Subjective Logic

José C. Oliveira
Bernardo T. Amorim

April 27, 2020

Obs: This notes are basically a clone of José's with some extra teaks and notes of my own.

1 Introduction

Question 1.1. What is $p(y \parallel x)$ and $p(x \mathbin{\widetilde{\parallel}} y)$.

2 Elements of Subjective Opinions

2.1 Motivation for the Opinion Representation

For decision makers it can make a big difference whether probabilities are confident or uncertain. Decision makers should instead request additional evidence so the analysts can produce more confident conclusion probabilities about hypotheses of interest.

2.2 Flexibility of Representation

There can be multiple equivalent formal representations of subjective opinions.

2.3 Domains and Hyperdomains

Definition 2.1. *(Hyperdomain)* Let $\mathbb{X}$ be a domain, and let $\mathcal{P}(\mathbb{X})$ denote the powerset of $\mathbb{X}$. The powerset contains all subsets of $\mathbb{X}$, including the empty set $\{\varnothing\}$, and the domain $\mathbb{X}$ itself. The *hyperdomain* denoted $\mathcal{R}(\mathbb{X})$ is the reduced powerset of $\mathbb{X}$, i.e. the powerset excluding the empty-set $\{\varnothing\}$ and the domain value $\{\mathbb{X}\}$. The hyperdomain is expressed as

$$\text{Hyperdomain: } \mathcal{R}(\mathbb{X}) = \mathcal{P} \setminus \{\{\mathbb{X}\}, \{\varnothing\}\} \tag{2.1}$$

Question 2.1. I don't know if this is important, but I don't understand exactly how indexing works by the way that is explained in the book.

Definition 2.2. *(Composite set)* Let $\mathbb{X}$ be a domain of cardinality k, where $\mathcal{R}(\mathbb{X})$ is its hyperdomain of cardinality κ. Every proper subset $x \subset \mathbb{X}$ of cardinality $|x| \geq 2$ is a *composite value*. The set of composite values is the *composite set*, denoted $\mathcal{C}(\mathbb{X})$ and defined as:

$$\text{Composite set: } \mathcal{C}(\mathbb{X}) = \{x \subset \mathbb{X} \text{ where } |x| \geq 2\} \tag{2.2}$$

2.4 Random Variables and Hypervariables

Definition 2.3. *(Hypervariable)* Let $\mathbb{X}$ be a domain with coresponding hyperdomain $\mathcal{R}(\mathbb{X})$ A variable X takes its value from $\mathcal{R}(\mathbb{X})$ is a hypervariable.

Note 2.1. The events analyzed must be mutually exclusive.

2.5 Belief Mass Distribution and Uncertainty Mass

Definition 2.4. *(Belief Mass Distribution)* Let $\mathbb{X}$ be a domain with corresponding hyperdomain $\mathcal{R}(\mathbb{X})$, and let X be a variable over those domains. A belief mass distribution denote $\mathbf{b}_X$ assigns belief mass to possible values of the variable X. In the case of a random variable $X \in \mathbb{X}$, the belief mass distribution applies to domain $\mathbb{X}$, and in the case of a hypervariable $X \in \mathcal{R}(\mathbb{X})$ the belief mass distribution applies to hyperdomain $\mathcal{R}(\mathbb{X})$. This is formally defined as follows.

$$\begin{aligned}&\text{Multinomial belief mass distribution: } \mathbf{b}_X : \mathbb{X} \to [0,\ 1],\\ &\text{with the additivity requirement: } u_X + \textstyle\sum_{x\in\mathbb{X}} \mathbf{b}_X(x) = 1.\end{aligned} \tag{2.3}$$

$$\begin{aligned}&\text{Hypernominal belief mass distribution: } \mathbf{b}_X : \mathcal{R}(\mathbb{X}) \to [0,\ 1],\\ &\text{with the additivity requirement: } u_X + \textstyle\sum_{x\in\mathcal{R}(\mathbb{X})} \mathbf{b}_X(x) = 1.\end{aligned} \tag{2.4}$$

The sub-additivity of belief mass distributions is complemented by *uncertainty mass* denoted u_X.

2.6 Base Rate Distributions

Definition 2.5. *(Base Rate Distribution)* Let $\mathbb{X}$ be a domain, and let X be a random variable in $\mathbb{X}$. The base rate distribution $\mathbf{a}_X$ assigns base rate probability to possible values of $X \in \mathbb{X}$, and is an additive probability distribution, formally expressed as:

$$\begin{aligned}&\text{Base rate distribution: } \mathbf{a}_X : \mathbb{X} \to [0,\ 1],\\ &\text{with the additivity requirement: } \textstyle\sum_{x\in\mathbb{X}} \mathbf{a}_X(x) = 1.\end{aligned} \tag{2.5}$$

Definition 2.6. *(Base Rate Distribution over Values in a Hyperdomain)* Let $\mathbb{X}$ be a domain with corresponding hyperdomain $\mathcal{R}(\mathbb{X})$, and let X be a variable over those domains. Assume the base rate distribution $\mathbf{a}_X$ over the domain $\mathbb{X}$ according to Definition 2.5. The base rate $\mathbf{a}_X$ for a composite value $x \in \mathcal{R}(\mathbb{X})$ can be computed as follows:

$$\text{Base rate over composite values: } \mathbf{a}_X(x_i) = \sum_{\substack{x_j \in \mathbb{X} \\ x_j \subseteq x_i}} \mathbf{a}_X(x_j), \ \forall x_i \in \mathcal{R}(\mathbb{X}). \tag{2.6}$$

Definition 2.7. *(Relative Base Rate)* Assume a domain $\mathbb{X}$ of cardinality k, and the corresponding hyperdomain $\mathcal{R}(\mathbb{X})$. Let X be a hypervariable over $\mathcal{R}(\mathbb{X})$. Assume that a base rate distribution $\mathbf{a}_X$ is defined over $\mathbb{X}$ according to Definition 2.6. Then the base rate of a value x relative to a value v_i is expressed as the relative base rate $\mathbf{a}_X(x|x_i)$ defined below.

$$\mathbf{a}_X(x|x_i) = \frac{\mathbf{a}_X(x \cap x_i)}{\mathbf{a}_X(x_i)}, \ \forall x, x_i \in \mathcal{R}(\mathbb{X}), \text{ where } \mathbf{a}_X(x_i) \neq 0. \tag{2.7}$$

In the case when $\mathbf{a}_X(x_i) = 0$, then $\mathbf{a}_X(x|x_i) = 0$. Alternatively it can simply be assumed that $a_X(x_i) > 0$, for every $x_i \in \mathbb{X}$, meaning that everything we include in the domain has a non-zero base rate of occurrence in general.

2.7 Probability Distributions

Definition 2.8. *(Probability Distribution)* Let $\mathbb{X}$ be a domain with corresponding hyperdomain $\mathcal{R}(\mathbb{X})$, and let X denote a variable in $\mathbb{X}$ or in $\mathcal{R}(\mathbb{X})$. The standard probability distribution $\mathbf{p}_X$ assigns probabilities to possible values of $X \in \mathbb{X}$. The hyper-probability distribution $\mathbf{p}^{\mathrm{H}}_X$ assigns probabilities to possible values of $X \in \mathcal{R}(\mathbb{X})$. These distributions are formally defined below:

$$\begin{aligned} &\text{Probability distribution: } \mathbf{p}_X : \mathbb{X} \to [0,\ 1], \\ &\text{with the additivity requirement: } \textstyle\sum_{x \in \mathbb{X}} \mathbf{p}_X(x) = 1. \end{aligned} \tag{2.8}$$

$$\begin{aligned} &\text{Hyper-probability distribution: } \mathbf{p}^{\mathrm{H}}_X : \mathcal{R}(\mathbb{X}) \to [0,\ 1], \\ &\text{with the additivity requirement: } \textstyle\sum_{x \in \mathcal{R}(\mathbb{X})} \mathbf{p}^{\mathrm{H}}_X(x) = 1. \end{aligned} \tag{2.9}$$

3 Opinion Representations

3.1 Opinion Classes

The opinion itself is a composite function $\omega^A_X = (\mathbf{b}_X, u_X, \mathbf{a}_X)$, consisting of the belief mass distribution $\mathbf{b}_X$, the uncertainty mass u_X, and the base rate distribution $\mathbf{a}_X$.

