

Lecture 11: Worked examples: linear rules; the beta-binomial model

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 10: Small-neighbourhood (MAP) and LINEX losses.

Today

- The Bayes Principle
 - Worked Examples on Bayes Estimation

Example 11.1 (Bayes risk of a linear rule, normal model)

In the setting of the example on linear estimators of a normal mean ($X \sim N(\theta, 1)$, SELF, rules $\delta_c(X) = cX$), suppose further that θ is a random variable with prior $N(0, \tau^2)$. Find the Bayes risk of δ_c and the value of c minimising it.

Solution

From the linear-estimator example, $R(\theta, \delta_c) = c^2 + (1 - c)^2\theta^2$. Hence

$$r(g, \delta_c) = \mathbb{E}_g[R(\theta, \delta_c)] = \mathbb{E}_g[c^2 + (1 - c)^2\theta^2] = c^2 + (1 - c)^2 \mathbb{E}_g(\theta^2).$$

Since $\theta \sim N(0, \tau^2)$ we have $\mathbb{E}_g(\theta^2) = \text{Var}_g(\theta) + \{\mathbb{E}_g(\theta)\}^2 = \tau^2 + 0 = \tau^2$, so

$$r(g, \delta_c) = c^2 + (1 - c)^2\tau^2.$$

Solution (cont.)

Minimising with respect to c :

$$\frac{\partial}{\partial c} r(g, \delta_c) = 2c + 2(1 - c)\tau^2 \cdot (-1) = 2c - 2\tau^2(1 - c) = 0,$$

so $c + \tau^2 c = \tau^2$, giving

$$c_0 = \frac{\tau^2}{1 + \tau^2}$$

The second derivative is $2 + 2\tau^2 > 0$, confirming a minimum.

Solution (cont.)

The minimum Bayes risk is

$$\begin{aligned} r(g, \delta_{c_0}) &= c_0^2 + (1 - c_0)^2 \tau^2 = \left(\frac{\tau^2}{1 + \tau^2} \right)^2 + \left(1 - \frac{\tau^2}{1 + \tau^2} \right)^2 \tau^2 \\ &= \frac{\tau^4}{(1 + \tau^2)^2} + \left(\frac{1}{1 + \tau^2} \right)^2 \tau^2 = \frac{\tau^4}{(1 + \tau^2)^2} + \frac{\tau^2}{(1 + \tau^2)^2} \\ &= \frac{\tau^2(\tau^2 + 1)}{(1 + \tau^2)^2} = \frac{\tau^2}{1 + \tau^2}. \end{aligned}$$

Thus $r(g, \delta_{c_0}) = \tau^2/(1 + \tau^2)$, which numerically coincides with c_0 . This coincidence is peculiar to the present parametrisation and should not be taken as a general rule; the cancellation is displayed above so that the reader may verify it.

The best linear rule is the overall Bayes rule

Remark 11.2

Since $\delta_{c_0}(X) = \tau^2 X / (1 + \tau^2)$ is Bayes within the class $\{\delta_c\}$, and since (as will be shown in the normal–normal example) the overall Bayes estimator for a $N(\theta, 1)$ observation with $N(0, \tau^2)$ prior is exactly $\tau^2 X / (1 + \tau^2)$, the restriction to linear rules costs nothing here.

- The reason is that the prior mean is 0, so the overall Bayes rule happens to lie in the class $\{cX\}$: with a $N(\mu_0, \tau^2)$ prior, $\mu_0 \neq 0$, the Bayes estimator is $(\tau^2 X + \mu_0) / (1 + \tau^2)$, which is affine but not of the form cX .
- Nor is the phenomenon special to the normal–normal pair: conjugate priors in exponential families generally give posterior means that are affine in the data, e.g. $(x + 1) / (n + 2)$ in the beta–binomial example and $(S + \alpha) / (n + \beta)$ in the Poisson–gamma example.

Example 11.3 (Beta-binomial: Bayes estimator of a proportion)

Let $X \sim \text{Binomial}(n, p)$ and let the prior for p be uniform on $(0, 1)$, i.e. $g(p) = 1$, $0 < p < 1$. Under squared error loss $L(p, a) = (a - p)^2$, obtain the Bayes estimator of p , the posterior expected loss, and the Bayes risk of the problem.

Solution

The likelihood is

$$f(x | p) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, \dots, n.$$

Solution (cont.)

Step 1: marginal of X . Here we write $m(x)$ for the marginal probability function of X (the $p(x)$ of the Structure Theorem), to avoid a clash with the parameter p :

$$m(x) = \int_0^1 \binom{n}{x} p^x (1-p)^{n-x} \cdot 1 \, dp = \binom{n}{x} B(x+1, n-x+1),$$

where $B(\alpha, \beta) = \int_0^1 u^{\alpha-1} (1-u)^{\beta-1} du = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$. Using $\Gamma(k) = (k-1)!$ for integer k ,

$$B(x+1, n-x+1) = \frac{x! (n-x)!}{(n+1)!},$$

Solution (cont.)

Substituting $B(x+1, n-x+1) = x!(n-x)!/(n+1)!$ into $m(x) = \binom{n}{x}B(x+1, n-x+1)$, we obtain

$$m(x) = \frac{n!}{x!(n-x)!} \cdot \frac{x!(n-x)!}{(n+1)!} = \frac{n!}{(n+1)!} = \frac{1}{n+1}, \quad x = 0, 1, \dots, n.$$

That is, the marginal distribution of X is *discrete uniform* on $\{0, 1, \dots, n\}$ — a fact we shall need in Step 4.

