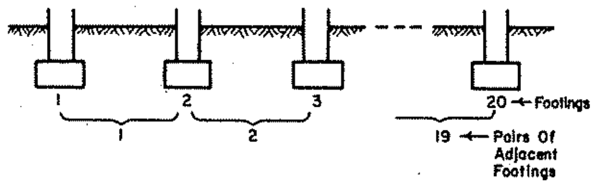

Figure P7.14

exists between the settlement behavior of two adjacent footings. The following is a set of data on the settlement of a series of footings on sand.

Footing	Settlement (in.)	Footing	Settlement (in.)
1	0.59	11	0.93
2	0.60	12	0.78
3	0.54	13	0.78
4	0.70	14	0.77
5	0.75	15	0.79
6	0.80	16	0.79
7	0.79	17	0.78
8	0.95	18	0.77
9	1.00	19	0.63
10	0.92	20	0.73

From a row of 20 footings, 19 pairs of adjacent footings can be obtained as shown in Fig. P7.14. The degree of dependence between the settlements of adjacent footings is described by the correlation coefficient.

(a) Estimate this correlation based on the 19 pairs of data. *Ans. 0.766.*
(b) Estimate the coefficient of variation of the settlement of a footing. *Ans. 0.157.*

8. The Bayesian Approach

8.1. INTRODUCTION

In Chapter 5 we presented the methods of point and interval estimation of distribution parameters, based on the *classical statistical* approach. Such an approach assumes that the parameters are constants (but unknown) and that sample statistics are used as estimators of these parameters. Because the estimators are invariably imperfect, errors of estimation are unavoidable; in the classical approach, confidence intervals are used to express the degree of these errors.

As implied earlier, accurate estimates of parameters require large amounts of data. When the observed data are limited, as is often the case in engineering, the statistical estimates have to be supplemented (or may even be superseded) by judgmental information. With the classical statistical approach there is no provision for combining judgmental information with observational data in the estimation of the parameters.

For illustration, consider a case in which a traffic engineer wishes to determine the effectiveness of the road improvement at an intersection. Based on his experience with similar sites and traffic conditions, and on a traffic-accident model, he estimated that the average occurrences of accidents at the improved intersection would be about twice a year. However, during the first week after the improved intersection is opened to traffic, an accident occurs at the intersection. A dichotomy, therefore, may arise: The engineer may hold strongly to his judgmental belief, in which case he would insist that the accident is only a chance occurrence and the average accident rate remains *twice a year*, in spite of the most recent accident. However, if he only considers actual observed data, he would estimate the average accident rate to be *once a week*. Intuitively, it would seem that both types of information are relevant and ought to be used in determining the average accident rate. Within the classical method of statistical estimation, however, there is no formal basis for such analysis. Problems of this type are formally the subject of Bayesian estimation.

The *Bayesian method* approaches the estimation problem from another point of view. In this case, the unknown parameters of a distribution are assumed (or modeled) to be also random variables. In this way, uncertainty associated with the estimation of the parameters can be combined formally

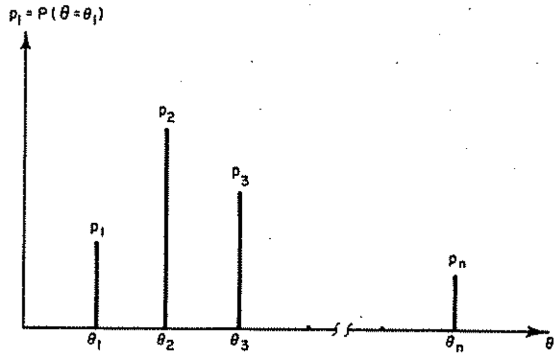

Figure 8.1 Prior PMF of parameter θ

(through Bayes' theorem) with the inherent variability of the basic random variable. With this approach, subjective judgments based on intuition, experience, or indirect information are incorporated systematically with observed data to obtain a balanced estimation. The Bayesian method is particularly helpful in cases where there is a strong basis for such judgments. We introduce the basic concepts of the Bayesian approach in the following sections.

8.2. BASIC CONCEPTS—THE DISCRETE CASE

The *Bayesian approach* has special significance to engineering design, where available information is invariably limited and subjective judgment is often necessary. In the case of parameter estimation, the engineer often has some knowledge (perhaps inferred intuitively from experience) of the possible values, or range of values, of a parameter; moreover, he may also have some intuitive judgment on the values that are more likely to occur than others. For simplicity, suppose that the possible values of a parameter θ were assumed to be a set of discrete values θ_i, $i = 1, 2, \ldots, n$, with relative likelihoods $p_i = P(\Theta = \theta_i)$ as illustrated in Fig. 8.1 (Θ is the random variable whose values represent possible values of the parameter θ).

Then if additional information becomes available (such as the results of a series of tests or experiments), the prior assumptions on the parameter θ may be modified formally through Bayes' theorem as follows.

Let ϵ denote the observed outcome of the experiment. Then applying

Bayes' theorem of Eq. 2.20, we obtain the updated PMF for Θ as

$$P(\Theta = \theta_i \mid \epsilon) = \frac{P(\epsilon \mid \Theta = \theta_i)\, P(\Theta = \theta_i)}{\sum_{i=1}^{n} P(\epsilon \mid \Theta = \theta_i)\, P(\Theta = \theta_i)} \qquad i = 1, 2, \ldots, n \tag{8.1}$$

The various terms in Eq. 8.1 can be interpreted as follows:

$P(\epsilon \mid \Theta = \theta_i)$ = the likelihood of the experimental outcome ϵ if $\Theta = \theta_i$; that is, the conditional probability of obtaining a particular experimental outcome assuming that the parameter is θ_i

$P(\Theta = \theta_i)$ = the *prior* probability of $\Theta = \theta_i$; that is, prior to the availability of the experimental information ϵ

$P(\Theta = \theta_i \mid \epsilon)$ = the *posterior* probability of $\Theta = \theta_i$; that is, the probability that has been revised in the light of the experimental outcome ϵ

Denoting the *prior* and *posterior* probabilities as $P'(\Theta = \theta_i)$ and $P''(\Theta = \theta_i)$, respectively, Eq. 8.1 becomes

$$P''(\Theta = \theta_i) = \frac{P(\epsilon \mid \Theta = \theta_i) P'(\Theta = \theta_i)}{\sum_{i=1}^{n} P(\epsilon \mid \Theta = \theta_i) P'(\Theta = \theta_i)} \tag{8.1a}$$

Equation 8.1*a*, therefore, gives the posterior probability mass function of Θ. (In general, we shall use $'$ and $''$ to denote the prior and posterior).

The *expected value* of Θ is then commonly used as the *Bayesian estimator** of the parameter; that is,

$$\hat{\theta}'' = E(\Theta \mid \epsilon) = \sum_{i=1}^{n} \theta_i P''(\Theta = \theta_i) \tag{8.2}$$

We may point out that in Eq. 8.2 observational data and judgmental information are both used and combined in a systematic way to estimate the underlying parameter.

In the Bayesian framework, the significance of judgmental information is reflected also in the calculation of relevant probabilities. In the case above, where subjective judgments were used in the estimation of the parameter θ, such judgments would be reflected in the calculation of the probability, for example, $P(X \leq a)$, through the theorem of total proba-

* There are other Bayesian estimators depending on the assumed form of the "loss function" (discussed in Vol. II). Moreover, other parameters of the posterior distribution may serve as the estimator instead; for example, the *mode*.

bility using the posterior PMF of Eq. 8.1a. That is,

$$P(X \leq a) = \sum_{i=1}^{n} P(X \leq a \mid \Theta = \theta_i)\, P''(\Theta = \theta_i) \qquad (8.3)$$

This represents the up-to-date probability of the event $(X \leq a)$ based on all available information. It may be emphasized that in Eq. 8.3 the uncertainty associated with the error of estimating the parameter [as reflected in $P''(\Theta = \theta_i)$] is combined with the inherent variability of the random variable X.

To clarify these general concepts, consider the following examples.

EXAMPLE 8.1

Piles for a building foundation were initially designed for 250-ton capacity each; however, this did not include the effect of high winds that occur only very rarely. On such rare occasions, it is estimated that some of the piles may be subjected to loads as high as 300 tons. In order to assess the safety of the initial design, the engineer in charge wishes to determine the probability of the piles failing under a maximum load of 300 tons.

Suppose that from the engineer's experience with this type of piles and the soil condition at the site, he estimated (judgmentally) that the probability p would range from 0.2 to 1.0 with 0.4 as the most likely value; more specifically, p is described by the prior PMF shown in Fig. E8.1a. The values of p are discretized at 0.2 intervals to simplify the presentation.

