & Coolen, F.P.A., (2011). Adaptive Utility and Trial Aversion. *Journal of Statistical Planning and Inference, 141*, 734-747.

Kmietowicz, Z.W., & Pearman, A.D., (1981). *Decision Theory and Incomplete Knowledge.* Aldershot Gower.

Pratt, J.W., (1964). Risk Aversion in the Small and in the Large. *Econometrica, 32*, 122-136.

Raiffa, H., & Schlaifer, R., (1961). *Applied Statistical Decision Theory.* Harvard University, Boston.

Savage, L.J., (1954). *The Foundations of Statistics.* John Wiley and Sons.

Wald, A., (1950). *Statistical Decision Functions.* John Wiley & Sons.