Dealing With Uncertainty

Probability
Measure for Uncertain Events

• We will be interested in quantifying our uncertainty about various ‘inputs’ or variables in a system.

• This can be done using probability.
Overview of Probability

• Probability is about measuring the ‘belief’ we have regarding uncertain variables.
• Formal probability is a branch of pure mathematics, concerned with the properties of such measures.
• Intuitively, we have an understanding of how to ‘do’ probability calculations.
Example

• Let $X$ denote the profit that will be made.
• Then we are interested in ‘measuring’ the chance that $X$ is $\geq$ zero.
• Typically, if $z = [0, +\infty)$
  – This is just $P(z)$. 
Properties

• Coherence

• Let $X$ be an uncertain quantity.

  – $P(x)$ is a value between 0 and 1. (Cx)

• In general, $x$ is a subset of the possible values that $X$ can take.
Properties (ctd)

• Additivity

• For disjoint or exclusive $x$ and $y$:

  \[ P(x \cup y) = P(x) + P(y). \quad (Ad) \]

• In particular, $P(x) + P(y) = 1$, if $x$ and $y$ ‘cover’ $X$. 


Properties (ctd)

• **Multiplicative**

• The notation $P(a|b)$ means the restricted measure of ‘a’, to the set where ‘b’ is true.

• This can be pronounced, “The probability of $a$, given $b$.”

• For general $x,y$, and where $P(y)$ not zero,
  – $P(x,y) = P(y)P(x|y)$. (Mu)

• Thus, Cx, Ad and Mu are properties of probability.
On Conditionals

• P(a|b) is also spoken, “The probability of a conditional on b.”

• Where we are discussing ‘real’ events, and probability is linked to ‘belief’, then we always have a conditioning set.

• Typically, this is our prior knowledge, or our history, H.

• Thus, we should always write P(x|H).
Foundations

- Cx, Ad and Mu are foundations.
- We can build arguments upon these, and derive other results.
- However, if we are being consistent, we cannot breach these.
- Some other results follow – but these are derived.
- The foundations are axioms.
Extend the Conversation

• Let e, f, g be mutually exclusive, and exhaustive events – that is to say that they do not contain material overlap, and the combination of them covers the domain.

• Then for any X of interest;
  \[ P(X) = P(X|e)P(e) + P(X|f)P(f) + P(X|g)P(g) \]
Example

- Let us say that response rate can either be low (l) or high (h).
- Then, with X being profit, as before, and z being non-negative values;
  - \( P(z) = P(z|h)P(h) + P(z|l)P(l) \)
- Thus, in order to answer questions about profit, we allow considerations regarding response rate to enter our discussion.
Bayes Theorem

• This theorem allows the inversion of conditionals.
• Thus, if we start by considering the probability of profit, given response rates we could ‘invert’ to say something about response rates, given profit.
• It is stated thus;
  – \( P(A|B)P(B)=P(B|A)P(A) \)
• This may be derived using Mu.
Example

• Let D be the event of dying before 50, and L the event of living beyond 50. Let S be the event that someone smokes, and N the event that they don’t.

• In a cohort of 1,000,000, there were 412,000 smokers. 153,000 deaths were observed in this group before the age of 50, 78,000 of these were smokers.

• Write down P(S), P(D), P(S|D) and P(D|S).
Example - Classic

- A test for a disease is 99% accurate (that is 99% of the time it will give the ‘correct’ answer, be it positive or negative.)
- The disease has a prevalence of 1 in 1,000,000 (very rare) in the general population.
- Write down the various probabilities described and of interest here. Write down the probability of having the disease, given that you are a random individual who has tested positive.
Problem

• A man can’t find his wallet. However, he is 75% sure it is in his work suit – the other possibility is that it could be in his dinner jacket.
• If it is in his dinner jacket, there is a 50% chance of either pocket.
• With his work suit, it is twice as likely to be in his left than his right pocket (because of habit.)
• What are the probabilities of being in each of the possible locations?
Interpretations

• Everything we have looked at so far is ‘classical’ probability
  – large population
  – “frequent”

• The next step is more controversial.

• Probability coherently measures beliefs, and allows us to reason with uncertainty.

• Thus, for example, the man specified various ‘beliefs’ and then infers that the most likely location for the wallet to be was his left suit pocket.
Continuous Events

- The example earlier discussed ‘High’ and ‘Low’ response rates.
- However, response rate is a variable that has realisations on a continuum.
- Thus, we need probability measures that work for such variables.
- These exist, and follow the rules as described earlier.
Cts Probabilities

• In this case, however, we consider a ‘function’ which gives us an idea of how likely various values are for the variable of interest.

• To get the probability that a variable takes a value in a range, $X$, we integrate the function across the range.

• Thus $P(X) = \int_X f(x) \, dx$
Follow Through

• For finite X, then,
  – (Cx) $\int_x f(x)dx$ is in $[0,1]$
  – (Ad) Exc $f(xUy)=f(x)+f(y)$
  – (Mu) $f(xy) = f(x|y)f(y)$.

• Marginalisation and Bayes
  – Ext $f(x)=\int f(x|y)f(y)dy$
  – Bay $f(x|y)f(y) = f(y|x)f(x)$
Summary

• Probability can be used to quantify uncertainty.
• Probability can be described using a number of axioms.
• Some rules follow from the axioms.
• Continuous as well as discrete versions exist.