On the basis of this prior PMF, the estimated probability of a pile failing at a load of 300 tons would be (by virtue of the total probability theorem)

$$\begin{aligned} \hat{p}' &= (0.2)(0.3) + (0.4)(0.4) + (0.6)(0.15) + (0.8)(0.10) + (1.0)(0.05) \\ &= 0.44 \end{aligned}$$

In order to supplement his judgment, the engineer ordered a pile of the same type test-loaded at the site to a maximum load of 300 tons. The outcome of the test shows that the pile failed to carry the maximum load. Based on this single test result, the PMF of p would be revised according to Eq. 8.1a, obtaining the posterior PMF as follows:

$$\begin{aligned} P''(p = 0.2) &= \frac{(0.2)(0.3)}{(0.2)(0.3) + (0.4)(0.4) + (0.6)(0.15) + (0.8)(0.1) + (1.0)(0.05)} \\ &= 0.136 \end{aligned}$$

and, similarly,

$$\begin{aligned} P''(p = 0.4) &= 0.364 \\ P''(p = 0.6) &= 0.204 \\ P''(p = 0.8) &= 0.182 \\ P''(p = 1.0) &= 0.114 \end{aligned}$$

which are shown graphically in Fig. E8.1b.
The Bayesian estimate for p, Eq. 8.2, therefore is

$$\begin{aligned} \hat{p}'' = E(p \mid \varepsilon) &= 0.2(0.136) + 0.4(0.364) + 0.6(0.204) + 0.8(0.182) + 1.0(0.114) \\ &= 0.55 \end{aligned}$$

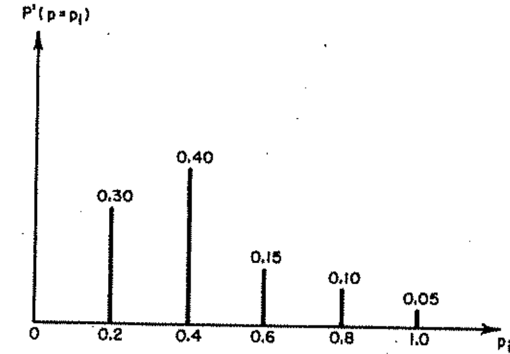

Figure E8.1a Prior PMF of p

In Fig. E8.1b, we see that as a result of the single unsuccessful load test, the probabilities for higher values of p_i are increased from those of the prior distribution, resulting in a higher estimate for p, namely, $\hat{p}'' = E(p \mid \varepsilon) = 0.55$, whereas the prior estimate was 0.44. Observe that the failure of one test pile does not imply the impossibility of such piles carrying the 300-ton load; instead, the test result merely serves to increase the estimated probability by 0.11 (from 0.44 to 0.55). Figure E8.1c illustrates how the PMF of p changes with increasing number of consecutive test pile failures; the distribution shifts toward $p = 1.0$ as $n \to \infty$.

Figure E8.1d shows the corresponding Bayesian estimate for p; observe that after a sequence of 6 consecutive failures the estimate for p is 0.90. If a long sequence of failures is observed, the Bayesian estimate of p approaches 1.0—a result that tends to the classical estimate; in such a case, there is overwhelming amount of observed data

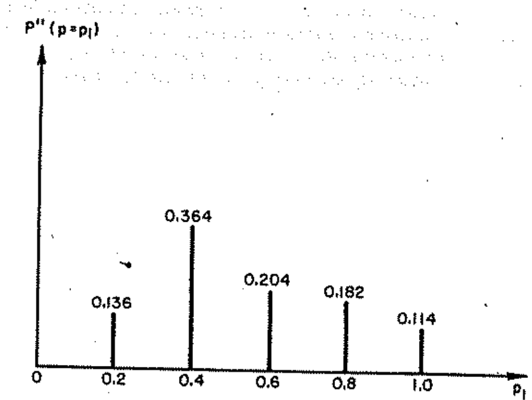

Figure E8.1b Posterior PMF of p

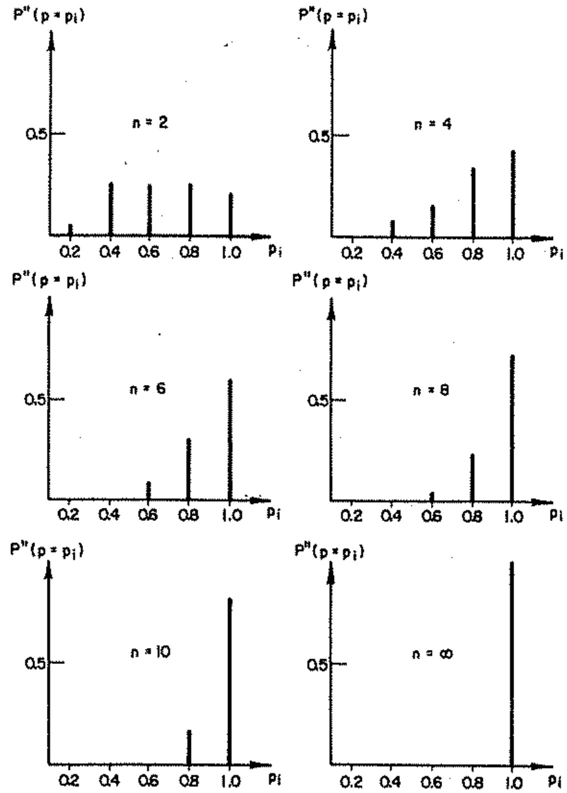

Figure E8.1c PMF of p for increasing number of test pile failures

to supersede any prior judgment. Ordinarily, however, where observational data are limited, judgment would be important and is reflected properly in the Bayesian estimation process.

Now suppose that each main column is supported on a group of three piles. If the piles carry equal loads and are statistically independent, the probability that none of the piles supporting a column will fail at a total column load of 900 tons (300 tons per pile) can be obtained by Eq. 8.3. Based on the posterior PMF of Fig. E8.1*b*, and denoting X as the number of piles failing, the required probability is

$$\begin{aligned} P(X = 0) &= P(X = 0 \mid p = 0.2)P''(p = 0.2) + P(X = 0 \mid p = 0.4)P''(p = 0.4) \\ &\quad + \cdots + P(X = 0 \mid p = 1.0)P''(p = 1.0) \\ &= (0.8)^3(0.136) + (0.6)^3(0.364) + (0.4)^3(0.204) + (0.2)^3(0.182) \\ &= 0.163 \end{aligned}$$

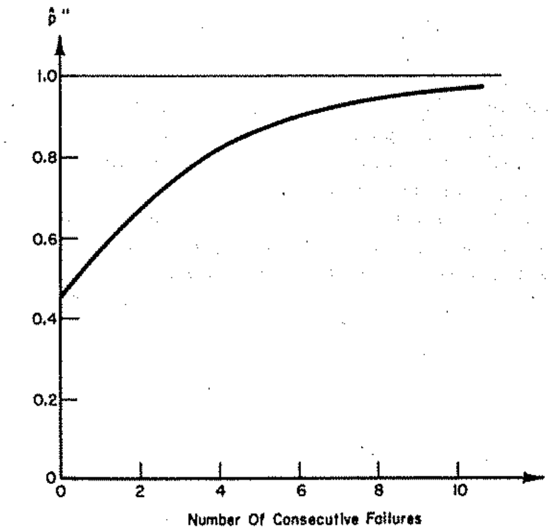

Figure E8.1d $\hat{p}''$ vs. no. of consecutive failures

EXAMPLE 8.2

A traffic engineer is interested in the average rate of accidents ν at an improved road intersection. Suppose that from his previous experience with similar road and traffic conditions, he deduced that the expected accident rate would be between one and three per year, with an average of two, and the prior PMF shown in Fig. E8.2. Occurrence of accidents is assumed to be a Poisson process.

During the first month after completion of the intersection, one accident occurred.

(a) In the light of this observation, revise the estimate for ν.

(b) Using the result of part (a), determine the probability of no accident in the next six months.

Solutions

(a) Let ε be the event that an accident occurred in one month. The posterior probabilities then are

$$\begin{aligned} P''(\nu = 1) &= \frac{P(\varepsilon \mid \nu = 1)P'(\nu = 1)}{P(\varepsilon \mid \nu = 1)P'(\nu = 1) + P(\varepsilon \mid \nu = 2)P'(\nu = 2) + P(\varepsilon \mid \nu = 3)P'(\nu = 3)} \\ &= \frac{e^{-1/12}(1/12)(0.3)}{e^{-1/12}(1/12)(0.3) + e^{-1/6}(1/6)(0.4) + e^{-1/4}(1/4)(0.3)} \\ &= 0.166 \end{aligned}$$

Similarly,

$$P''(\nu = 2) = 0.411$$
$$P''(\nu = 3) = 0.423$$

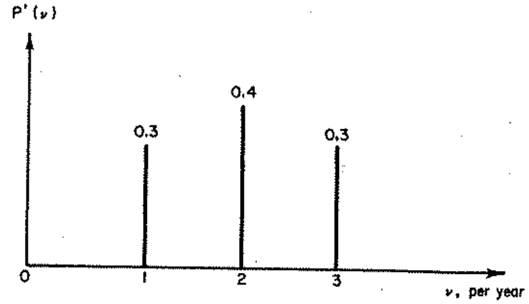

Figure E8.2 Prior distribution of ν

Hence the updated value of ν is

$$\begin{aligned} \nu'' = E(\nu \mid \varepsilon) &= (0.166)(1) + (0.411)(2) + (0.423)(3) \\ &= 2.26 \text{ accidents per year} \end{aligned}$$

