The Probability Calculation

Condition probability is defined by the relationship $P(V|S) = \frac{P(V \cap S)}{P(S)}$

Hence $P(V|S)P(S) = P(V \cap S)$

and since $P(V \cap S) = P(S \cap V)$ we have $P(V|S)P(S) = P(S|V)P(V)$

giving the simple Bayes $P(S|V) = \frac{P(V|S)P(S)}{P(V)}$

Now let *V* be the set of values for a derived variables and $v \subset V$
Also let *S* be the set of states for the system $s \subset S = (fraud, legal)$

The system can only be in one of the states states *s.* In this particular case there are only two possible states and we have $P(fraud) + P(legal) = 1$

In general we seek $P(s = fraud|v)$ and from Bayes we have

$$P(fraud|v) = \frac{P(v|fraud)P(fraud)}{P(v)} \qquad (1)$$

> We know that $P(fraud) + P(legal) = 1$ and as either *legal* or *fraud* must occur we can therefore also write
> $P(fraud|v) + P(legal|v) = 1$

<u>This then leads to the following:</u>

As $P(fraud \cap v) = P(fraud|v)P(v)$ and $P(legal \cap v) = P(legal|v)P(v)$

we can therefore write

$$P(fraud \cap v) + P(legal \cap v) = [P(fraud|v) + P(legal|v)]P(v) = P(v)$$

and then using $P(fraud \cap v) = P(fraud|v)P(v)$,etc we have

$$P(v|fraud)P(fraud) + P(v|legal)P(legal) = P(v)$$

So we can write (1) as:

$$P(fraud|v) = \frac{P(v|fraud)P(fraud)}{P(v|fraud)P(fraud) + P(v|legal)P(legal)} = \frac{1}{1 + \frac{P(v|legal)P(legal)}{P(v|fraud)P(fraud)}}$$

If we consider this formula in the context of the frequency distributions calculated for fraud *F(v)* and legal *L(v)* then we have:

$$P(fraud \mid v) = \frac{(F(v)/N_F)(N_F/N)}{(F(v)/N_F)(N_F/N) + (L(v)/N_L)(N_L/N)} = \frac{F(v)}{F(v)+L(v)}$$

This is the calculation performed by the Risk Engine.