Classes:

- *Binomial*: Domain $\mathbb{X}$ and variable X are binary.

- *Multinomial*: Domain larger than binary and the variable is a random variable $X \in \mathbb{X}$.
- *Hypernomial*: Domain larger than binary and the variable is a hypervariable $X \in \mathcal{R}(\mathbb{X})$.

Levels of confidence of a opinion:

- *Vacuous*: $u_X = 1$.
- *Uncertain*: $0 < u_X < 1$.
- *Dogmatic*: $u_X = 0$.
- *Absolute*: One single value is TRUE by assigning belief mass 1 to that value.

3.2 Aleatory and Epistemic Opinions

- *Aleatory Uncertainty*, which is the same as statistical uncertainty, express that we do not know the outcome each time we run the same experiment, we only know the long-term relative frequency of outcomes. E.g.: Flip a coin.
- *Epistemic Uncertainty*, aka systematic uncertainty, express that we could in principle know the outcome of a specific or future or past event, but that we do not have enough evidence to know it exactly. E.g.: Assassination of President Kennedy.

Question 3.1. First-order and second-order opinions?

Question 3.2. Projected probability?

High aleatory/epistemic uncertainty is consistent with both high and low uncertainty mass.

- **An aleatory Opinion** applies to a variable governed by a frequentist process, and that represents the (uncertain) likelihood of values of the variable in any unknown past or future instance of the process. An aleatory opinion can naturaly have an arbitrary uncertainty mass.
- **An epistemic Opinion** applies to a variable that is assumed to be non-frequentist, and that represents the (uncertain) likelihood of values of the variable in a specific unknown past or future instance.

3.3 Binomial Opinions

3.3.1 Binomial Opinion Representation

Definition 3.1. *Binomial Opinion* Let $\mathbb{X} = \{x, \overline{x}\}$ be a binary domain with binomial random variable $X \in \mathbb{X}$. A binomial opinion about the truth/presence of value x is the ordered quadruplet $\omega_x = (b_x, d_x, u_x, a_x)$, where the additivity requirement

$$b_x + d_x + u_x = 1 \tag{3.1}$$

is satisfied, and where the respective parameters are defined as

- b_x: *belief mass* in support of x being TRUE (i.e. $X = x$),
- d_x: *disbelief mass* in support of x being FALSE (i.e. $X = \overline{x}$)
- u_x: *uncertainty mass* representing the vacuity of evidence,
- a_x: *base rate*, i.e. prior probability of x without any evidence.

The projected probability of a binomial opinion about value x is defined by the following equation.

$$\mathrm{P}(x) = b_x + a_x u_x. \tag{3.2}$$

The variance of binomial options is expressed as

$$\mathrm{Var}(x) = \frac{\mathrm{P}(x)(1 - \mathrm{P}(x))u_x}{W + u_x}, \tag{3.3}$$

where W denotes non-informative prior weight, which must be set to $W = 2$ as explained in Section 3.5.2. Binomial opinion variance is derived from the variance of the Beta PDF.

3.3.2 The Beta Binomial Model

Definition 3.2. *(Beta Probability Density Function)* Assume a binaru domain $\mathbb{X} = \{x, \overline{x}\}$ and a random variable $X \in \mathbb{X}$. Let p denote the continuous probability function $p \; : \; X \to [0, 1]$ where $p(x) + p(\overline{x}) = 1$. For compactness of notation e define $p_x \equiv p\,(x)$ and $p_{\overline{x}} \equiv p\,(\overline{x})$.

The parameter α represents evidence/observations of $X = x$, and the parameter β represents evidence/observations of $X = \overline{x}$. With p_x as variable, the Beta probability density function $\mathrm{Beta}(p_x, \alpha, \beta)$ is the function expressed as

$$\mathrm{Beta}(p_x, \alpha, \beta) \; : \; [0, 1] \to \mathbb{R}_{\leq 0}, \text{ where} \tag{3.4}$$

$$\mathrm{Beta}(p_x, \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} (p_x)^{\alpha - 1} (1 - p_x)^{\beta - 1}, \; \alpha > 0, \; \beta > 0, \tag{3.5}$$

with the restrictions that $p_(x) \neq 0$ if $\alpha < 1$, and $p(x) \neq 1$ if $\beta \leq 1$.

Note 3.1. The part of the equation that evolves the gamma PDF exists only to ensure that the integral of the PDF equals 1.

Assume that x represents a frequentist event. Let r_x (or r_s) denote the number of observations of x (or $\overline{x}$). With the evidence observations, the base rate a_x and the non-informative prior weight W, the α and β parameters can be expressed as:

$$\begin{cases} \alpha = r_x + a_x W, \\ \beta = s_x + (1 - a_x). \end{cases} \tag{3.6}$$

The evidence notation of the Beta PDF is denoted by $\mathrm{Beta}^{\mathrm{e}}(p_x, r_x, s_x, a_x)$.

Note 3.2. $W = 2$ because, with $r_x = 0$ and $s_x = 0$, and $a_x = \frac{1}{2}$, the Beta PDF $(p_x^{\alpha-1}(1-p_x)^{\beta-1})$ becomes a constant, which is equivalent to the uniform PDF [0,1]. This makes sense intuitively, if we don't have any evidence and our base rates are the same for both events, any event is perceived as equally likely.

Expected probability:

$$\mathrm{E}(x) = \frac{r_x + a_x W}{r_x + s_x + W} \tag{3.7}$$

Variance:

$$\mathrm{Var}(x) = \frac{\mathrm{P}(x)(1 - \mathrm{P}(x))u_x}{W + u_x} \tag{3.8}$$

3.3.3 Mapping Between a Binomial Opinion and a Beta PDF

Definition 3.3. *(Mapping: Binomial Opinion $\leftrightarrow$) Beta PDF)* Let $\omega_x = (b_x, d_x, u_x, a_x)$ be a binomial opinion, and let $p(x)$ be a probability distribution, both over the same binomial random variable X. Let $\mathrm{Beta}^{\mathrm{e}}(p_x, r_x, s_x, a_x)$ a Beta PDF over the probability variable p_x defined as a function of r_x, s_x and a_x according. The opinion ω_x and the Beta PDF $\mathrm{Beta}^{\mathrm{e}}(p_x, r_x, s_x, a_x)$ are equivalent through the following mapping:

$$\begin{cases} b_x = \dfrac{r_x}{W + r_x + s_x}, \\ d_x = \dfrac{s_x}{W + r_x + s_x}, \\ u_x = \dfrac{W}{W + r_x + s_x} \end{cases} \Leftrightarrow \begin{cases} \begin{cases} r_x = \dfrac{b_x W}{u_x}, \\ s_x = \dfrac{d_x W}{u_x}, \\ 1 = b_x + d_x + u_x \end{cases} & \text{if } u \neq 0 \\ \begin{cases} r_x = b_x \cdot \infty, \\ s_x = d_x \cdot \infty, \\ 1 = b_x + d_x. \end{cases} & \text{if } u \neq 0 \end{cases} \tag{3.9}$$

The equivalence between binomial opinions and Beta PDFs is very powerful, because subjective-logic operators (SL operators) can then be applied to Beta PDFs, and statistics operations for Beta PDFs can be applied to opinions. In addiction, it makes it possible to determine binomial opinions from statistical observations.