(b) Let A be the event of no accidents in the next six months. Then

$$\begin{aligned} P(A) &= P(A \mid \nu = 1)P''(\nu = 1) + P(A \mid \nu = 2)P''(\nu = 2) + P(A \mid \nu = 3)P''(\nu = 3) \\ &= e^{-1/2}(0.166) + e^{-1}(0.411) + e^{-3/2}(0.423) \\ &= 0.346 \end{aligned}$$

8.3. THE CONTINUOUS CASE

8.3.1. General formulation

In Section 8.2 the possible values of the parameter θ (such as p in Example 8.1 and ν in Example 8.2) were limited to a discrete set of values; this was purposely assumed to simplify the presentation of the concepts underlying

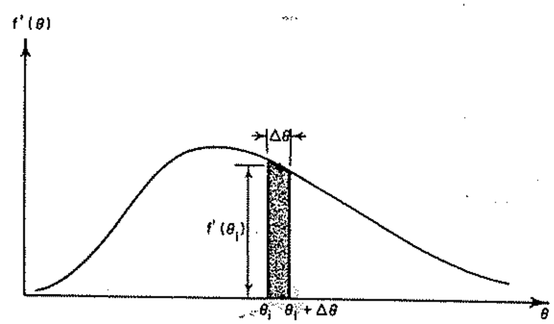

Figure 8.2 Continuous prior distribution of parameter θ

the Bayesian method of estimation. In many situations, however, the value of a parameter could be in a continuum of possible values. Thence, it would be appropriate to assume the parameter to be a continuous random variable in the Bayesian estimation. In this case we develop the corresponding results, analogous to Eqs. 8.1 through 8.3, as follows.

Let Θ be the random variable for the parameter of a distribution, with a prior density function $f'(\theta)$ shown in Fig. 8.2. The prior probability that θ will be between θ_i and $\theta_i + \Delta\theta$ then is $f'(\theta_i)\Delta\theta$. Then, if ϵ is an observed experimental outcome, the prior distribution $f'(\theta)$ can be revised in the light of ϵ using Bayes' theorem, obtaining the posterior probability that θ will be in $(\theta_i, \theta_i + \Delta\theta)$ as

$$f''(\theta_i)\Delta\theta = \frac{P(\epsilon \mid \theta_i)f'(\theta_i)\Delta\theta}{\sum_{i=1}^{n} P(\epsilon \mid \theta_i)f'(\theta_i)\Delta\theta}$$

where $P(\epsilon \mid \theta_i) = P(\epsilon \mid \theta_i < \theta \leq \theta_i + \Delta\theta)$. In the limit, this yields

$$f''(\theta) = \frac{P(\epsilon \mid \theta)f'(\theta)}{\int_{-\infty}^{\infty} P(\epsilon \mid \theta)f'(\theta)\, d\theta} \tag{8.4}$$

The term $P(\epsilon \mid \theta)$ is the conditional probability or likelihood of observing the experimental outcome ϵ assuming that the value of the parameter is θ. Hence $P(\epsilon \mid \theta)$ is a function of θ and is commonly referred to as the *likelihood function* of θ and denoted $L(\theta)$. The denominator is independent of θ; this is simply a normalizing constant required to make $f''(\theta)$ a proper density function. Equation 8.4 then can be expressed as

$$f''(\theta) = kL(\theta)f'(\theta) \tag{8.5}$$

where the normalizing constant $k = \left[\int_{-\infty}^{\infty} L(\theta)f'(\theta)\, d\theta\right]^{-1}$; and

$L(\theta)$ = the likelihood of observing the experimental outcome ϵ assuming a given θ.

We observe from Eq. 8.5 that both the prior distribution and the likelihood function contribute to the posterior distribution of Θ. In this way, as in the discrete case, the significance of judgment and of observational data is combined properly and systematically; the former through $f'(\theta)$ and the latter in $L(\theta)$.

Analogous to the discrete case, Eq. 8.2, the expected value of Θ is commonly used as the point estimator of the parameter. Hence the updated estimate of the parameter θ, in the light of observational data ϵ, is given by

$$\hat{\theta}'' = E(\Theta \mid \epsilon) = \int_{-\infty}^{\infty} \theta f''(\theta)\, d\theta \tag{8.6}$$

The uncertainty in the estimation of the parameter can be included in the calculation of the probability associated with a value of the underlying random variable. For example, if X is a random variable

$$P(X \le a) = \int_{-\infty}^{\infty} P(X \le a \mid \theta) f''(\theta)\, d\theta \tag{8.7}$$

Physically, Eq. 8.7 is the average probability of $(X \le a)$ weighted by the posterior probabilities of the parameter θ.

EXAMPLE 8.3

Consider again the problem of Example 8.1, in which the probability of pile failure at a load of 300 tons is of concern; this time, however, assume that the probability p is a continuous random variable. If there is no (prior) factual information on p, a uniform prior distribution may be assumed (known as the *diffuse prior*), namely,

$$f'(p) = 1.0 \qquad 0 \le p \le 1$$

On the basis of a single test, the likelihood function is simply the probability of the event ε = capacity of test pile less than 300 tons, which is simply p. Hence the posterior distribution of p, according to Eq. 8.5, is

$$f''(p) = kp(1.0) \qquad 0 \le p \le 1$$

in which the constant

$$k = \left[\int_0^1 p\, dp\right]^{-1} = 2$$

Thus

$$f''(p) = 2p \qquad 0 \le p \le 1$$

The Bayesian estimate of p then is

$$\hat{p}'' = E(p \mid \varepsilon) = \int_0^1 p \cdot 2p\, dp$$
$$= 0.667$$

If a sequence of n piles were tested, out of which r piles failed at loads less than the maximum test load, then the likelihood function is the probability of observing r failures among the n piles tested. If the failure probability of each pile is p, and statistical independence is assumed between piles, the likelihood function would be

$$L(p) = \binom{n}{r} p^r (1-p)^{n-r}$$

Then, with the diffuse prior, the posterior distribution of p becomes

$$f''(p) = k\binom{n}{r} p^r (1-p)^{n-r} \qquad 0 \le p \le 1$$

where

$$k = \left[\int_0^1 \binom{n}{r} p^r (1-p)^{n-r}\, dp\right]^{-1}$$

Thus the Bayesian estimator is

$$\hat{p}'' = E(p \mid \varepsilon) = \frac{\int_0^1 p \binom{n}{r} p^r (1-p)^{n-r}\, dp}{\int_0^1 \binom{n}{r} p^r (1-p)^{n-r}\, dp}$$

$$= \frac{\int_0^1 p^{r+1} (1-p)^{n-r}\, dp}{\int_0^1 p^r (1-p)^{n-r}\, dp}$$

Repeated integration-by-parts of the above integrals yields

$$\hat{p}'' = \frac{r+1}{n} \frac{\int_0^1 (p^n - p^{n+1})\, dp}{\int_0^1 (p^{n-1} - p^n)\, dp}$$

$$= \frac{r+1}{n+2}$$

From this result, we may observe that as the number of tests n increases (with the ratio r/n remaining constant), the Bayesian estimate for p approaches that of the classical estimate; that is,

$$\frac{r+1}{n+2} \to \frac{r}{n} \qquad \text{for large } n$$

EXAMPLE 8.4

An engineer is designing a temporary structure subjected to wind load on a newly developed island in the Pacific. Of interest is the probability p that the annual maximum wind speed will not exceed 120 km/hr. Records for the annual maximum wind speed in the island are available only for the last five years; and among these, the 120 km/hr wind was exceeded only once. However, an adjacent island has a longer record of wind speeds. After a comparative study of the geographical condition in the two islands, the engineer inferred from this longer record that the average value of p for the newly developed island is 2/3 with a COV of 27%. Since p is bounded between 0 and 1.0, the following beta distribution (consistent with the above statistics) is also assumed for the prior distribution:

$$f'(p) = 20p^3(1-p) \qquad 0 \le p \le 1$$

In this case, the likelihood that the annual maximum wind speed will exceed 120 kph in one out of five years is

$$L(p) = \binom{5}{4} p^4 (1-p)$$

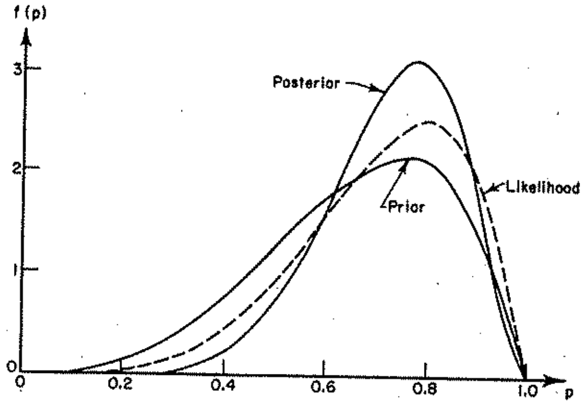

Figure E8.4 Prior, likelihood, and posterior functions

Hence the posterior density function of p is

$$f''(p) = kL(p)f'(p)$$
$$= k\left[\binom{5}{4}p^4(1-p)\right][20p^3(1-p)]$$
$$= 100kp^7(1-p)^2$$

where

$$k = \left[\int_0^1 100p^7(1-p)^2\,dp\right]^{-1} = 3.6$$

Thus

$$f''(p) = 360p^7(1-p)^2 \qquad 0 \le p \le 1$$

In this case, the prior density function is equivalent to the assumption of one exceedance in four years, whereas the resulting posterior distribution is tantamount to two exceedances in nine years. In fact, the above posterior distribution is the same as that obtained for a case in which two exceedances were observed in nine years and a diffused prior distribution is assumed. This example should serve also to illustrate a property of the Bayesian approach—namely, that information from sources other than the observed data can be useful in the estimation process.

