**Prior Distribution**

Bayesian procedures formally utilize information available from sources other than the statistical investigation. Such information, available through expert judgment, past experience, or prior belief, is described by a probability distribution on the set of all possible values of the unknown parameter of the statistical model at hand. This probability distribution is called the prior distribution.

The choice of the prior distribution for the unknown parameter *θ* is of crucial importance in Bayesian statistics. The selected prior distribution for *θ* must be at least a reasonable approximation to the true beliefs about *θ*. In addition, the prior distribution must be such that the posterior distribution is tractable. Prior Distribution quantifies information about parameter prior to any data being gathered.

A prior probability is a marginal probability interpreted as a description of what is known about a variable in the absence of some evidence. The posterior probability is then the conditional probability of the variable taking the evidence into account. The posterior probability is computed from the prior and the likelihood function via Bayes’ Theorem.

**Mathematical Formulation**

The prior distribution for the parameter  is the function . Strictly, if  is discrete-valued parameter,  is the probability that the parameter takes any particular value , while in the more usual cases where  is a continuous variable and  is a probability-density function.

**Formulating Prior Distribution**

Where genuine, substantial prior information exists it needs to be based on defensible evidence and reasoning. This is clearly important when the new data are not so extensive as to overwhelm the prior information, so the Bayes’ theorem will give the prior a non-negligible weight in its synthesis with the data. Prior information of this kind exists routinely in all kinds of applications. If prior information is vague and insubstantial, then it will get negligible weight in the synthesis with the data and the posterior will in effect be based entirely on data information.So, the most important consideration in the use of prior information is to ensure that the prior distribution honestly reflects genuine information, not personal bias, prejudice, superstition or other such factors that are justly condemned in science as ‘subjectivity’.

Prior information should be based on sound evidence and reasoned judgments. A good way to think of this is to parody a familiar quotation: the prior distribution should be ‘the evidence, the whole evidence and nothing but the evidence’:

* ‘The evidence’: genuine information legitimately interpreted.
* ‘The whole evidence’: not omitting relevant information (preferably a consensus that pools the knowledge of a range of experts).
* ‘Nothing but the evidence’: Not contaminated by bias or prejudice.

To gain the full benefit of the Bayesian approach, genuine prior information should not be ignored. However, the process of converting that information into its formal expression as a prior distribution is not straightforward. A particularly difficulty arises when, as is often the case, the person who has useful prior information about the parameters in some problem is not a statistician but a specialist in the application area. Formulating a prior distribution then typically involves a dialogue between the expert and a statistically trained facilitator. This process is called ‘elicitation’, and has been studied both by statisticians and psychologists (O’Hagan, 2003).