3.4 Multinomial Opinions

3.4.1 The Multinomial Opinion Representation

Definition 3.4. *(Multinomial Opinion)* Let $\mathbb{X}$ be a domain larger than binary, i.e. so that $k = |X| > 2$. Let X be a random variable in $\mathbb{X}$. A multinomial opinion over the random variable X is the ordered triplet $\omega_X = (\mathbf{b}_X, u_X, \mathbf{a}_X)$ where

- $\mathbf{b}_X$ is a belief mass distribution over X,
- u_X is the uncertainty mass which represents the vacuity of evidence,
- $\mathbf{a}_X$ is a base rate distribution over $\mathbb{X}$,

and the multinomial additivity requirement of Eq.(2.3) is satisfied.

A multinomial opinion contains $(2k+1)$ parameters. However, given the belief and uncertainty mass additivity of Eq.(2.3), and the base rate additivity of Eq.(2.5), multinomial opinions only have $(2k-1)$ degrees of freedom.

The projected probability distribution of multinomial opinions is defined by:

$$\mathbf{P}_X(x) = \mathbf{b}_X(x) + \mathbf{a}_X(x)u_X, \ \forall x \in \mathbb{X}. \tag{3.10}$$

The variance of multinomial opinions is expressed as

$$\mathrm{Var}_X = \frac{\mathbf{P}_X(x)(1 - \mathbf{P}_X(x)u_X)}{W + u_X}, \tag{3.11}$$

where W denotes non-informative prior weight, which must be set to $W = 2$.

3.4.2 The Dirichlet Multinomial Model

Definition 3.5. *(Dirichlet Probability Density Function)* Let $\mathbb{X}$ be a domain consisting of k mutually disjoint values. Let α_X represent the strength vector over the values of $\mathbb{X}$, and let $\mathbf{p}_X$ denote the probability distribution over $\mathbb{X}$. With $\mathbf{p}_X$ as a k-dimensional variable, the Dirichlet PDF denoted $\mathrm{Dir}(\mathbf{p}_X, \alpha_X)$ is expressed as:

$$\mathrm{Dir}(\mathbf{p}_X, \alpha_X) = \frac{\Gamma\left(\sum\limits_{x \in \mathbb{X}} \alpha_X(x)\right)}{\prod\limits_{x \in \mathbb{X}} \Gamma(\alpha_X(x))} \prod_{x \in \mathbb{X}} \mathbf{p}_X(x)^{(\alpha_X(x)-1)}, \text{ where } \alpha_X(x) \geq 0, \tag{3.12}$$

with the restrictions that $\mathbf{p}_X(x) \neq 0$ if $\alpha_X(x) < 1$.

The evidence representation of the Dirichlet PDF is denoted by $\mathrm{Dir}^{\mathrm{e}}_X(\mathbf{p}_X, \mathbf{r}_X, \mathbf{a}_X)$, where the total strength $\alpha_X(x)$ for each value $x \in \curvearrowleft$ can be expressed as

$$\alpha_X(x) = \mathbf{r}_X(x) + \mathbf{a}_X(x)W, \text{ where } \mathbf{r}_X(x) \geq 0 \;\forall x \in \mathbb{X}. \tag{3.13}$$

The evidence-Dirichlet PDF is expressed in terms of the evidence vector $\mathbf{r}_X$, where $\mathbf{r}_X(x)$ is the evidence for outcome $x \in \mathbb{X}$. In addition, the base rate distribution $\mathbf{a}_X$ and the non-informative prior weight W are parameters in the expression for the evidence-Dirichlet PDF.

The expected distribution over $\mathbb{X}$ can be written as

$$\mathbf{E}_X(x) = \frac{\mathbf{r}_X(x) + \mathbf{a}_X(x)W}{W + \sum\limits_{x_j \in \mathbb{X}} \mathbf{r}_X(x_j)} \;\forall x \in \mathbb{X}. \tag{3.14}$$

The variance of the Dirichlet is defined by

$$\mathrm{Var}_X(x) = \frac{\mathbf{P}_X(x)(1 - \mathbf{P}_X(x))}{W + u_X}. \tag{3.15}$$

3.4.3 Visualising Dirichlet Probability Density Functions

Dirichlet PDFs over ternary domains are the largest that can be practically visualized.

The Figure 3.4 from the book shows graphical representations with non-informative prior Dirichlet PDF, and posterior Dirichlet PDF. At the second case, $\mathbf{r}_X$ is not constant on 0.

3.4.4 Coarsening Example: From Ternary to Binary

Question 3.3. What does coarsening mean?

Question 3.4. Can you always partition a multinomial domain in a binary one and transform a multinomial opinion in a binary one?

This subsubsection shows a ternary domain $\{x_1, x_2, x_3\}$ and it reduces to binary domain making $\overline{x}_1 = \{x_2, x_3\}$. This way makes it possible to visualize prior and posteriori Beta PDFs for $p(x_1)$. Here I realize that didn't understand at section 3.3.2 that a Beta (and Dirithlet here) is a function from probability to probability. The PDF I saw most so far was from Normal Distribution and was a function from a random variable to probability. The take-way here (actually from subsubsection 3.3.2) is that the expected probability (3.7 and 3.14) is the posteriori probability, i.e. with the new evidences.

3.4.5 Mapping Between Multinomial Opinion and Dirichlet PDF

Definition 3.6. *(Mapping: Multinomial Opinion $\leftrightarrow$ Dirichlet PDF)* Let $\omega_X = (\mathbf{b}_X, u_X, \mathbf{a}_X)$ be a multinomial opinion and let $\mathrm{Dir}^{\mathrm{e}}_X(\mathbf{p}_X, \mathbf{r}_X, \mathbf{a}_X)$ be a Dirichlet PDF, both over the same variable $X \in \mathbb{X}$. These are equivalent through the following mapping,

$$\forall x \in \mathbb{X}$$
$$\begin{cases} \mathbf{b}_X(x) & = \dfrac{\mathbf{r}_X(x)}{W + \sum\limits_{x_i \in \mathbb{X}} \mathbf{r}_X(x_i)} \\ u_X & = \dfrac{W}{W + \sum\limits_{x_i \in \mathbb{X} \mathbf{r}_X(x_i)}} \end{cases} \Leftrightarrow \begin{cases} \begin{cases} \mathbf{r}_X(x) = \dfrac{W \mathbf{b}_X(x)}{u_X} \\ 1 = u_X = \sum\limits_{x_i \in \mathbb{X}} \mathbf{b}_X(x_i) \end{cases} & \text{if } u_X \neq 0 \\ \begin{cases} \mathbf{r}_X(x) = \mathbf{b}_X(x) \cdot \infty \\ 1 = \sum\limits_{x_i \in \mathbb{X}} \mathbf{b}_X(x_i) \end{cases} & \text{if } u_X \neq 0 \end{cases} \tag{3.16}$$

Statistics tools and methods, such as collecting statistical observation evidence, can no be applied to opinions.

3.4.6 Uncertainty-Maximisation of Multinomial Opinions

Question 3.5. Didn't understood the method for maximizing uncertainty.

Given a multinomial opinion ω_X, with its projected probability distribution $\mathbf{P}_X$, the corresponding uncertainty-maximised opinion is denoted $\ddot{\omega}_X = (\ddot{\mathbf{b}}_X, \ddot{u}_X, \mathbf{a}_X)$. The theoretical maximum uncertainty mass $\ddot{u}_X$ is determined by converting as much belief as possible into uncertainty mass, while preserving consistent projected probabilities. On the simplex, the opinion $\ddot{\omega}_X$ will be on one of the planes, closer to u_X vertex.

