

**HW #3 (due at Canvas midnight on Thursday, September 14 ET)**

(There are 5 questions. The hints to Problem 1 and Problem 4(c) can be found on page #3 of this pdf file)

**Problem 1.** In Problem 1 of HW#1, we assume that  $Y_1, \dots, Y_n$  are iid normal  $N(\theta, 1)$ , and we want to guess  $\theta$  when the loss function is given by  $L(\theta, d) = (\theta - d)^2 / (1 + \theta^2)$ . One of the proposed procedures is:

$$\delta_{2,n}(Y_1, \dots, Y_n) = \frac{\bar{Y}_n + n^{-1}}{1 + n^{-1}}.$$

The purpose of this exercise is to show that this procedure is actually Bayes. Show that the procedure  $\delta_{2,n}$  is Bayes relative to the prior density  $\pi_a(\theta) = C_1(1 + \theta^2)\phi(\theta - 1)$ , where  $\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$  is the standard normal  $N(0, 1)$  density and  $C_1$  is a suitable constant. You need *not* determine the  $C_1$ ; it suffices to verify that the given function of  $\theta$  have finite integrals, so that one knows such  $C_1$  exist.

**Problem 2.** Suppose  $Y_1, Y_2, \dots, Y_n$  ( $n \geq 2$ ) are independent and identically distributed (iid) with a Uniform $[0, \theta]$ , and consider a Pareto prior distribution,  $\theta \sim PA(\alpha, \beta)$ , i.e.,  $\theta$  has a prior density

$$\pi(\theta) = \frac{\alpha\beta^\alpha}{\theta^{\alpha+1}}, \quad \text{for } \theta \geq \beta,$$

and mean  $\mathbf{E}(\theta) = \alpha\beta/(\alpha - 1)$ , for some known  $\alpha > 1$  and  $\beta > 0$ .

(a) Show that the posterior distribution of  $\theta$  has a Parato  $PA(\alpha^*, \beta^*)$  distribution with  $\alpha^* = \alpha + n$  and  $\beta^* = \max\{\beta, Y_{(n)}\}$ , where  $Y_{(n)} = \max(Y_1, \dots, Y_n)$ .

(b) Find the Bayes procedure  $\delta_B^*(X)$  under the “squared error loss” function  $L(\theta, d) = (\theta - d)^2$ .

(c) Find the Bayes procedure  $\delta_C^*(X)$  under the “absolute error loss” function  $L(\theta, d) = |\theta - d|$ .

[Hints for (a): the domain of  $\theta$  also includes  $Y_1 \leq \theta, \dots, Y_n \leq \theta$ , which is equivalent to  $\theta \geq Y_{(n)}$ .]

**Problem 3. (LINEX Loss. Problem 7.65 on page 367 of our text).** Suppose  $\Omega$  can be labeled according to the values of a real parameter  $\theta$ , and the decisions are  $D = \{d : -\infty < d < \infty\}$ , representing guesses as to the true value of  $\theta$ . The loss function is the LINEX (LINear-EXponential) loss given by

$$L(\theta, d) = e^{c(d-\theta)} - c(d - \theta) - 1,$$

where  $c$  is a positive constant. The LINEX loss was investigated by Zellner (1986), and can handle asymmetries in a smooth way: as the constant  $c$  varies, the loss function varies from very asymmetric to almost symmetric. We also assume that the prior density on  $\Omega = \{\theta : -\infty < \theta < \infty\}$  is  $\pi(\theta)$ .

(a) For  $c = 0.2, 0.5, 1$ , plot  $L(\theta, d)$  as a function of  $d - \theta$ . Feel free to use any computer software.

(b) “No data problem.” Suppose we had to make a guess of  $\theta$  based on no observations, but merely on the given prior law  $\pi(\theta)$ . Show that you minimize your expected loss  $\mathbf{E}(L(\theta, d)) = \int_{-\infty}^{\infty} L(\theta, d)\pi(\theta)d\theta$  by guessing  $\theta$  as  $\delta_b = -\frac{1}{c} \log \int_{-\infty}^{\infty} e^{-c\theta}\pi(\theta)d\theta = -\frac{1}{c} \log \mathbf{E}(e^{-c\theta})$ .

(c) (2 pts). “Problem with data.” Now suppose that we also observe data  $\mathbf{Y}$  which has distribution function  $F(\mathbf{y}|\theta)$ . Show that the Bayes procedure is given by  $\delta_c(\mathbf{Y}) = -\frac{1}{c} \log \mathbf{E}(e^{-c\theta}|\mathbf{Y})$ .

(d) Let  $Y_1, \dots, Y_n$  be iid  $N(\theta, \sigma^2)$ , where  $\sigma^2$  is known, and suppose that  $\theta$  has the so-called “noninformative” prior  $\pi(\theta) \equiv 1$ . Show that the Bayes procedure with respect to LINEX loss is given by  $\delta_d(\bar{Y}) = \bar{Y} - (c\sigma^2/(2n))$ . See the remarks/hints at the end of this question.

