Bayesian Probability Concepts

An important consideration in probability is the revision of the probability of an event *A* consequent to the occurrence of event *z*. The Bayesian approach uses several terms in updating probabilities. Suppose that the possible values of a parameter *A* were assumed to be a discrete set of values $a_i, i = 1,2,...,k$, with prior relative likelihoods $P(A = a_i)$ where *A* is the RV whose values represent possible values of the parameter a. As additional information becomes available, the prior assumptions on the parameter a [1]may be formally modified through Bayes Theorem.

To describe in equation form, let *z* = outcome of an experiment. Then applying Bayes theorem, we can obtain the updated probabilities as

$$P(A = a_i \mid z) = \frac{P(z \mid A = a_i)P(A = a_i)}{\sum_{i=1}^{k} P(z \mid A = a_i)P(A = a_i)} \qquad i = 1,2,...,k$$

z = experimental outcome

$P(z \mid A = a_i)$ = the likelihood of the experimental outcome z if $A = a_i$; that is, the conditional probability of obtaining a particular experimental outcome assuming the parameter is a_i.

$P(A = a_i)$ = the prior probability of $A = a_i$, that it, prior to the availability of experimental information z.

$P(A = a_i \mid z)$ = the posterior probability of $A = a_i$, that is, the probability that has been revised in light of the experimental outcome z.

$P(A = a_i)$ = the prior probability of $A = a_i$, that it, prior to the availability of experimental information z.

The expected value of A is commonly used as the Bayesian estimator of the parameter:

$$\hat{a} = E(A \mid z) = \sum_{i=1}^{k} a_i * [\text{posterior probability of } A = a_i]$$

In general, by updating prior probabilities, an engineer can assess the likelihood of design events by incorporating additional information given by conditioned posterior probabilities.

If one defines as "state" the unknown quantification of the population and considers some sample of observations is available, Bayes' theorem can be written as

[1] References: Kottegoda and Rosso; Ang and Tang.

$$P(\text{state}|\text{sample}) = \frac{P(\text{sample} \mid \text{state})P(\text{state})}{\sum_{\text{all states}} P(\text{sample}|\text{state})P(\text{state})}.$$

In practice, an engineer often has prior knowledge of the occurrences of different states of a population. In addition, there is frequently access to data from which one can estimate the likelihood of a measurable quantity or sample of data given the true state of the population.

Example.
An engineer designing a foundation for a tall structure needs to know the depth *h* of a soil above bedrock at the site. For preliminary design purposes the depth is divided into four states:

$$B_1 = \{h \le 5m\}$$
$$B_2 = \{5m < h \le 10m\}$$
$$B_3 = \{10m < h \le 15m\}$$
$$B_4 = \{h > 15m\}$$

The engineer consults a local geologist who, from knowledge of the geology of that area, assigns prior probabilities to the four states as follows:

$$P(B_1) = 0.60$$
$$P(B_2) = 0.20$$
$$P(B_3) = 0.15$$
$$P(B_4) = 0.05$$

For measuring the depth to bedrock, a seismic recorder is used, which is subject to some error. From previous experience the geologist estimates the conditional probabilities that the instrument indicates a particular states out of four states (The sum of the probabilities of which is 1.0). The likelihoods are given in the table below. The reading of the instrument is $h = 7\ m$, which is referred to as *Sample 1*. This corresponds with state B_2.

Likelihood that the seismic recorder showing the conditional probability of depths to bedrock indicates state B_i given that the true state is B_j.

	True State, B_j			
Measured State, B_i	$j=1$ $h \le 5m$	$j=2$ $5m < h \le 10m$	$j=3$ $10m < h \le 15m$	$j=4$ $h > 15m$
$i=1$ $h \le 5m$	0.90	0.05	0.03	0.02
$i=2$ $5m < h \le 10m$	0.07	0.88	0.15	0.06
$i=3$ $10m < h \le 15m$	0.03	0.05	0.85	0.12
$i=4$ $h > 15m$	0.00	0.02	0.06	0.80
Sum	1.00	1.00	1.00	1.00

The posterior probabilities of the actual states of nature are evaluated as

$$P(B_k \mid sample\ no.\ 1) = \frac{P(sample\ no.\ 1 \mid B_k)P(B_k)}{\sum_{i=1}^{4} P(sample\ no.\ 1 \mid B_i)P(B_i)}$$

where

$$\sum_{i=1}^{4} P(\{5m < h \leq 10m\} \mid B_i)(P(B_i) = 0.07(0.60) + 0.88(0.20) + 0.10(0.15)+).06(0.05) = 0.236$$

$$P(B_1 \mid \text{sample no. } 1) = \frac{0.07(0.60)}{0.236} = 0.178$$

$$P(B_2 \mid \text{sample no. } 1) = \frac{0.88(0.20)}{0.236} = 0.746$$

$$P(B_3 \mid \text{sample no. } 1) = \frac{0.10(0.15)}{0.236} = 0.063$$

$$P(B_4 \mid \text{sample no. } 1) = \frac{0.06(0.05)}{0.236} = 0.13$$

Because there is still a chance of about 25% that the true state may not be B2, a second test is made and the reading is 8m. Thus sample no. 2 also indicates state B2.

Using the foregoing posterior probabilities (after sample no. 1) as revised prior probabilities for the site, we find posterior probabilities (after sample no. 2) as follows:

$$\sum_{i=1}^{4} P(\{5m < h \leq 10m\} \mid B_i)P(B_i) = 0.07(0.178)+).88(0.746) + 0.10(0.063) + 0.06(0.013) = 0.675$$

$$P(B_1 \mid \text{sample nos. 1 and 2}) = \frac{0.07(0.178)}{0.675} = 0.018$$

$$P(B_2 \mid \text{sample nos. 1 and 2}) = \frac{0.88(0.746)}{0.675} = 0.972$$

$$P(B_3 \mid \text{sample nos. 1 and 2}) = \frac{0.10(0.063)}{0.675} = 0.009$$

$$P(B_4 \mid \text{sample nos. 1 and 2}) = \frac{0.06(0.013)}{0.675} = 0.001$$

It is now evident that the chance that the true state is NOT B_2 is very small. The engineer may therefore proceed on the assumption that the depth to rock is in the range of 5 to 10 m.