The components of the uncertainty-maximised opinion $\ddot{\omega}_X$ should satisfy the following requirements:

$$\ddot{u} = \frac{\mathbf{P}_X(x_{i_0})}{\mathbf{a}_X(x_{i_0})}, \text{ for some } i_0 \in \{1, \dots, k\}, \text{ and} \tag{3.17}$$

$$\mathbf{P}_X(x_i) \geq \mathbf{a}_X(x_i) u_X, \text{ for every } i \in \{1, \dots, k\}. \tag{3.18}$$

The requirement of Eq.(3.18) ensures that all the belief masses determined according to Eq.(3.10) are non-negative. These requirements lead to the theoretical uncertainty maximum:

$$\ddot{u}_X = \min_i \left[\frac{\mathbf{P}_X(x_i)}{\mathbf{a}_X(x_i)} \right] \tag{3.19}$$

Non uncertainty-maximised only can be aleatory opinion. Uncertainty-maximised can be and an aleatory opinion or an epistemic opinion.

3.5 Hyper-opinions

3.5.1 The Hyper-opinion Representation

Definition 3.7. Let $\mathbb{X}$ be a domain of cardinality $k > 2$, with corresponding hyperdomain $\mathscr{R}(\mathbb{X})$. Let X be a hypervariable in $\mathscr{R}(\mathbb{X})$. A hyper-opinion on the hypervariable X is the ordered triplet $\omega_X = (\mathbf{b}_X, u_X, \mathbf{a}_X)$ where:

- $\mathbf{b}_X$ is a *belief mass distribution* over $\mathscr{R}(\mathbb{X})$,
- u_X is the *uncertainty mass* which represents the vacuity of evidence,
- $\mathbf{a}_X$ is a *base rate distribution* over $\mathbb{X}$,

and the hypernomial additivity of 2.4 is satisfied.

The representation of subjective opinion is the same as in previous domains.

The belief mass distribution $\mathbf{b}_X$ over $\mathscr{R}(\mathbb{X})$ has $2^k - 2$ parameters, whereas the base rate distribution $\mathbf{a}_X$ over $\mathbb{X}$ only has k parameters and the uncertainty u_X is a simple scalar. A general opinion thus contain $2^k + k - 1$ parameters. However it has $2^k - k - 3$ degrees of freedom.

By using the concept of relative base rate 2.7, the projected probability distribution $\mathbf{P}_X$ of hyper-opinions can be expressed as:

$$\mathbf{P}_X(x) = \sum_{x_i \in \mathscr{R}(\mathbb{X})} \mathbf{a}_X(x|x_i)\mathbf{b}_X(x_i) + \mathbf{a}_X(x)u_X, \qquad \forall x \in \mathbb{X} \tag{3.20}$$

The sum of projected probabilities over hyper-opinions is, in general, super-additive. This comes from the fact that composite values are partially based on overlapping singleton values, thus some elements are counted multiple times.

3.5.2 Projecting Hyper-opinions to Multinomial Opinions

Given a hyper-opinion, it can be useful to project it into a multinomial opinion. The procedure goes as follows.

If $\mathbf{b}'_X(x)$ is a belief mass distribution defined by the sum in Eq.(3.20), i.e.

$$\mathbf{b}'_X(x) = \sum_{x' \in \mathscr{R}(\mathbb{X})} \mathbf{a}_X(x|x')\mathbf{b}_X(x'), \tag{3.21}$$

then $\mathbf{b}'_X$ together with u_X satisfies the additivity property, i.e. $\omega'_X = (\mathbf{b}'_X, u_X, \mathbf{a}_X$ is a multinomial opinion. From the equations above we obtain $\mathbf{P}(\omega_X) = \mathbf{P}(\omega'_X)$. This means that every hyper-opinion can be approximated with a multinomial opinion which has the same projected probability distribution as the initial hyper-opinion.

Question 3.6. What guarantees me that this is a good approximation and what am I loosing by doing this projection, since the hyper-opinion representation also dismisses, in some sense, the properties of a hyper-domain.

3.5.3 The Dirichlet Model Applied to Hyperdomains

The traditional Dirichlet model applies naturally to a multinomial domain $\mathbb{X}$ of cardinality k, and there is a simple bijective mapping between multinomial opinions and Dirichlet PDFs. The question is if it can also be done for hyperdomains. A problem with applying the Dirichlet model to hyperdomains is that those aren't mutually exclusive, which is contrary to the assumption in the tradicional Dirichlet model.

The approach that we follow is to artificially assume that hyperdomain $\mathscr{R}(\mathbb{X})$ is exclusive $(x_i \cap x_j = \emptyset)$.

The set of input arguments to the Dirichlet PDF over $\mathscr{R}(\mathbb{X})$ then becomes a sequence of strength parameters of the k possible values $x \in \mathscr{R}(\mathbb{X})$ represented as k positive real numbers $\alpha_X(x_i), i = 1, ..., k$, each corresponding to one of the possible values $x \in \mathscr{R}(\mathbb{X})$. Because this is a Dirichlet PDF over a hypervariable, it is called a Dirichlet Hyper-PDF, or Dirichlet HPDF for short.

$$Dir_X^H(\mathbf{p}_X^H, \alpha_X) = \frac{\Gamma\left(\sum\limits_{x \in \mathscr{R}(\mathbb{X})} \alpha_X(x)\right)}{\prod\limits_{x \in \mathscr{R}(\mathbb{X})} \Gamma(\alpha_X(x))} \prod_{x \in \mathscr{R}(\mathbb{X})} \mathbf{p}_X^H(x)^{\alpha_X(x)-1} \tag{3.22}$$

where $\alpha_X(x) \geq 0$ and $p_X^H(x) \neq 0$ if $\alpha_X(x) < 1$.

The strength vector α_X represents the prior as well as the observation evidence, now assumed applicable to values $x \in \mathscr{R}(\mathbb{X})$. Since the values of $\mathscr{R}(\mathbb{X})$ can contain multiple singletons from $\mathbb{X}$, a value of $\mathscr{R}(\mathbb{X})$ has a base rate equal to the sum of the base rates of the singletons it contains. Thus the strengh $\alpha_X(x)$ can be expressed as:

$$\alpha_X(x) = \mathbf{r}_X(x) + \mathbf{a}_X(x)W \tag{3.23}$$

where:

- $\mathbf{r}_X(x) \geq 0$,
- $\mathbf{a}_X(x) = \sum\limits_{x_j \subseteq x_{x_j \in X}} \mathbf{a}(x_j)$,
- $W = 2$, $\forall x \in \mathscr{R}(\mathbb{X})$.

Question 3.7. Apparently, by the formula of the strength of a value in a hyper-domain, the observation of some evidence only affects one composite values that contains it. Which I believe derives from the assumption that the values are mutually exclusive. In this sense, the hyper-opinion is only "hyper", when we analyze the base rate. Thus how useful is the use of the Hyper-Dirichlet if it apparently hides from us one of the most important parts of a hyper-domain, the dependence of its composite values?

The Dirichlet HPDF over of set of k possible states $x_i \in \mathscr{R}(\mathbb{X})$ can thus be expressed as a function of the observation evidence $\mathbf{r}_X$ and the base rate distribution $\mathbf{a}_X x$, where $x \in \mathscr{R}(\mathbb{X})$. The subscript 'eH' in the notation Dir_X^{eH} indicates that it is expressed as a function of the evidence parameter vector $\mathbf{r}_X$ (not of the strength parameter vector α_X), and that it is a Dirichlet HPDF.

$$\mathrm{Dir}_X^{\mathrm{eH}}(\mathbf{p}_X^H, \mathbf{r}_X, \mathbf{a}_X) = \frac{\Gamma\left(\sum\limits_{x \in \mathscr{R}(\mathbb{X})} (\mathbf{r}_X(x) + \mathbf{a}_X(x)W)\right)}{\prod\limits_{x \in \mathscr{R}(\mathbb{X})} \Gamma(\mathbf{r}_X(x) + \alpha_X(x)W)} \prod_{x \in \mathscr{R}(\mathbb{X})} \mathbf{p}_X^H(x)^{\mathbf{r}_X + \mathbf{a}_X W - 1} \tag{3.24}$$

where $\mathbf{r}_X(x) + \mathbf{a}_X(x)W \geq 0$, with the restriction that $\mathbf{p}_X^H(x) \neq 0$ if $(\mathbf{r}_X(x) + \mathbf{a}_X(x)W < 1)$

The expected probability of any of the k values $x \in \mathbb{X}$ can be written as.

$$\mathbf{E}_X(x) = \frac{\sum\limits_{x_i \in \mathscr{R}(\mathbb{X})} \mathbf{a}_X(x|x_i)\mathbf{r}(x_i) + W\mathbf{a}_X(x)}{W + \sum\limits_{x_i \in \mathscr{R}(\mathbb{X})} \mathbf{r}(x_i)} \tag{3.25}$$

$\forall x \in \mathbb{X}$.