(e) Under the assumption of part (d), calculate the posterior expected loss for both  $\delta_d(\bar{Y})$  and  $\delta_e(\bar{Y}) = \bar{Y}$  using LINEX loss. Recall that for a given procedure  $\delta(\mathbf{Y})$ , the posterior expected loss is defined as  $h_\pi(\mathbf{y}, \delta(\mathbf{y}))$  where  $h_\pi(\mathbf{y}, d) = \int_{-\infty}^{\infty} L(\theta, d)\pi(\theta|\mathbf{y})d\theta$ .

- (f) Repeat part (e) using squared error loss, i.e., calculate the posterior expected loss for both  $\delta_d(\bar{Y})$  and  $\bar{Y}$  when the loss function is  $L(\theta, d) = (\theta - d)^2$ .

[Remarks/Hints for (d): Rigorously speaking,  $\pi(\theta) \equiv 1$  is not a probability density on  $\Omega = (-\infty, \infty)$ , but for our purpose, here we pretend that it can be thought of as a valid prior distribution. To simplify your computation, you can assume that you just observe “one new observation”  $\bar{Y}_n \sim N(\theta, \frac{\sigma^2}{n})$  (this simplification is reasonable because  $\bar{Y}_n$  is a sufficient statistic which will be discussed later). For a data  $\bar{Y}$ , what is the marginal distribution  $m(\bar{y}) = \int_{-\infty}^{\infty} f_{\theta}(\bar{y})\pi(\theta)d\theta$  when  $\pi(\theta) \equiv 1$ ? What is the posterior distribution  $\pi(\theta|\bar{y}) = f_{\theta}(\bar{y})\pi(\theta)/m(\bar{y})$ ? Is it well-defined?]

**Problem 4.** Suppose that a random variable  $Y$  has a  $\text{Gamma}(q, 1/\theta)$  distribution with density  $f_{\theta}(y) = \frac{\theta^q y^{q-1} e^{-\theta y}}{\Gamma(q)}$ , for  $y \geq 0$ , where  $\Gamma(q) = \int_0^{\infty} y^{q-1} e^{-y} dy = (q-1)\Gamma(q-1)$  and  $q$  is positive and known. Here  $S = \{y : y \geq 0\}, \Omega = \{\theta : \theta \geq 0\}, D = \{d : d \geq 0\}$ . Under a Bayesian setting, suppose that the prior distribution  $\theta$  is  $\text{Gamma}(\alpha, \beta)$ , i.e.,  $\theta$  has a prior density  $\pi(\theta) = \frac{\theta^{\alpha-1} e^{-\theta/\beta}}{\Gamma(\alpha)\beta^{\alpha}}$  for  $\theta \geq 0$ .

- (a) Show that the posterior distribution of  $\theta$  has a  $\text{Gamma}(\alpha', \beta')$  distribution with  $\alpha' = q + \alpha$  and  $\beta' = \frac{\beta}{1 + \beta y}$ .

This implies that  $\pi(\theta|y) = \frac{\pi(\theta)f_{\theta}(y)}{\int_0^{\infty} \pi(\theta)f_{\theta}(y)d\theta}$  can be written as  $\pi(\theta|y) = \frac{\theta^{\alpha'-1} e^{-\theta/\beta'}}{\Gamma(\alpha')(\beta')^{\alpha'}}$ .

- (b) (**Point Estimation**) Find the Bayes procedure  $\delta_{\text{Bayes}}(Y)$  under the loss function  $L(\theta, d) = (\theta - d)^2$ .

- (c) (**Confidence Interval**) In the fixed-width confidence interval estimation setting, suppose one wants to estimate  $\theta$  by the interval  $[d - 2, d + 2]$  for some decision  $d \geq 2$ . This corresponds to estimate  $\theta$  under a new binary loss function  $L^*(\theta, d) = \begin{cases} 1 & \text{if } |\theta - d| > 2; \\ 0 & \text{if } |\theta - d| \leq 2. \end{cases}$  That is, the loss is the same for all “bad” decisions (those which mis-estimate  $\theta$  by more than 2 units) and is 0 for all “correct” decisions. Show that the Bayes procedure is given by

$$\begin{aligned} \delta^*(Y) &= \begin{cases} 2, & \text{if } q + \alpha \leq 1; \\ \text{the solution } d \text{ of } \pi(d + 2|Y) = \pi(d - 2|Y), & \text{if } q + \alpha > 1. \end{cases} \\ &= \begin{cases} 2, & \text{if } q + \alpha \leq 1; \\ -2 + 4/[1 - \exp(-\frac{4(Y+1/\beta)}{\alpha+q-1})], & \text{if } q + \alpha > 1. \end{cases} \end{aligned}$$

In other words, the Bayes confidence interval with fixed-width of 4 is  $[\delta^*(Y) - 2, \delta^*(Y) + 2]$ .

