

Lecture 7: Prior distribution, Bayes risk and Bayes rule

Statistical Inference–II ■ Unit I: Decision Theory

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

M.Sc. Statistics, Semester III
Academic Session 2026–2027

Previous lecture

Lecture 6: Minimax and Bayes rules among non-randomised rules: the Bernoulli example.

Today

- The Bayes Principle
 - Prior, Bayes Risk and Bayes Rule

The Bayes approach is an alternative method of reducing the risk function to a single number. Here the parameter θ is itself treated as a random variable with a specified distribution, called the *prior distribution*, which expresses whatever is known or believed about θ before the data are observed.

Definition 7.1 (Prior distribution)

A *prior distribution* for θ is a probability distribution on Θ , with density (or mass function) $g(\theta)$ with respect to a suitable dominating measure.

Definition 7.2 (Bayes risk)

The *Bayes risk* of a decision rule δ with respect to the prior g is the average of the risk function over the prior:

$$r(g, \delta) = \mathbb{E}_g[R(\theta, \delta)] = \int_{\Theta} R(\theta, \delta) g(\theta) d\theta,$$

the integral being replaced by a sum when Θ is discrete.

Definition 7.3 (Bayes rule)

A rule $\delta_g \in \mathcal{D}$ is a *Bayes rule* with respect to the prior g if it has the smallest Bayes risk in $\mathcal{D}$, that is,

$$r(g, \delta_g) \leq r(g, \delta) \quad \text{for all } \delta \in \mathcal{D}.$$

The value

$$r(g) = \inf_{\delta \in \mathcal{D}} r(g, \delta) = r(g, \delta_g)$$

is called the *Bayes risk of the problem* with respect to g .

Remark 7.4 (Four distinct quantities)

Students very frequently confuse four different objects. It is worth displaying them together:

Quantity	Notation	Depends on
Loss	$L(\theta, \delta(\mathbf{X}))$	θ and $\mathbf{X}$ (random)
Risk (frequentist)	$R(\theta, \delta) = \mathbb{E}_\theta[L]$	θ (and δ)
Bayes risk	$r(g, \delta) = \mathbb{E}_g[R(\theta, \delta)]$	neither: a number
Posterior expected loss	$\rho(\pi, a \mathbf{x}) = \mathbb{E}^{\theta \mathbf{x}}[L(\theta, a)]$	$\mathbf{x}$ and a

($\pi(\theta | \mathbf{x})$ and $\rho(\pi, a | \mathbf{x})$ are defined formally in the Structure Theorem below.) The Bayes risk is a single number obtained by a *double* expectation — first over $\mathbf{X}$ given θ , then over θ . The posterior expected loss is a function of $\mathbf{x}$ (and a) and is *not* the Bayes risk.

Example 7.5 (Lot sampling)

Consider the inspection of a lot of N manufactured items containing an unknown number M of defectives. A random sample of predetermined size n is drawn without replacement, and X denotes the number of defectives observed. Then X has the hypergeometric distribution

$$\mathbb{P}(X = x \mid M) = \begin{cases} \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}, & x = 0, 1, \dots, n, \\ 0, & \text{otherwise,} \end{cases}$$

with the single unknown parameter M .

Example 7.5 (cont.)

If the production process has a long-run fraction defective p , and items are defective independently, a natural prior for M is binomial:

$$g(m) = \mathbb{P}(M = m) = \begin{cases} \binom{N}{m} p^m (1-p)^{N-m}, & m = 0, 1, \dots, N, \\ 0, & \text{otherwise,} \end{cases}$$

with N and p known, p being the process average fraction defective.

- The action space is $\mathcal{A} = \{\text{accept the lot, reject the lot}\}$, and a loss function is specified reflecting the cost of accepting a bad lot and of rejecting a good one.
- This is a genuine decision problem of industrial importance, and it was in such contexts that the theory was first applied.

Example 7.5 (cont.): is the sample informative?

With the binomial prior,

$$M - x \mid X = x \sim \text{Binomial}(N - n, p) \quad \text{whatever } x \text{ is}$$

(end-of-chapter Exercise 1.2), so the sample is uninformative about the $N - n$ unsampled items. If instead M is given a beta-binomial prior (the fraction defective p itself uncertain, with a beta distribution), sampling becomes informative about them.

Practice (Lecture Notes, Chapter 1)

- Short-answer questions (Section 1.12): Q8.
- Exercises (Section 1.13): 1.14.

Next lecture

Lecture 8: Structure of the Bayes rule: posterior expected loss.