3.5.4 Mapping Between a Hyper-opinion and a Dirichlet HPDF

The hyper opinion ω_X and the Dirichlet HPDF $\mathrm{Dir}_X^{\mathrm{eH}}(\mathbf{p}_X^H, \mathbf{r}_X, \mathbf{a}_X)$ are equivalent through the following mapping:

$\forall x \in \mathscr{R}(\mathbb{X})$

- $\mathbf{b}_X(x) = \frac{\mathbf{r}_X(x)}{W + \sum_{x_i \in \mathscr{R}(\mathbb{X})} \mathbf{r}_X(x_i)}$
- $u_x = \frac{W}{W + \sum_{x_i \in \mathscr{R}(\mathbb{X})} \mathbf{r}_X(x_i)}$

$\Updownarrow$

For $u_X \neq 0$:

- $\mathbf{r}_X(x) = \frac{W\mathbf{b}_X(x)}{u_X}$
- $1 = u_X + \sum\limits_{x_i \in \mathscr{R}(\mathbb{X})} \mathbf{b}_X(x_i)$

For $u_X = 0$

- $\mathbf{r}_X(x) = \mathbf{b}_X(x).\infty$
- $1 = \sum\limits_{x_i \in \mathscr{R}(\mathbb{X})} \mathbf{b}_X(x_i)$

The Dirichlet HPDF described above applies to the hyperdomain $\mathscr{R}(\mathbb{X})$ and is not suitable for representing probability over the corresponding domain $\mathbb{X}$. What we need is a PDF that somehow represents the parameters of the Dirichlet HPDF over the domain $\mathbb{X}$.

A PDF that does exactly this can be obtained by integrating the evidence parameters for the Dirichlet HPDF to produce the parameters for a PDF over the probability variable $\mathbf{p}_X$. In other words, the evidence on singleton values of the random variable must be computed as a function of the evidence of composite values of the hypervariable. This method will be described by an equation below. One takeaway of this method is that it needs a normalizing factor $B(\mathbf{r}_X, \mathbf{a}_X)$ that must be computed numerically.

The hyper-Dirichlet PDF is denoted $\mathrm{HDir}^{\mathrm{e}}_X(\mathbf{p}^H_X, \mathbf{r}_X, \mathbf{a}_X)$. Its mathematical expression is given by the equation below:

$$\mathrm{HDir}^{\mathrm{e}}_X(\mathbf{p}^H_X, \mathbf{r}_X, \mathbf{a}_X) = B(\mathbf{r}_X, \mathbf{a}_X)^{-1} \left(\prod_{i=1}^{k} \mathbf{p}_X(x_i)^{\mathbf{r}_X(x_i) + \mathbf{a}_X(x_i)W - 1} \prod_{j=k+1}^{K} \mathbf{p}_X(x_j)^{\mathbf{r}_X(x_j)} \right) \tag{3.26}$$

Where $B(\mathbf{r}_X, \mathbf{a}_X)$ equals:

$$B(\mathrm{r}_X, \mathbf{a}_X) = \int\limits_{\substack{\mathbf{p}_X(x) \geq 0 \\ \sum_{j=(k+1)}^{\kappa} \mathbf{p}_X(x_j) \geq 1}} \left(\prod_{i=1}^{k} \mathbf{p}_X(x_i)^{(\mathbf{r}_X(xi) + \mathbf{a}_X(x_i)W - 1)} \prod_{j=(k+1)}^{\kappa} \mathbf{p}_X(x_j)^{\mathbf{r}_X(x_j)} \right) \mathrm{d}(\mathbf{p}_X(x_1), \ldots, \mathbf{p}_X(k_\kappa)). \tag{3.27}$$

Note 3.3. The equation above strictly distinguishes between composite values and singleton values (the notation is slightly different).

Question 3.8. I got pretty confused with this part. The formulas aren't making sense to me. Also the simplification he did in the HPDF seems to take the whole purpose of the hyper-opinion.

3.6 Alternative Opinion Representations

3.6.1 Probabilistic Notation of Opinions

Definition 3.8. *(Probabilistic Opinion Notation)* Assume domain $\mathbb{X}$ with random variable X, and let $\omega_X = (\mathbf{b}_X, u_X, \mathbf{a}_X)$ be a binomial or multinomial opinion on X. Let $\mathbf{P}_X$ be the corresponding projected probability distribution over X defined according to Eq.(3.10). The probabilistic notation for multinomial opinions is given below.

$$\text{Probabilistic opinion: } \pi_X = (\mathbf{P}_X, u_X, \mathbf{a}_X)$$
$$\text{Constraints: } \begin{cases} \mathbf{a}_X(x)u_X \leq \mathbf{P}_X(x) \leq (\mathbf{a}_X(x)u_X + 1 - u_X), \ \forall x \in \mathbb{X}. \\ \sum_{x \in \mathbb{X}} \mathbf{P}_X(x) = 1, \ \forall x \in \mathbb{X}. \end{cases} \tag{3.28}$$

Question 3.9. Why $\mathbf{P}_X(x) \leq (\mathbf{a}_X(x)u_X + 1 - u_X)$?

Definition 3.9. Let $\omega_X = (\mathbf{b}_X, u_X, \mathbf{a}_X)$ be a multinomial belief opinion, and let $\pi_X = (\mathbf{P}_X, u_X, \mathbf{a}_X)$ be a multinomial probabilistic opinion, both over the same variable $X \in \mathbb{X}$. The multinomial opinions ω_X and π_X are equivalent because the belief mass distribution $\mathbf{b}_X$ is uniquely determined through Eq.(3.10) which is rearranged in Eq.(3.29):

$$\mathbf{b}_X(x) = \mathbf{P}_X(x) - \mathbf{a}_X(x)u_X. \tag{3.29}$$

3.6.2 Qualitative Opinion Representation

Table 3.4 from the book shows qualitative levels os likelihood and confidence. Figure 3.8 shows 2 possibles mappings between binomial opinion and qualitative representation of opinion. These mapping differ from each other by the base rate. Table 3.4 shows also the difference between likelihood and confidence.

Let ω_X be a binomial or multinomial opinions with uncertainty mass u_X, then we define:

$$\text{Confidence}(\omega_X) = c_X = 1 - u_X. \tag{3.30}$$

For hyper-opinions, low uncertainty mass does not necessarily indicate high confidence because belief mass can express vagueness.

Naturally, some mappings will always be impossible for a given base rate (see Figure 3.8), but these are logically inconsistent and should be excluded from selection.