The relation between the likelihood function and the prior and posterior distributions of the parameter p is illustrated in Fig. E8.4. Observe that the posterior distribution is "sharper" than either the prior distribution or the likelihood function. This implies that more information is "contained" in the posterior distribution than in either the prior or the likelihood function.

EXAMPLE 8.5

The occurrences of earthquakes may be modeled as a Poisson process with mean occurrence rate ν (Benjamin, 1968). Suppose that historical record for a region A shows that n_0 earthquakes have occurred in the past t_0 years. The corresponding likelihood function is then given by

$$L(\nu) = P(n_0 \text{ quakes in } t_0 \text{ years} \mid \nu)$$
$$= \frac{(\nu t_0)^{n_0}}{n_0!}e^{-\nu t_0} \qquad \nu \ge 0$$

If there is no other information for estimating ν, a uniform diffuse prior may be assumed; this implies that $f'(\nu)$ is independent of the values of ν and thus can be absorbed into the normalizing constant k. Then the posterior distribution of ν becomes

$$f''(\nu) = kL(\nu)$$
$$= k\frac{(\nu t_0)^{n_0}}{n_0!}e^{-\nu t_0} \qquad \nu \ge 0$$

Upon normalization, $k = t_0$; this result may also be obtained by comparing the foregoing $f''(\nu)$ with the gamma density function of Eq. 3.44*b* (for the random variable ν).

The probability of the event (E = n earthquakes in the next t years in region A) is then given by Eq. 8.7 as follows:

$$P(E) = \int_0^\infty P(E \mid \nu)f''(\nu)\,d\nu$$
$$= \int_0^\infty \frac{(\nu t)^n}{n!}e^{-\nu t}\cdot\frac{t_0(\nu t_0)^{n_0}}{n_0!}e^{-\nu t_0}\,d\nu$$
$$= \left(\int_0^\infty \frac{(t+t_0)[\nu(t+t_0)]^{n+n_0}}{(n+n_0)!}e^{-\nu(t+t_0)}\,d\nu\right)\frac{(n+n_0)!}{n!\,n_0!}\,\frac{t^n t_0^{(n_0+1)}}{(t+t_0)^{n+n_0+1}}$$

Since the integrand inside the parentheses is a gamma density function, the integral is equal to 1.0. Hence

$$P(E) = \frac{(n+n_0)!}{n!\,n_0!}\,\frac{t^n t_0^{(n_0+1)}}{(t+t_0)^{n+n_0+1}} = \frac{(n+n_0)!}{n!\,n_0!}\,\frac{(t/t_0)^n}{(1+t/t_0)^{n+n_0+1}}$$

a result that was first derived by Benjamin (1968).

As an illustration, suppose that historical records in region A show that two earthquakes with intensity exceeding VI (MM scale) had occurred in the last 60 years. The probability that there will be no earthquakes with this intensity in the next 20 years, therefore, is

$$P(E) = \frac{(0+2)!}{0!\,2!}\,\frac{(20/60)^0}{(1+20/60)^3}$$
$$= 0.42$$

8.3.2. A special application of Bayesian up-dating process

An interesting application of the Bayesian updating process is in the inspection and detection of material defects (Tang, 1973). Fatigue and fracture failures in metal structures are frequently the result of unchecked propagation of flaws or cracks in the joints (welds) or base metals. Periodic inspection and repair can be used to minimize the risk of fracture

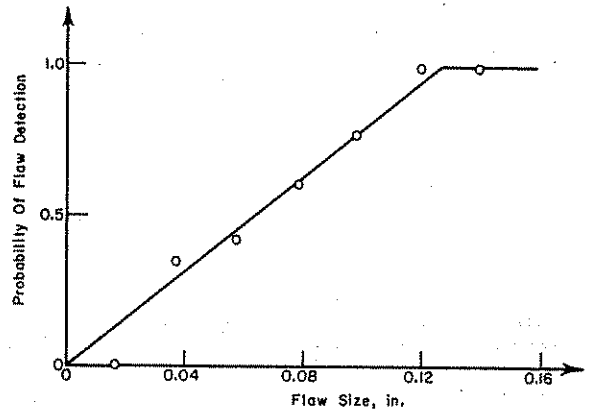

Figure 8.3 Detectability versus actual flaw depth (data from Packman et al., 1968)

failure by limiting the existing flaw sizes. Methods of detecting flaws, such as nondestructive testing (NDT), however, are invariably imperfect; consequently, not all flaws may be detected during an inspection.

The probability of detecting a flaw generally increases with the flaw size and the detection power of the device. An example of a detectability curve for ultrasonics method is shown in Fig. 8.3. Hence, even when a structure is inspected and all detected flaws are repaired, it is difficult to ensure that there are no flaws larger than some specified size.

Suppose that an NDT device is used to inspect a set of welds in a structure and all detected flaws are fully repaired. On the basis of this assumption, the flaws that remain in the weld would be those that were not detected. Let X be the flaw size and D the event that a flaw is detected. The probability that a flaw size (for example, depth) will be between x and $(x + dx)$ given that the flaw was not detected is, therefore,

$$P(x < X \leq x + dx \mid \bar{D}) = \frac{P(\bar{D} \mid x < X \leq x + dx)P(x < X \leq x + dx)}{P(\bar{D})}$$

This can be expressed also in terms of density functions as

$$f_X(x \mid \bar{D}) = kP(\bar{D} \mid x) f_X(x) \tag{8.8}$$

in which $f_X(x)$ is the distribution of the flaw size prior to inspection and repair, whereas $f_X(x \mid \bar{D})$ is the corresponding dist-ᵗʰution after inspection and repair. Also $P(\bar{D} \mid x) = 1 - P(D \mid x)$, whe '$(D \mid x)$ is simply the

probability of detecting a flaw with depth x, which is the function defined by the detectability curve, such as that shown in Fig. 8.3. Comparing Eq. 8.8 with Eq. 8.5, we observe that Eq. 8.8 is of the same form as Eq. 8.5, with the following equivalences:

$$f_X(x \mid \bar{D}) \sim \text{the posterior distribution}$$

$$P(\bar{D} \mid x) \sim \text{the likelihood function}$$

$$f_X(x) \sim \text{the prior distribution}$$

EXAMPLE 8.6

As an illustration, suppose the initial (prior) distribution of flaw depths X in a series of welds has a triangular shape described as follows (see Fig. E8.6):

$$f_X(x) = \begin{cases} 208.3x & 0 < x \leq 0.06 \\ 20 - 125x & 0.06 < x \leq 0.16 \\ 0 & x > 0.16 \end{cases}$$

Assume also that the NDT device used in the inspection has the detectability curve shown in Fig. 8.3; mathematically, this curve is given by

$$P(D \mid x) = \begin{cases} 0 & x \leq 0 \\ 8x & 0 < x \leq 0.125 \\ 1.0 & x > 0.125 \end{cases}$$

Substituting the appropriate expressions for each interval of X into Eq. 8.8, we

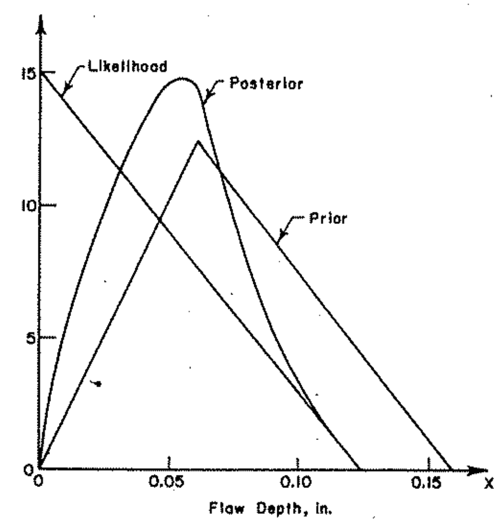

Figure E8.6 Distribution of flaw depth obtain the updated density function of flaw depths

$$f_X(x \mid D) = \begin{cases} 0 & x \le 0 \\ k(1 - 8x)(208.3x) & 0 < x \le 0.06 \\ k(1 - 8x)(20 - 125x) & 0.06 < x \le 0.125 \\ 0 & x > 0.125 \end{cases}$$

which, after normalization, becomes

$$f_X(x \mid D) = \begin{cases} 0 & x \le 0 \\ 495x - 3964x^2 & 0 < x \le 0.06 \\ 47.6 - 678x + 2379x^2 & 0.06 < x \le 0.125 \\ 0 & x > 0.125 \end{cases}$$

The above prior, likelihood, and posterior functions are plotted in Fig. E8.6. It can be observed that the likelihood function, which is the "complementary function" of Fig. 8.3, behaves as a filter; it cuts off flaws larger than 0.125 in. and also eliminates many of the remaining larger flaws. Thus, after the inspection and repair program, the distribution of flaw depth is shifted toward smaller values.