Solution (cont.)

Step 2: posterior.

$$\pi(p \mid x) = \frac{f(x \mid p)g(p)}{m(x)} = \frac{p^x(1-p)^{n-x}}{B(x+1, n-x+1)}, \quad 0 < p < 1,$$

which is the $\text{Beta}(x+1, n-x+1)$ density.

Step 3: Bayes estimator. By the corollary “Bayes estimator under SELF is the posterior mean”, the Bayes estimator is the posterior mean. For $\text{Beta}(\alpha, \beta)$ the mean is $\alpha/(\alpha + \beta)$; here $\alpha = x+1$, $\beta = n-x+1$, so

$$\hat{p}_{\text{Bayes}} = \delta_g(x) = \frac{x+1}{(x+1) + (n-x+1)} = \frac{x+1}{n+2}$$

This is the classical *Laplace rule of succession*.

Solution (cont.)

Step 4: posterior expected loss and Bayes risk. The posterior expected loss at x is the posterior variance. For $\text{Beta}(\alpha, \beta)$ the variance is $\alpha\beta/\{(\alpha + \beta)^2(\alpha + \beta + 1)\}$, hence

$$\rho(\pi, \delta_g(x) \mid x) = \text{Var}(p \mid x) = \frac{(x+1)(n-x+1)}{(n+2)^2(n+3)}.$$

Note that this *depends on x* ; it is therefore not the Bayes risk (see the remark “A common error”).

Solution (cont.)

The Bayes risk of the problem is the expectation of the posterior expected loss over the marginal of X , which by Step 1 is uniform on $\{0, \dots, n\}$:

$$\begin{aligned} r(g) &= \sum_{x=0}^n \frac{(x+1)(n-x+1)}{(n+2)^2(n+3)} \cdot \frac{1}{n+1} \\ &= \frac{1}{(n+1)(n+2)^2(n+3)} \sum_{x=0}^n (x+1)(n+1-x). \end{aligned}$$

Solution (cont.)

Put $k = x + 1$, so k runs from 1 to $n + 1$ and $n + 1 - x = n + 2 - k$:

$$\begin{aligned}\sum_{x=0}^n (x+1)(n+1-x) &= \sum_{k=1}^{n+1} k(n+2-k) \\&= (n+2) \sum_{k=1}^{n+1} k - \sum_{k=1}^{n+1} k^2 \\&= (n+2) \cdot \frac{(n+1)(n+2)}{2} - \frac{(n+1)(n+2)(2n+3)}{6} \\&= \frac{(n+1)(n+2)}{6} [3(n+2) - (2n+3)] \\&= \frac{(n+1)(n+2)}{6} (n+3).\end{aligned}$$

Solution (cont.)

Substituting $\sum_{x=0}^n (x+1)(n+1-x) = \frac{(n+1)(n+2)(n+3)}{6}$ into

$$r(g) = \frac{1}{(n+1)(n+2)^2(n+3)} \sum_{x=0}^n (x+1)(n+1-x),$$

we therefore obtain

$$r(g) = \frac{1}{(n+1)(n+2)^2(n+3)} \cdot \frac{(n+1)(n+2)(n+3)}{6} = \boxed{\frac{1}{6(n+2)}}$$

Remark 11.4 (Cross-check via the frequentist risk)

As a verification, compute the risk function of δ_g directly. With $\delta_g(X) = (X + 1)/(n + 2)$ and $\mathbb{E}_p(X) = np$, $\text{Var}_p(X) = np(1 - p)$,

$$\begin{aligned} R(p, \delta_g) &= \text{Var}_p\left(\frac{X + 1}{n + 2}\right) + \left\{ \mathbb{E}_p\left(\frac{X + 1}{n + 2}\right) - p \right\}^2 \\ &= \frac{np(1 - p)}{(n + 2)^2} + \left\{ \frac{np + 1 - (n + 2)p}{n + 2} \right\}^2 \\ &= \frac{np(1 - p)}{(n + 2)^2} + \frac{(1 - 2p)^2}{(n + 2)^2} = \frac{np(1 - p) + (1 - 2p)^2}{(n + 2)^2}. \end{aligned}$$

Remark (cont.)

Averaging over the uniform prior, using

$$\int_0^1 p(1-p) dp = \frac{1}{6} \quad \text{and} \quad \int_0^1 (1-2p)^2 dp = \frac{1}{3},$$

$$r(g, \delta_g) = \frac{n/6 + 1/3}{(n+2)^2} = \frac{n+2}{6(n+2)^2} = \frac{1}{6(n+2)},$$

in agreement with the direct computation.

Remark (cont.)

Moreover,

$np(1-p) + (1-2p)^2 = np(1-p) + 1 - 4p(1-p) = 1 + (n-4)p(1-p)$, so

$$R(p, \delta_g) = \frac{1 + (n-4)p(1-p)}{(n+2)^2},$$

which is constant in p only when $n = 4$.

Whether δ_g is minimax is settled in the remark “Is Laplace’s estimator minimax?” in the section on relations among Bayes, minimax and admissible rules.

Practice (Lecture Notes, Chapter 1)

- Exercises (Section 1.13): 1.2, 1.8.

Next lecture

Lecture 12: Normal–normal and Poisson–gamma models; different losses, different estimates.