4 Decision Making Under Vagueness and Uncertainty

4.1 Aspects of Belief and Uncertainty in Opinions

4.1.1 Sharp Belief Mass

Definition 4.1. *(Sharp Belief Mass)* Let $\mathbb{X}$ be a domain with hyperdomain $\mathcal{R}(\mathbb{X})$ and variable X. Given an opinion ω_X, the sharp belief mass of value $x \in \mathcal{R}(\mathbb{X})$ is the function $\mathbf{b}^{\mathrm{S}}_X : \mathcal{R}(\mathbb{X}) \rightarrow [0, 1]$ expressed as

$$\text{Sharp belief mass: } \mathbf{b}^{\mathrm{S}}_X = \sum_{x_i \subseteq x} \mathbf{b}_X(x_i), \ \forall x \in \mathcal{R}(\mathbb{X}). \tag{4.1}$$

Definition 4.2. *(Total Sharp Belief Mass)* Let $\mathbb{X}$ be a domain with variable X, and let ω_X be an opinions on $\mathbb{X}$. The total sharp belief mass contained in the opinion ω_X is the function $\mathbf{b}_X^{\mathrm{TS}} : \mathbb{X} \to [0,1]$ expressed as

$$\text{Total Sharp belief mass: } b_X^{\mathrm{TS}} = \sum_{x_i \subseteq \mathbb{X}} \mathbf{b}_X(x_i). \tag{4.2}$$

4.1.2 Vague Belief Mass

The vague belief mass on a value $\mathbf{x} \in \mathcal{R}(\mathbb{X})$ is defined asthe weighted sum of belief masses on the composite values of which x is a member,where the weights are determined by the base rate distribution.

Definition 4.3. *(Vague Belief Mass)* Let $\mathbb{X}$ be a domain with hyperdomain $\mathcal{R}(\mathbb{X})$ and composite set $\mathcal{C}(\mathbb{X})$. Given anopinion ω_X, the vague belief mass on $x \in \mathcal{R}(\mathbb{X})$ is the function $\mathbf{b}_X^{\mathrm{V}} : \mathcal{R}(\mathbb{X}) \to [0,1]$:

$$\text{Vague belief mass: } \mathbf{b}_X^{\mathrm{V}}(x) = \sum_{\substack{x_i \in \mathcal{C}(\mathbb{X}) \\ x_i \not\subseteq x}} \mathbf{a}_X(x|x_i)\mathbf{b}_X(x_i),\ \forall x \in \mathcal{R}(\mathbb{X}). \tag{4.3}$$

Definition 4.4. *(Total Vague Belief Mass)* Let $\mathbb{X}$ be a domain with variable X, andlet ω_X be an opinions on $\mathbb{X}$. The total vagueness contained in the opinion ω_X is thefunction ${}_X^{\mathrm{TV}} : \mathcal{C}(\mathbb{X}) \to [0,1]$ expressed as:

$$\text{Total vague belief mass: } b_X^{\mathrm{TV}} = \sum_{x \in \mathcal{C}(\mathbb{X})} \mathrm{b}_X(x). \tag{4.4}$$

Note 4.1. The ideia of sharp and vague is relative to the object you are analyzing.

Question 4.1. I didn't understood how could we model loosing evidence and how new observations can replace old ones, although it seems pretty useful.

4.1.3 Dirichlet Visualization of Opinion Vagueness

Example: The singletons and composite values of R(X) are listed below.

$$\left\{\begin{array}{lll} \text{Domain:} & \mathbb{X} & = \{x_1, x_2, x_3\}, \\ \text{Hyperdomain:} & \mathcal{R}(\mathbb{X}) & = \{x_1, x_2, x_3, x_4, x_5, x_6\}, \\ \text{Composite set:} & \mathcal{R}(\mathbb{X}) & = \{x_4, x_5, x_6\}, \end{array}\right. \text{ where } \left\{\begin{array}{l} x_4 = \{x_1, x_2\}, \\ x_5 = \{x_1, x_3\}, \\ x_6 = \{x_2, x_3\}. \end{array}\right. \tag{4.5}$$

$$\begin{array}{c} \text{Belief mass distribution} \\ \left\{\begin{array}{ll} \mathbf{b}_X(x_6) & = 0.8, \\ u_X & = 0.2. \end{array}\right. \end{array} \quad \begin{array}{c} \text{Base rate distribution} \\ \left\{\begin{array}{l} \mathbf{a}_X(x_1) = 0.33, \\ \mathbf{a}_X(x_2) = 0.33, \\ \mathbf{a}_X(x_3) = 0.33. \end{array}\right. \end{array} \tag{4.6}$$

$$
\begin{aligned}
\mathbf{P}_X(x_1) = & \sum_{x_i \in \mathcal{R}(\mathbb{X})} \mathbf{a}_X(x_1|x_i)\mathbf{b}_X(x_i) + \mathbf{a}_X(x_1)u_X \\
= & \frac{\mathbf{a}_X(\{x_1\} \cup \{x_2, x_3\})}{\mathbf{a}_X(\{x_2, x_3\})}\mathbf{b}_X(x_6) + \mathbf{a}_X(x_1)u_X \\
= & \; 0 + 0.33 \cdot 0.2 \\
= & \; 0.066
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{P}_X(x_2) = & \sum_{x_i \in \mathcal{R}(\mathbb{X})} \mathbf{a}_X(x_2|x_i)\mathbf{b}_X(x_i) + \mathbf{a}_X(x_1)u_X \\
= & \frac{\mathbf{a}_X(\{x_2\} \cap \{x_2, x_3\})}{\mathbf{a}_X(\{x_2, x_3\})}\mathbf{b}_X(x_6) + \mathbf{a}_X(x_1)u_X \\
= & \frac{0.33}{0.66} \cdot 0.8 + 0.33 \cdot 0.2 \\
= & \; 0.467
\end{aligned}
$$

$$
\mathbf{P}_X(x_3) = 0.467
$$

$$
\begin{aligned}
\mathbf{b}^{\mathrm{V}}_X(x_1) = & \sum_{\substack{x_i \in \mathcal{C}(\mathbb{X}) \\ x_i \not\subseteq x}} \mathbf{a}_X(x_1|x_i)\mathbf{b}_X(x_i) \\
= & \frac{\mathbf{a}_X(\{x_1\} \cap \{x_2, x_3\})}{\mathbf{a}_X(\{x_2, x_3\})}\mathbf{b}_X(x_6) \\
= & \; 0 \cdot 0.8 \\
= & \; 0
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{b}^{\mathrm{V}}_X(x_2) = & \sum_{\substack{x_i \in \mathcal{C}(\mathbb{X}) \\ x_i \not\subseteq x}} \mathbf{a}_X(x_2|x_i)\mathbf{b}_X(x_i) \\
= & \frac{\mathbf{a}_X(\{x_2\} \cap \{x_2, x_3\})}{\mathbf{a}_X(\{x_2, x_3\})}\mathbf{b}_X(x_6) \\
= & \frac{0.33}{0.66} \cdot 0.8 \\
= & \; 0.4
\end{aligned}
$$

$$
\mathbf{b}^{\mathrm{V}_X}(x_3) = 0.4
$$

$$
\begin{array}{ll}
\text{Projected probability distribution} & \text{Vague belief mass} \\
\begin{cases} \mathbf{P}_X(x_1) = 0.066, \\ \mathbf{P}_X(x_2) = 0.467, \\ \mathbf{P}_X(x_3) = 0.467. \end{cases} &
\begin{cases} \mathbf{b}^{\mathrm{V}}_X(x_1) = 0.0, \\ \mathbf{b}^{\mathrm{V}}_X(x_2) = 0.4, \\ \mathbf{b}^{\mathrm{V}}_X(x_3) = 0.4. \end{cases}
\end{array}
\tag{4.7}
$$

Figure 4.2 from the book shows the hyper-Dirichlet PDF for this vague opinion.