8.4. BAYESIAN CONCEPTS IN SAMPLING THEORY

8.4.1. General formulation

If the experimental outcome ϵ in Eq. 8.4 is a set of observed values $x_1, x_2, \ldots, x_n$, representing a random sample (see Section 5.2.1) from a population X with underlying density function $f_X(x)$, the probability of observing this particular set of values, assuming that the parameter of the distribution is θ, is

$$P(\epsilon \mid \theta) = \prod_{i=1}^{n} f_X(x_i \mid \theta)\, dx$$

Then, if the prior density function of Θ is $f'(\theta)$, the corresponding posterior density function becomes, according to Eq. 8.4,

$$f''(\theta) = \frac{\left[\prod_{i=1}^{n} f_X(x_i \mid \theta)\, dx\right] f'(\theta)}{\int_{-\infty}^{\infty} \left[\prod_{i=1}^{n} f_X(x_i \mid \theta)\, dx\right] f'(\theta)\, d\theta}$$

$$= kL(\theta) f'(\theta) \tag{8.9}$$

in which the normalizing constant is

$$k = \left[\int_{-\infty}^{\infty} \left(\prod_{i=1}^{n} f_X(x_i \mid \theta)\right) f'(\theta)\, d\theta\right]^{-1}$$

whereas the likelihood function $L(\theta)$ is the product of the density function

of X evaluated at $x_1, x_2, \ldots, x_n$, or

$$L(\theta) = \prod_{i=1}^{n} f_X(x_i \mid \theta) \tag{8.10}$$

Using the posterior density function for Θ of Eq. 8.9 in Eq. 8.6, we therefore obtain the Bayesian estimator of the parameter θ. It is interesting to observe that the likelihood function of Eq. 8.10 is the same as that given earlier in Eq. 5.4 in connection with the classical method of maximum likelihood estimation. Furthermore, if a diffuse prior distribution is assumed (for example, as in Eq. 8.13), then the mode of the posterior distribution, Eq. 8.9, gives the maximum likelihood estimator.

8.4.2. Sampling from normal population

In the case of a Gaussian population with known standard deviation σ, the likelihood function for the parameter μ, according to Eq. 8.10, is

$$L(\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2\right] = \prod_{i=1}^{n} N_\mu(x_i, \sigma)$$

where $N_\mu(x_i, \sigma)$ denotes the density function of μ with mean value x_i and standard deviation σ. It can be shown (for instance, Tang, 1971) that the product of m normal density functions with respective means μ_i and standard deviations σ_i is also a normal density function with mean and variance

$$\mu^* = \frac{\sum_{i=1}^{m} (\mu_i/\sigma_i^2)}{\sum_{i=1}^{m} 1/\sigma_i^2} \quad \text{and} \quad (\sigma^*)^2 = \frac{1}{\sum_{i=1}^{m} 1/\sigma_i^2} \tag{8.11}$$

Therefore the likelihood function $L(\mu)$ becomes

$$L(\mu) = N_\mu\left(\frac{\sum_{i=1}^{n} (x_i/\sigma^2)}{\sum_{i=1}^{n} (1/\sigma^2)}, \frac{1}{\sqrt{\sum_{i=1}^{n} (1/\sigma^2)}}\right) = N_\mu\left(\frac{(1/\sigma^2)\sum_{i=1}^{n} x_i}{n/\sigma^2}, \frac{1}{\sqrt{n/\sigma^2}}\right)$$

$$= N_\mu\left(\bar{x}, \frac{\sigma}{\sqrt{n}}\right) \tag{8.12}$$

where $\bar{x}$ is the sample mean of Eq. 5.1.

Without prior information. In the absence of prior information on μ, a diffuse prior distribution may be assumed. In such a case we obtain the posterior distribution for μ, as

$$\begin{aligned} f''(\mu) &= kL(\mu) \\ &= kN_\mu\left(\bar{x}, \frac{\sigma}{\sqrt{n}}\right) \\ &= N_\mu\left(\bar{x}, \frac{\sigma}{\sqrt{n}}\right) \end{aligned} \tag{8.13}$$

where k is necessarily equal to 1.0 upon normalization. Therefore, without prior information, the posterior distribution of μ is Gaussian with a mean value equal to the sample mean $\bar{x}$ and standard deviation $\sigma/\sqrt{n}$.

Using the expected value of μ as the Bayesian estimator we obtain, in accordance with Eq. 8.6,

$$\hat{\mu}'' = E(\mu \mid \epsilon) = \bar{x}$$

That is, the sample mean $\bar{x}$ is the point estimate of the population mean. We recognize that this is the same as the classical estimate of Eq. 5.1. Therefore, in the absence of prior information, the Bayesian and classical methods give the same estimates for the population mean. Conceptually, however, the Bayesian basis for this estimate differs from that of the classical approach. Whereas Eq. 8.13 says that the posterior distribution of μ is Gaussian with mean $\bar{x}$ and standard deviation $\sigma/\sqrt{n}$, the classical approach (of Sect. 5.2) says the sample mean $\bar{X}$ is a Gaussian random variable with mean μ and standard deviation $\sigma/\sqrt{n}$.

Significance of prior information. In contrast to the classical approach, however, prior information can be included in the estimation of the parameter μ. This is accomplished explicitly through the prior distribution $f'(\theta)$; we demonstrate this for the case of a Gaussian population as follows.

In the case where X is Gaussian with known variance, it is mathematically convenient to assume also a Gaussian prior (see Sect. 8.4.4). Suppose that $f'(\mu)$ is $N(\mu', \sigma')$. Then, with the likelihood function of Eq. 8.12, the posterior distribution of $\bar{\mu}$ becomes

$$\begin{aligned} f''(\mu) &= kL(\mu)f'(\mu) \\ &= kN_\mu\left(\bar{x}, \frac{\sigma}{\sqrt{n}}\right)N_\mu(\mu', \sigma') \end{aligned}$$

which is a product of two normal density functions. Again, it can be shown that $f''(\mu)$ is also Gaussian with mean

$$\mu'' = \frac{[\bar{x}/(\sigma/\sqrt{n})^2] + [\mu'/(\sigma')^2]}{[1/(\sigma/\sqrt{n})^2] + [1/(\sigma')^2]} = \frac{\bar{x}(\sigma')^2 + \mu'(\sigma^2/n)}{(\sigma')^2 + (\sigma^2/n)} \tag{8.14}$$

and standard deviation

$$\sigma'' = \sqrt{\frac{(\sigma')^2(\sigma^2/n)}{(\sigma')^2 + (\sigma^2/n)}} \tag{8.15}$$

In this case the Bayesian estimator of μ, Eq. 8.6, yields

$$\hat{\mu}'' = \mu''$$

That is, the Bayesian estimate of the mean value is an average of the prior mean μ' and the sample mean $\bar{x}$, weighted inversely by the respective variances.

Equation 8.14 is an example of how prior information is combined systematically with observed data—in the present case, to estimate the mean value μ.

It is important to observe that the posterior variance of μ, as given by Eq. 8.15, is always less than* $(\sigma')^2$ or (σ^2/n); that is, the variance of the posterior distribution is always less than that of the prior distribution or of the likelihood function.

On the basis of the posterior distribution of μ, that is, $N_\mu(\bar{x}, \sigma/\sqrt{n})$ of Eq. 8.13 or $N_\mu(\mu'', \sigma'')$ with Eqs. 8.14 and 8.15, we may also determine the interval for μ corresponding to a specified probability. For example, the probability that μ is between a and b is given by

$$P(a < \mu \leq b) = \int_a^b f''(\mu)\, d\mu$$

8.4.3. Error in estimation

Any error in the estimation of a parameter θ can be combined with the inherent variability of the underlying random variable, for example X, to obtain the total uncertainty associated with X. Accounting for the error in the estimation of θ, the density function of X becomes (by virtue of the

* Since $(\sigma')^2 \geq 0$, and $\sigma^2/n \geq 0$

$$(\sigma'^2)^2 + (\sigma'^2)\left(\frac{\sigma^2}{n}\right) \geq (\sigma'^2)\left(\frac{\sigma^2}{n}\right)$$

$$(\sigma')^2\left(\sigma'^2 + \frac{\sigma^2}{n}\right) \geq (\sigma'^2)\left(\frac{\sigma^2}{n}\right)$$

or

$$(\sigma')^2 \geq \frac{(\sigma'^2)(\sigma^2/n)}{(\sigma')^2 + \sigma^2/n} = \sigma''^2$$

Similarly, it can be shown that $(\sigma'')^2 \leq \sigma^2/n$.

total probability theorem)

$$f_X(x) = \int_{-\infty}^{\infty} f_X(x \mid \theta) f''(\theta)\, d\theta \tag{8.16}$$

In the case of a Gaussian variate X, with known σ, and μ estimated from sample data,

$$f_X(x) = \int_{-\infty}^{\infty} f_X(x \mid \mu) f''(\mu)\, d\mu$$

where $f_X(x \mid \mu) = N_X(\mu, \sigma)$, and $f''(\mu)$ is given by Eq. 8.13. Again it can be shown (for instance, Tang, 1971) that this last integral yields the normal density function $N_X(\bar{x}, \sqrt{\sigma^2 + \sigma^2/n})$; that is,

$$f_X(x) = N(\bar{x}, \sqrt{\sigma^2 + \sigma^2/n}) \tag{8.17}$$

The overall uncertainty in X here is reflected in its variance, $\sigma^2 + \sigma^2/n$, which is composed of the variance of the basic random variable X and that of the parameter μ. Effectively, the error in the estimation of μ serves to increase the total uncertainty in X, by an amount that decreases with the sample size n.