4.1.4 Focal Uncertainty Mass

Definition 4.5. *(Focal Uncertainty Mass)* Let $\mathbb{X}$ be a domain and $\mathcal{R}(\mathbb{X})$ denote itshyperdomain. Given an opinion ω_X, the focal uncertainty mass of an value $x \in \mathcal{R}(\mathbb{X})$is computed with the function $\mathbf{u}_X^{\mathrm{F}} : \mathcal{R}(\mathbb{X}) \to [0, 1]$ defined as

$$\text{Focal uncertainty mass: } \mathbf{u}_X^{\mathrm{F}}(x) = \mathbf{a}_X(x) u_X. \tag{4.8}$$

4.2 Mass Sum

4.2.1 Mass Sum of a Value

The sum of sharp belief mass, vague belief mass and focal uncertainty mass is of a value x is equal to the value's projected probability, expressed as:

$$\mathbf{b}_X^S(x) + \mathbf{b}_X^V(x) + \mathbf{u}_X^F(x) = \mathbf{P}_X(x) \tag{4.9}$$

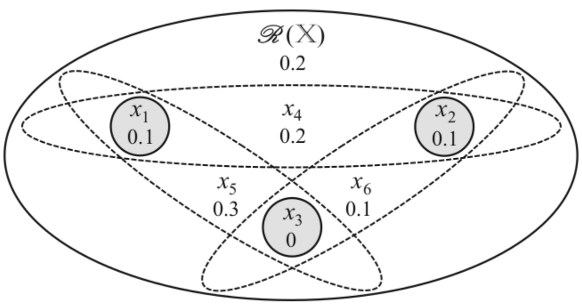
Equation 4.9 shows that the projected probability can be split in three parts: sharp belief mass, vague belief mass and focal uncertainty mass. The composition of these three parts is called *mass-sum* and is denoted $M_X(x)$. The concept of mass sum is defined next.

Definition 4.6. Let $\mathbb{X}$ be a domain with hyperdomain $\mathscr{R}(\mathbb{X})$ and assume that the opinion ω_X is specified. Consider a value $x \in \mathscr{R}(\mathbb{X})$ with sharp belief mass $\mathbf{b}_X^S(x)$, vague belief mass $\mathbf{b}_X^V(x)$ and focal uncertainty mass $\mathbf{u}_X^F(x)$. The mass-sum function of valuef x is the triplet denoted $M_X^E(x)$ expressed as:

$$\text{Mass-sum of } x\text{: } \mathbf{M}_X(x) = \big(\mathbf{b}_X^S(x), \mathbf{b}_X^V(x), \mathbf{u}_X^F(x)\big).$$

Thus, each value x in the hyperdomain has its own mass-sum.

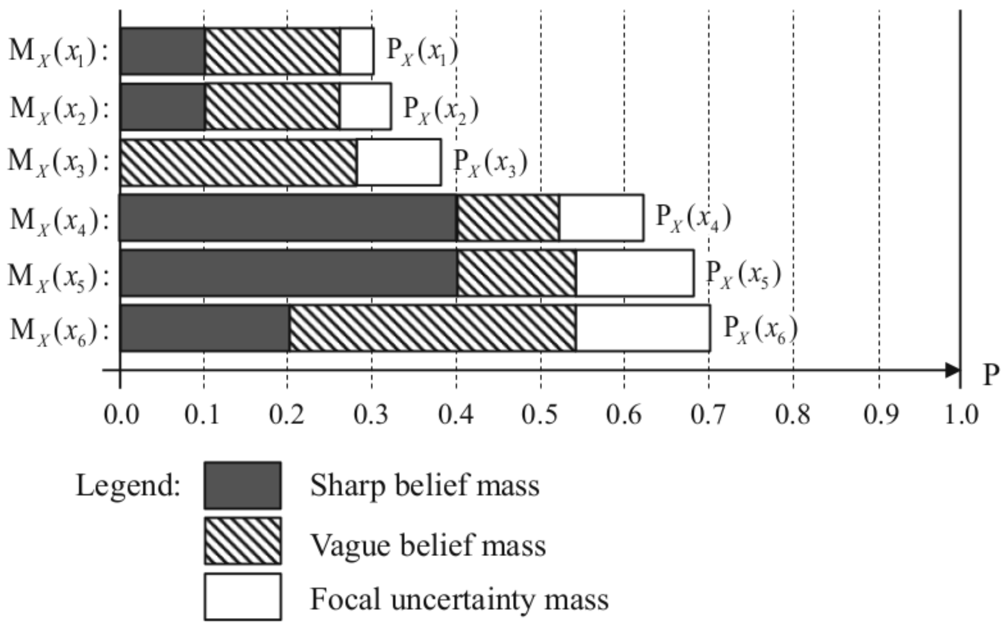
In order to visualize it, consider the ternary domain $\mathbb{X} = \{x_1, x_2, x_3\}$:



The following table shows the projected probability of each value of x.

Value x	Belief mass $\boldsymbol{b}_X(x)$	Base rate $\boldsymbol{a}_X(x)$	Sharp belief mass $\boldsymbol{b}_X^{\mathrm{S}}(x)$	Vague belief mass $\boldsymbol{b}_X^{\mathrm{V}}(x)$	Focal uncertainty mass $\boldsymbol{u}_X^{\mathrm{F}}(x)$	Projected probability $\mathbf{P}_X(x)$
x_1	0.10	0.20	0.10	0.16	0.04	0.30
x_2	0.10	0.30	0.10	0.16	0.06	0.32
x_3	0.00	0.50	0.00	0.28	0.10	0.38
x_4	0.20	0.50	0.40	0.12	0.10	0.62
x_5	0.30	0.70	0.40	0.14	0.14	0.68
x_6	0.10	0.80	0.20	0.34	0.16	0.70
u_X	0.20					

The *mass-sums* can be easily visualized by a mass sum diagram. Since opinions in higher dimension can't be visualized by a simplex, a mass sum diagram makes easier to compare opinions and appreciate its nature.



Although x_3 has the highest projected probability, it has no sharp belief mass. Differentiating between those parts of an opinion might be important for decision making and will be showed bellow.

4.2.2 Total Mass-Sum

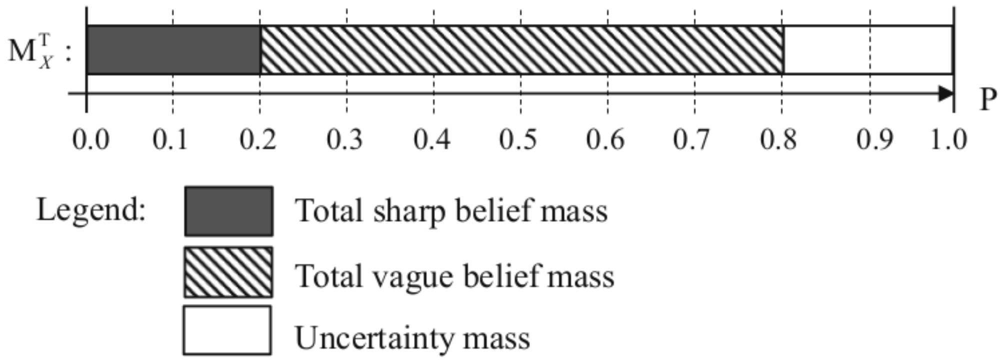
The belief mass of an opinion as a whole can be decomposed into sharp belief mass which provides distinctive support for singletons, and vague belief mass which provides vague support for singletons. These two belief masses are then complementary to the uncertainty mass. For any opinion equation 4.10 holds.

$$b_X^{TS} + b_X^{TV} + u_X = 1. \tag{4.10}$$

Equation 4.10 shows that the belief and uncertainty mass can be split into the three parts of sharp/vague belief mass and uncertainty mass. The composition of these three parts is called total mass-sum, denoted $\mathbf{M}_X^T$, can be represented with a triplet and is defined below.

Definition 4.7. Total mass-sum: $M_X^T = (b_X^{TS}, b_X^{TV}, u_X)$.

The total mass-sum of opinion ω_X is illustrated bellow:



4.3 Utility and Normalization

Assume a random variable X with an associated projected probability distribution $\mathbf{P}_X$. Utility is typically associated with outcomes of a random variable, in the sense that for each outcome x there is an associated utility $\boldsymbol{\lambda}_X(x)$ expressed on some scale such as monetary value, which can be positive or negative. Given utility $\boldsymbol{\lambda}_X(x)$ in case of outcome x, then the expected utility for x is

$$\text{Expected utility: } \mathbf{L}_X(x) = \boldsymbol{\lambda}_X(x)\mathbf{P}_X(x).$$

Total expected utility for the variable X is then:

$$\text{Total expected utility: } L_X^T = \textstyle\sum_{x\in\mathbb{X}} \boldsymbol{\lambda}_X(x)\mathbf{P}_X(x)$$

In classical utility theory, decisions are based on expected utility for possible options. It is also possible to eliminate the notion of utility by integrating it into the probabilities for the various options, which produces a *utility-normalised probability vector*. This approach greatly simplifies decision-making models, because every option can be represented as a simple utility-probability. Normalisation is useful when comparing options of variables from different domains, where the different variables have different associated probability distributions and utility vectors. The normalisation factor must be appropriate for all variables, so that the utility-normalised probability vectors are within a given range. Note that in case of negative utility for a specific outcome, the utility-normalised probability for that outcome is also negative. In that sense, utility-normalised probability represents synthetic probability, and not realistic probability.

Given a set of variables, with associated probability distributions and utility vectors, let λ^+ denote the absolute value of the greatest absolute utility of all utilities in all vectors. The utility-normalised probability vector $\mathbf{P}_X^T$ is defined below.

Definition 4.8. Assume a random variable X with a projected probability distribution $\mathbf{P}_X$. Let $\mathbf{b}_X^S(x)$ denote the sharp belief mass of x, let $\mathbf{b}_X^V(x)$ denote the vague belief mass of x, and let $\mathbf{u}_X^F(x)$, denote the focal uncertainty mass of x. Assume the utility vector $\boldsymbol{\lambda}_X$, as well as λ^+. The utility-normalized masses are expressed as

Utility-normalized sharp belief mass: $\mathbf{b}_X^{NS}(x) = \frac{\boldsymbol{\lambda}_X(x)\mathbf{b}_X^S(x)}{\lambda^+}, \forall x \in \mathbb{X}$

Utility-normalized vague belief mass: $\mathbf{b}_X^{NV}(x) = \frac{\boldsymbol{\lambda}_X(x)\mathbf{b}_X^V(x)}{\lambda^+}, \forall x \in \mathbb{X}$

Utility-normalized focal uncertainty mass: $\mathbf{u}_X^{NF}(x) = \frac{\boldsymbol{\lambda}_X(x)\mathbf{u}_X^F(x)}{\lambda^+}, \forall x \in \mathbb{X}$

There is a additivity property on belief masses:

$$\text{Utility-normalized probability: } \mathbf{b}_X^{NS}(x) + \mathbf{b}_X^{NV}(x) + \mathbf{u}_X^{NF}(x) = \mathbf{P}_X^N(x).$$

Having defined utility-normalized probability, it is possible to directly compare options without involving utilities, because utilities are integrated into the utility normalized probabilities. It is also possible to define a utility normalized mass-sum:

Definition 4.9. Let $\mathbb{X}$ be a domain with hyperdomain $\mathscr{R}(\mathbb{X})$, and assume that the opinion ω_X is specified. Also assume that utility vector $\boldsymbol{\lambda}_X$ is specified. Consider a value $x \in \mathscr{R}(\mathbb{X})$ with the utility normalized masses $b_X^{NS}(x)$, $b_X^{NV}(x)$ and $u_X^{NS}(x)$. The utility-normalized mass-sum function of x is the triplet denoted $\mathbf{M}_X^N(x)$ expressed as:

$$\text{Utility-normalized mass-sum: } \mathbf{M}_X^N(x) = \big(b_X^{NS}(x), b_X^{NV}(x), u_X^{NS}(x)\big)$$

Note 4.2. Note that utility-normalised sharp belief mass, vague belief mass and focal uncertainty mass do not represent realistic masses, and must be considered as purely synthetic.

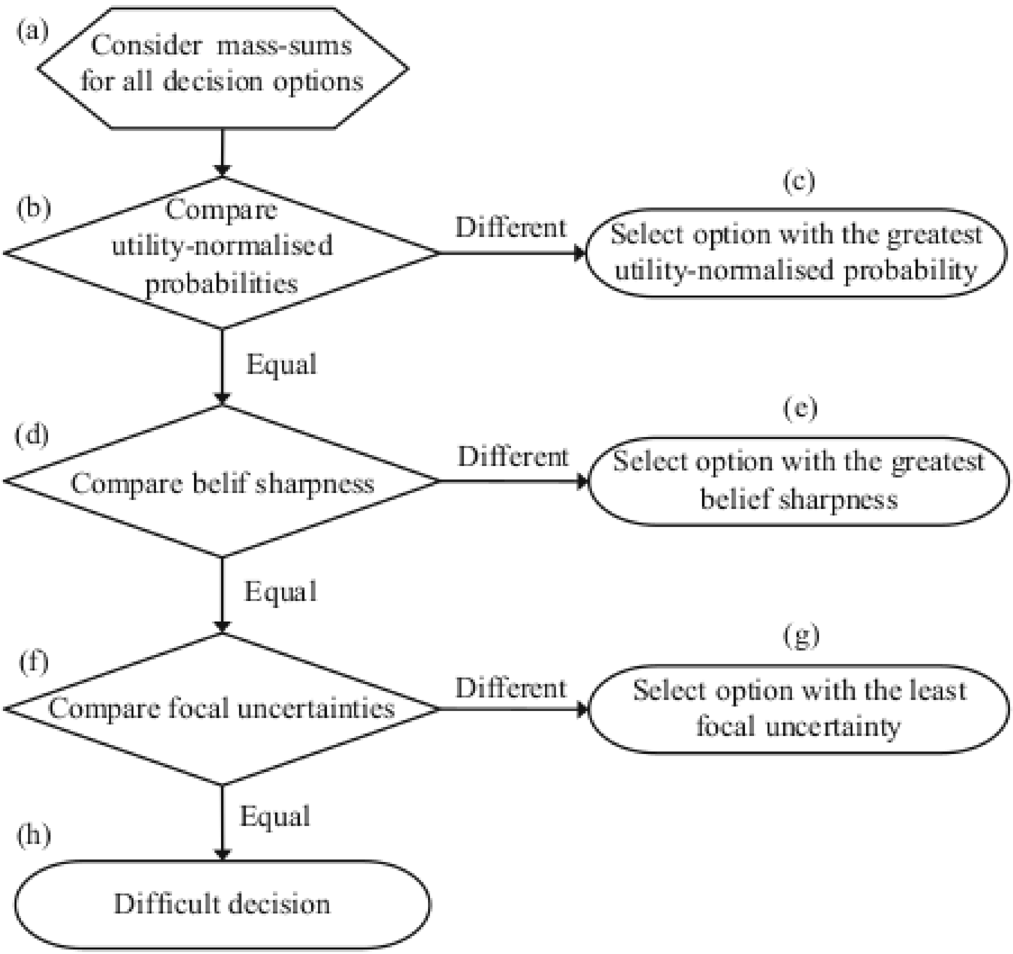
Now the book provides a good example on how to decide based on the concepts defined above and also some other ideas in the case that the projected utility is equal among some elements. These ideas will be better described in the chapter bellow.

Question 4.2. Why does it normalized by the greater element, doesn't it affect how you compare different domains with different maximum elements?

Note 4.3. When the utility is equal for all options, the mass-sum equals the utility-normalized mass-sum. Thus in this case it is not needed to normalize it.

4.4 Decision criteria

When deciding using opinions in subjective logic, the analyst should decide according to a priority scale defined by the author in the following flow chart:



Thus, dividing our opinion in different parts provided us a better way to decide upon uncertainty, since we can now differentiate where the belief came from.