EXAMPLE 8.7

A toll bridge was recently opened to traffic. For the past two weeks, records on rush-hour traffic during the last 10 workdays showed a sample mean of 1535 vehicles per hour (vph). Suppose that rush-hour traffic has a normal distribution with a standard deviation of 164 vph. Based on this observational information, the posterior distribution of the mean rush-hour traffic μ is, according to Eq. 8.13, $N(1535, 164/\sqrt{10})$ or $N(1535, 51.9)$ vph. The point estimate of μ, therefore, is 1535 vph.

The probability that μ will be between 1500 and 1600 vph is given by

$$\begin{aligned} P(1500 < \mu \le 1600) &= \Phi\left(\frac{1600 - 1535}{51.9}\right) - \Phi\left(\frac{1500 - 1535}{51.9}\right) \\ &= \Phi(1.253) - \Phi(-0.674) \\ &= 0.6445 \end{aligned}$$

Of greater interest are probabilities associated with the rush-hour traffic (rather than its mean) on a given workday. Suppose that for the present toll collection procedure, serious problems would arise if the rush-hour traffic exceeds 1700 vph on a given day. Then the probability that this will occur on any given day, based on Eq. 8.17, is given by

$$\begin{aligned} P(X > 1700) &= 1 - \Phi\left(\frac{1700 - 1535}{\sqrt{(164)^2 + (51.9)^2}}\right) \\ &= 1 - \Phi(0.958) \\ &= 0.169 \end{aligned}$$

In other words, in about 17% of the working days, the present toll collection system

will be inadequate during rush hours. Observe that the error in the estimation of μ has been included in computing this probability.

Now suppose that before the toll bridge was opened for traffic, simulation was performed to predict the rush-hour traffic on the bridge. Based on the simulation results alone, it was estimated that the mean rush-hour traffic on a workday would be 1500 ± 100 with 90% confidence. How can this information be used with the observed traffic flow in the estimation of μ?

Assuming a Gaussian prior and with the foregoing simulation results, we obtain the prior distribution of the mean rush-hour traffic μ to be $N(1500, 60.8)$ vph. Then, applying Eqs. 8.14 and 8.15, the posterior distribution of μ is Gaussian with

$$\mu'' = \frac{1535(60.8)^2 + 1500(51.9)^2}{(60.8)^2 + (51.9)^2} = 1520 \text{ vph}$$

and

$$\sigma'' = \sqrt{\frac{(60.8)^2(51.9)^2}{(60.8)^2 + (51.9)^2}} = 39.5 \text{ vph}$$

Therefore, by incorporating the result of simulation, the estimated mean rush-hour traffic is 1520 vph and corresponding standard deviation is 39.5 vph.

EXAMPLE 8.8

Five repeated measurements of the elevation (relative to a fixed datum) of a bridge pier under construction were made as follows:

20.45 m
20.38 m
20.51 m
20.42 m
20.46 m

Assume that the measurement error is Gaussian with zero mean and standard deviation 0.08 m.

(a) Estimate the actual elevation of the pier based on the given measurements.

(b) Suppose that the elevation of the pier was previously measured by another surveying crew; the elevation was estimated to be 20.42 ± 0.02 m (that is, the mean measurement was 20.42 m with a standard error of 0.02 m). Estimate the elevation of the pier taking advantage of this prior information.

Solution

The estimation of an actual dimension δ in surveying and photogrammetry is equivalent to the estimation of the mean value of a random variable (see Section 5.2.3). Measurement error is invariably assumed to be Gaussian with zero mean; this means tacitly that a set of measurements constitute a sample from a normal population. Therefore the results derived in Section 8.4.2 are applicable to the estimation of geometric quantities in surveying and photogrammetry.

(a) The sample mean of the five measurements is

$$\begin{aligned} \bar{d} &= \tfrac{1}{5}(20.45 + 20.38 + 20.51 + 20.42 + 20.46) \\ &= 20.444 \text{ m} \end{aligned}$$

Hence, on the basis of the five observations, the actual elevation of the pier has a

Gaussian distribution $N(20.444, 0.08/\sqrt{5})$ or $N(20.444, 0.036)$ m. In the convention of surveying and photogrammetry, the elevation of the pier would be given as 20.444 ± 0.036 m.

(b) In the case where prior information is available, such information can be incorporated through the prior distribution of δ. In the present case, using the pier elevation estimated earlier by another crew, the prior distribution of δ can be modeled as $N(20.420, 0.020)$ m. Then applying Eqs. 8.14 and 8.15, the Bayesian estimate of the elevation is

$$\hat{d}'' = \frac{(20.420)(0.036)^2 + (20.444)(0.020)^2}{(0.036)^2 + (0.020)^2}$$
$$= 20.426 \text{ m}$$

and the corresponding standard error is

$$\sigma_d'' = \sqrt{\frac{(0.036)^2(0.020)^2}{(0.036)^2 + (0.020)^2}}$$
$$= 0.017 \text{ m}$$

EXAMPLE 8.9

The annual maximum flow of a stream has been recorded for the last five years as follows:

21.5, 19.2, 23.4, 20.1, 18.1 (100 m³/sec)

Based on extensive data from adjacent streams, the annual maximum stream flow may be modeled by a log-normal distribution. Assume that the parameter ζ in the log-normal distribution is equal to the value obtained from the five sample values. The problem here is to estimate the parameter λ.

In Chapter 4 (Example 4.2) it is shown that if a random variable Y is log-normal, then $X = \ln Y$ is normal. Hence the logarithm of the stream flow will be Gaussian with mean λ and known standard deviation ζ.

The natural logarithm of the above data values are, respectively,

3.07, 2.96, 3.15, 3.00, 2.90

from which we obtain the sample mean $\bar{x} = 3.016$, and sample standard deviation $\zeta = 0.097$.

Without any prior information, the posterior distribution of λ, according to Eq. 8.13, is $N(\bar{x}, \zeta/\sqrt{5})$ or $N(3.016, 0.097/\sqrt{5}) = N(3.016, 0.043)$.

If prior information is available, it can be incorporated through the prior distribution of λ. For example, suppose that $f'(\lambda)$ is assumed to be $N(2.9, 0.06)$; then from Eqs. 8.14 and 8.15 the posterior distribution $f''(\lambda)$ will be normal with

$$\mu_\lambda'' = \frac{3.016(0.06)^2 + 2.9(0.0435)^2}{(0.06)^2 + (0.0435)^2} = 2.98$$

and

$$\sigma_\lambda'' = \sqrt{\frac{(0.06)^2(0.0435)^2}{(0.06)^2 + (0.0435)^2}} = 0.035$$

That is, in this latter case, the posterior distribution of λ is $N(2.98, 0.035)$.

8.4.4. Use of conjugate distributions

In deriving the posterior distribution of a parameter by Eq. 8.5 or 8.9, considerable mathematical simplification can be achieved if the distribution of the parameter is appropriately chosen with respect to that of the underlying random variable. We saw this in Sect. 8.4.2 in the case of the Gaussian random variable X with known σ; by assuming the prior distribution of μ to be also Gaussian, the posterior distribution of μ remains Gaussian. This was similarly demonstrated for the discrete case in Example 8.4, in which the random variable has a binomial distribution and the prior distribution for p was assumed to be a beta distribution (with parameters $q' = 4$ and $r' = 2$). The resulting posterior distribution for p is also a beta distribution, with updated parameters $q'' = 8$ and $r'' = 3$.

Such pairs of distributions are known in the Bayesian terminology as *conjugate pairs* or *conjugate distributions*. By choosing a prior distribution that is a conjugate of the distribution of the underlying random variable, convenient posterior distribution, which is usually of the same mathematical form as the prior, is obtained. This has been illustrated earlier in the case of the normal-normal and the binomial-beta distributions. Other pairs of conjugate distributions may be developed; Table 8.1 summarizes some of these involving certain common distributions.

It should be emphasized that conjugate distributions are chosen solely for mathematical convenience and simplicity. For a random variable with a specified distribution, its conjugate prior distribution may be adopted if there is no other basis for the choice of the prior distribution. However, if there is evidence to support a particular prior distribution, then such a distribution ought to be used, mathematical complications notwithstanding.

EXAMPLE 8.10

The occurrence of flaws in a weld joint may be modeled by a Poisson process with a mean occurrence rate of μ flaws per meter of weld. Actual observation with a powerful device (assume it would not miss detecting any significant flaw) detected 5 flaws in a weld of 9.2 meters. However, from previous experience with the same type of weld and quality of workmanship, the mean flaw rate is believed to be 0.5 flaw/m with a COV of 40%. Determine the mean and COV of μ for this type of weld, using the observed data as well as the information from prior experience.

Since the number of flaws in a given weld length is described by the Poisson distribution, it is convenient, according to Table 8.1, to prescribe its conjugate gamma distribution as the prior distribution for the parameter μ. From the information given above, and observing from Section 3.2.8 the mean and variance of the gamma distribution, we have

$$E'(\mu) = \frac{k'}{\nu'} = 0.5$$

Table 8.1 Conjugate Distributions

Basic random variable	Parameter	Prior and posterior distributions of parameter	Mean and Variance of Parameter	Posterior Statistics
Binomial $p_X(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$	θ	Beta $f_\Theta(\theta) = \frac{\Gamma(q+r)}{\Gamma(q)\Gamma(r)}\theta^{q-1}(1-\theta)^{r-1}$	$E(\Theta) = \frac{q}{q+r}$ $\mathrm{Var}(\Theta) = \frac{qr}{(q+r)^2(q+r+1)}$	$q'' = q' + x$ $r'' = r' + n - x$
Exponential $f_X(x) = \lambda e^{-\lambda x}$	λ	Gamma $f_\Lambda(\lambda) = \frac{\nu(\nu\lambda)^{k-1}e^{-\nu\lambda}}{\Gamma(k)}$	$E(\lambda) = \frac{k}{\nu}$ $\mathrm{Var}(\lambda) = \frac{k}{\nu^2}$	$\nu'' = \nu' + \sum_i x_i$ $k'' = k' + n$
Normal $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$ (with known σ)	μ	Normal $f_M(\mu) = \frac{1}{\sqrt{2\pi}\sigma_\mu}\exp\left[-\frac{1}{2}\left(\frac{\mu-\mu_\mu}{\sigma_\mu}\right)^2\right]$	$E(\mu) = \mu_\mu$ $\mathrm{Var}(\mu) = \sigma_\mu^2$	$\mu_\mu'' = \frac{\mu_\mu'(\sigma^2/n) + \bar{x}\sigma_\mu'^2}{\sigma^2/n + (\sigma_\mu')^2}$ $\sigma_\mu'' = \sqrt{\frac{(\sigma_\mu')^2(\sigma^2/n)}{(\sigma_\mu')^2 + \sigma^2/n}}$
Normal $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$	μ, σ	Gamma-Normal $f(\mu, \sigma)$ $= \left\{\frac{1}{\sqrt{2\pi}\sigma/n}\exp\left[-\frac{1}{2}\left(\frac{\mu-\bar{x}}{\sigma/\sqrt{n}}\right)^2\right]\right\}$ $\cdot\left\{\frac{[(n-1)/2]^{(n+1)/2}}{\Gamma[(n+1)/2]}\left(\frac{s^2}{\sigma^2}\right)^{(n-1)/2}\right.$ $\left.\cdot\exp\left(-\frac{n-1}{2}\frac{s^2}{\sigma^2}\right)\right\}$	$E(\mu) = \bar{x}$ $\mathrm{Var}(\mu) = s^2\left[\frac{n-1}{n(n-3)}\right]$ $E(\sigma) = s\sqrt{\frac{n-1}{2}}\frac{\Gamma[(n-2)/2]}{\Gamma[(n-1)/2]}$ $\mathrm{Var}(\sigma) = s^2\left(\frac{n-1}{n-3}\right) - E^2(\sigma)$	$n'' = n' + n$ $n''\bar{x}'' = n'\bar{x}' + n\bar{x}$ $(n''-1)s''^2 + n''\bar{x}''^2$ $= [(n'-1)s'^2 + n'\bar{x}'^2]$ $+ [(n-1)s^2 + n\bar{x}^2]$
Poisson $p_X(x) = \frac{(\mu t)^x}{x!}e^{-\mu t}$	μ	Gamma $f_M(\mu) = \frac{\nu(\nu\mu)^{k-1}e^{-\nu\mu}}{\Gamma(k)}$	$E(\mu) = \frac{k}{\nu}$ $\mathrm{Var}(\mu) = \frac{k}{\nu^2}$	$\nu'' = \nu' + t$ $k'' = k' + x$
Lognormal $f_X(x) = \frac{1}{\sqrt{2\pi}\zeta x}\cdot$ $\cdot\exp\left[-\frac{1}{2}\left(\frac{\ln x - \lambda}{\zeta}\right)^2\right]$ (with known ζ)	λ	Normal $f_\Lambda(\lambda) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{\lambda-\mu}{\sigma}\right)^2\right]$	$E(\lambda) = \mu$ $\mathrm{Var}(\lambda) = \sigma^2$	$\mu'' = \frac{\mu'(\zeta^2/n) + \sigma^2\overline{\ln x}}{\zeta^2/n + \sigma^2}$ $\sigma'' = \sqrt{\frac{\sigma^2(\zeta^2/n)}{\sigma^2 + \zeta^2/n}}$

and

$$\delta'(\mu) = \frac{\sqrt{k'/\nu'^2}}{k'/\nu'} = \frac{1}{\sqrt{k'}} = 0.4$$

Thus the prior parameters of the gamma distribution are $k' = 6.25$; and $\nu' = 12.5$.

It follows then that the posterior distribution of μ is also gamma. From the relationships given in Table 8.1 between the prior and posterior statistics, and the sample data, we evaluate the parameters k'' and ν'' of the posterior gamma distribution as follows:

$$k'' = k' + x = 6.25 + 5 = 11.25$$
$$\nu'' = \nu' + t = 12.5 + 9.2 = 21.7$$

Hence the updated mean and COV of the average flaw rate μ are

$$E''(\mu) = \frac{k''}{\nu''} = \frac{11.25}{21.7} = 0.52 \text{ flaw/m}$$

$$\delta''(\mu) = \frac{1}{\sqrt{k''}} = \frac{1}{\sqrt{11.25}} = 0.30$$

8.5. CONCLUDING REMARKS

In the process of engineering planning and design, judgmental assumptions and inferential information are often useful and necessary. The significance of such prior information and its role (in combination with observational data) in the process of estimation are formally the subject of Bayesian statistics. The basic concepts of the Bayesian approach have been introduced here with special reference to sampling and estimation. Applications of these concepts in Bayesian statistical decision will be covered in Vol. II.

Philosophically, there are fundamental differences between the Bayesian and classical statistics. Within the Bayesian context, a probability or a probability statement is an expression of the *degree-of-belief*, whereas in the classical sense, probability is a verifiable measure of *relative frequency*. Furthermore, in estimation, the Bayesian approach assumes that a parameter is a random variable, whereas in the classical approach it is an unknown constant.

Relative to engineering planning and design, the Bayesian approach offers the following advantages:

1. It provides the formal framework for incorporating engineering judgment (expressed in probability terms) with observational data.
2. It systematically combines uncertainties associated with randomness and those arising from errors of estimation and prediction (see Vol. II).
3. It provides a formal procedure for systematic updating of information.

PROBLEMS

8.1 A new structure is subjected to proof testing. Assume that the maximum proof load is specified at a reasonably high level so that the calculated probability of the structure surviving the maximum proof load is 0.90. However, it is felt that this calculation is only 70% reliable, and there is a 25% chance that the true probability may be 0.50; moreover, there is even a 5% chance that it may be only 0.10.

(a) What is the expected probability of survival before the proof test?
(b) If only one structure is proof-tested, and it survives the maximum proof load, determine the updated distribution of the survival probability.
(c) What is the expected probability of survival after the proof test?
(d) If three structures were proof-tested, and two of the structures survived whereas one failed under the maximum proof load, determine the updated expected probability of survival.

8.2 A new waste-treatment process has been developed. In order to evaluate its effectiveness, the treatment process is installed for a trial period. Each day the output from the treatment process is inspected to see if it satisfies the specified standard. Suppose that the outputs between days are statistically independent, and there is a probability p that the daily output will be acceptable. If the prior PMF is as shown in Fig. P8.2, determine the posterior distribution of p with each of the following observations.

(a) The output on the first day of the trial period is of unacceptable quality.
(b) For a three-day trial period, the quality is unacceptable in only one day.
(c) For a three-day trial period, the first two days are satisfactory whereas the quality is unacceptable on the third day.

In each case, determine also the Bayesian estimate for p. *Ans. 0.536, 0.617, 0.617.*

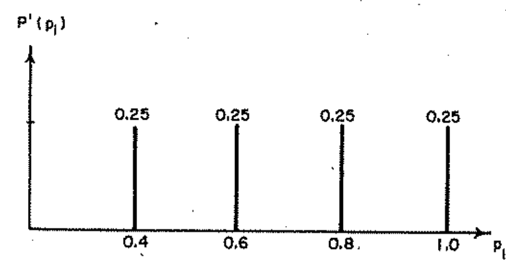

Figure P8.2

8.3 A hazardous street intersection has been improved by changing the geometric design to reduce the accident and fatality rates. For simplicity, assume that accident and fatality rates can be classified as high H or low L, leading to the following possible conditions: $H_A H_F$ (high accident rate, high fatality rate), $H_A L_F$, $L_A H_F$, and $L_A L_F$. Preliminary evaluation revealed that the relative likelihoods for these four conditions are 3:3:2:2.

An accident rate prediction model, for example, the Tharp model, was used to obtain a better evaluation of the accident potential at this (improved) intersection. Because of possible inaccuracies in the prediction model, a predicted condition may not be actually realized. Furthermore, the probability of a correct prediction depends on the underlying actual condition, as indicated in the following table of conditional probabilities.

Predicted \ Actual	H_AH_F	H_AL_F	L_AH_F	L_AL_F
$H_A'\ H_F'$	0.30	0.40	0.20	0.25
$H_A'\ L_F'$	0.30	0.30	0.20	0.25
$L_A'\ H_F'$	0.20	0.20	0.50	0.25
$L_A'\ L_F'$	0.20	0.10	0.10	0.25

(a) What is the probability that the model will indicate $H_A'H_F'$?
(b) Suppose that the model predicted $H_A'H_F'$; what is the probability that the condition of the improved intersection will actually be H_AH_F?
(c) If the model predicted $L_A'L_F'$; what is the updated relative likelihoods of the four possible conditions?

8.4 An instrument is used to check the accuracy of a set of measurements. However, it can only record three readings, namely $x = 1, 2,$ or 3. The reading $x = 2$ implies that the previously measured value is within a tolerable error, whereas $x = 1$ and $x = 3$ denote that the measurement is on the low and high side, respectively. Suppose the distribution of X is given as follows:

$$p_X(x_i) = \begin{cases} \dfrac{1-m}{2} & x_i = 1 \\ m & x_i = 2 \\ \dfrac{1-m}{2} & x_i = 3 \end{cases}$$

where m is the parameter. For a particular set of measurements, the engineer estimated that the value of m would be 0.4 or 0.8 with equal likelihood. However, on checking a set of measurements, the first one indicates $x = 2$.
(a) What should be the engineer's revised distribution of m?
(b) Estimate the probability that at least two out of the next three measurements will be accurate.

8.5 An engineer plans to build a log cabin in the middle of a forest where logs of similar size are available. He assumes that the bending capacity M of each log follows a Rayleigh distribution

$$f_M(m) = \frac{m}{\lambda^2} e^{-(1/2)(m/\lambda)^2} \qquad m \geq 0$$

where the parameter λ is the modal value of the distribution.
From previous experience with similar logs, he feels that λ would be 4(kip-ft) with probability 0.4 or 5 (kip-ft) with probability 0.6. Not entirely satisfied with these subjective probabilities, he decided to get a better measure of the parameter λ. Being pressed for time and with limited supply of logs, he can only afford to test the bending capacity of two logs by simple load test on the site. The test results yielded 4.5 kip-ft and 5.2 kip-ft for the two tests.

(a) Determine the posterior distribution (discrete) for the parameter λ.
(b) Derive the distribution of the bending capacity of the logs M, based on the posterior distribution of λ.
(c) What is the probability that M is less than 2 kip-ft?

8.6 The absolute error E (in cm) of each measurement from a surveying instrument is governed by the triangular distribution shown in Fig. P8.6, where α denotes the upper limit of the error.

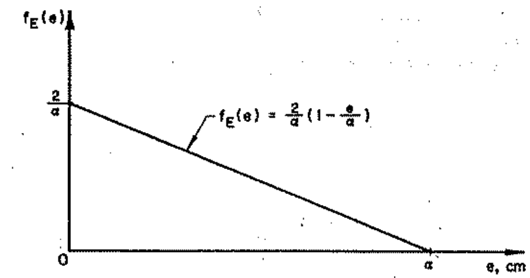

Figure P8.6

Two measurements were made and the errors are 1 and 2 cm, respectively.
(a) Suppose that α is assumed to be 2 or 3 cm with equal likelihood prior to the two measurements; determine the updated distribution of α. Estimate the value of α based on this updated distribution.
(b) Now suppose that the prior density function of α is uniform between 2 and 3; determine and plot the updated distribution of α, and evaluate the corresponding Bayesian estimate for α.

8.7 Suppose that the prior density function of the mean accident rate ν in Example 8.2 is

$$\begin{aligned} f'(\nu) &= \frac{0.271}{\nu} && 0.5 \leq \nu < 20 \\ &= 0 && \text{elsewhere} \end{aligned}$$

Determine the posterior density function of ν based on the observation that an accident was recorded during the first month of operation.

8.8 In Problem 8.1 suppose that the survival probability has a prior density function as follows.
(i) Uniform between $p = 0$ to 0.9.
(ii) Uniform between $p = 0.9$ to 1.0.
(iii) It is more likely that p will exceed 0.9 than be less than 0.9; the relative likelihood between these two possibilities is 7 to 3.
(a) Determine the prior density function of p.
(b) If three structures were proof tested, and all three survived the maximum proof load, determine and plot the posterior density function of p.
(c) What is the estimated value of p in light of the results in part (b)?

8.9 The occurrence of fire in a city may be modeled by a Poisson process. Suppose the average occurrences of fires ν is assumed to be 15 or 20 times a year; the likelihood of 20 fires a year (on the average) is twice that of 15 fires a year.

(a) Determine the probability that there will be 20 occurrences of fire in the next year.

(b) If there are actually 20 fires in the next year, what will be the updated PMF of ν?

(c) What is the probability that there will be 20 occurrences of fire in the year after next, in the light of part (b)?

8.10 Consider a case where the mean compression index of a soil stratum is to be estimated. Assume that the compression index of a soil sample is $N(\mu, \sigma)$ and σ is assumed to be equal to 0.16. Laboratory tests on four speciments show the following compression index values: 0.75, 0.89, 0.91, and 0.81.

(a) What is the posterior distribution of μ if there is no other information except the observed data?

(b) Suppose there is prior information to indicate that μ is Gaussian with mean 0.8 and COV 25%. What will be the posterior distribution of μ if this prior information is taken into account?

(c) What is the probability that μ will be less than 0.95, using the data from part (b)? *Ans. 0.938.*

8.11 An air passenger is commuting between San Francisco and Los Angeles regularly. Lately, he started recording the time of each flight. He computed the average flight time from his five previous trips to be 65 minutes. Suppose the flight time T is a Gaussian random variable with known standard deviation of 10 minutes.

(a) Based purely on the data, what is the posterior distribution of μ_T?

(b) The passenger is now on a plane from San Francisco to Los Angeles. By coincidence, the passenger sitting next to him also has been keeping track of the flight time. From his record of 10 previous trips, he obtained an average of 60 minutes. Assume that these two passengers have never taken a plane together before. With this additional information, what would be the updated distribution of μ_T?

(c) What is the probability that their flight will take more than 80 minutes? *Ans. 0.038.*

8.12 Six measurements were made of an angle as follows:

32°04′	32°05′
31°59′	31°57′
32°01′	32°00′

Assume that the measurement error is Gaussian with zero mean; and the standard deviation of each measurement can be represented by the sample standard deviation of the six measurements above.

(a) Estimate the angle.

(b) Subsequently, the engineer discovered that the angle has been measured before, and recorded as 32°00′ ± 2′. Estimate the actual angle using both sets of measurements.

8.13 A distance L is measured independently by three surveyors with three sets of instruments. The respective measurements are 2.15, 2.20, 2.18 km. Suppose the ratio of the standard error of the three measurements is 1:2:3. Estimate the actual distance L on the basis of the three sets of measurements. Assume that measurement error is Gaussian with zero mean.

8.14 In designing reinforced concrete structural members to resist ultimate load, a capacity reduction factor ϕ is often used. Suppose that the structural member is a beam element and it is designed for pure flexure. The conventional value of ϕ is 0.9. However, a committee is investigating the effect on the probability of failure of beams against ultimate load if ϕ is increased to 0.95. Twelve beams are designed using $\phi = 0.95$, and each of them is subjected to the designed ultimate load in the laboratory. It is desired to estimate p, the probability of failure of such beams against ultimate load based on the prior judgmental information as well as experimental outcome.
Suppose that the prior distribution of p has a mean of 0.1 and standard deviation of 0.06; and one out of 12 beams tested failed the ultimate load. Suggest a suitable prior distribution and determine the mean and variance of p from these data.

8.15 The time between breakdowns of a certain type of construction equipment follows an exponential distribution

$$f_X(x) = \lambda e^{-\lambda x} \qquad x \geq 0$$

where the mean rate of failure λ was rated by the manufacturer to have a mean of 0.5 per year and a COV of 20%. A contractor owns two pieces of this construction equipment. The operational times until breakdown of the equipments were subsequently observed to be 12 and 18 months, respectively. Using conjugate distributions, determine the updated mean and COV of the parameter λ.