**Problem 5 (Bayes with constant risk).** Assume that  $Y_1, \dots, Y_n$  are iid Bernoulli( $\theta$ ) with  $0 \leq \theta \leq 1$ , i.e.,  $\mathbf{P}_{\theta}(Y_i = y) = \theta^y(1 - \theta)^{1-y}$  for  $y = 0$  or  $1$ , and suppose that we want to estimate  $\theta$  under the square error loss function  $L(\theta, d) = (\theta - d)^2$ . We have shown in class when  $\theta$  has a prior Beta( $\alpha, \beta$ ) distribution, i.e.,

$$\pi(\theta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1}, \quad 0 \leq \theta \leq 1$$

for two pre-specified constants  $\alpha > 0, \beta > 0$ , the Bayes procedure is given by

$$\delta_{\alpha, \beta, n}(Y_1, \dots, Y_n) = \frac{\alpha + \sum_{i=1}^n Y_i}{\alpha + \beta + n} = \frac{\alpha + \beta}{\alpha + \beta + n} \times \frac{\alpha}{\alpha + \beta} + \frac{n}{\alpha + \beta + n} \bar{Y}_n,$$

which is the weighted average of the prior mean  $\alpha/(\alpha + \beta)$  and the sample mean  $\bar{Y}_n = (Y_1 + \dots + Y_n)/n$ .

- (a) Compute the risk function of the Bayes procedure  $\delta_{\alpha, \beta, n}(Y_1, \dots, Y_n)$ .
- (b) An interesting scenario is to choose a specific pair of  $\alpha^* > 0$  and  $\beta^* > 0$  in such a way that the corresponding Bayes procedure has constant risks when  $\alpha = \alpha^*$  and  $\beta = \beta^*$ , i.e., the risk function in part (a) does not depend on  $\theta$ . Find such a pair  $(\alpha^*, \beta^*)$ . [A check:  $\alpha^* = \sqrt{n}/2$ .]
- (c) Calculate the Bayes risk of the Bayes procedure  $\delta_{\alpha^*, \beta^*, n}$  in part (b) when  $\theta$  has a prior distribution  $\pi(\theta) = \text{Beta}(\alpha^*, \beta^*)$ .
- (d) Compute the risk function of the sample mean  $\delta_0 = \delta_0(Y_1, \dots, Y_n) = \bar{Y}_n$ , and calculate the Bayes risk of  $\delta_0$  when  $\theta$  has a prior distribution  $\pi(\theta) = \text{Beta}(\alpha^*, \beta^*)$ .
- (e) Show that the Bayes procedure  $\delta_{\alpha^*, \beta^*, n}$  indeed has a smaller Bayes risk than the sample mean  $\delta_0 = \bar{Y}_n$  when  $\theta$  has a prior distribution  $\pi(\theta) = \text{Beta}(\alpha^*, \beta^*)$ , where  $\alpha^*$  and  $\beta^*$  are determined in part (b).

Hint to Problem 1: You can show that a Bayes procedure shall minimize

$$h_{\pi}^*(\mathbf{y}, d) = \int_{-\infty}^{\infty} L(\theta, d) f_{\theta}(y_1, \dots, y_n) \pi_a(\theta) d\theta = C_1 \int_{-\infty}^{\infty} (\theta - d)^2 f_{\theta}(y_1, \dots, y_n) \phi(\theta - 1) d\theta,$$

where the integral is just the Bayes risk when the loss function is squared error loss and  $\theta$  has a prior distribution  $N(\mu = 1, \tau^2 = 1)$ . Which decision  $d$  minimizes this integral? Did we find such a minimizer  $d$  in class?

Hints to Problem 4(c): Note that a  $\text{Gamma}(\alpha, \beta)$  random variable has mean  $\alpha\beta$ , see page 624 of our textbook for more properties of Gamma distribution. In part (c), we want to choose  $d$  to minimize

$$\int_0^{\infty} L^*(\theta, d) \pi(\theta|y) d\theta = 1 - \int_{d-2}^{d+2} \pi(\theta|y) d\theta. \quad (1)$$

Note that in this problem,  $\pi(\theta|y)$  is continuous and *unimodal*; that is, there is a *mode* (possibly 0) such that  $\pi(\theta|y)$  increases for  $\theta < \text{mode}$  and decreases for  $\theta > \text{mode}$ . Then it is an easy calculus exercise to show that relation (1) is minimized by that value  $d^*$  for which  $\pi(d^* - 2|y) = \pi(d^* + 2|y)$  if  $\pi(0|y) < \pi(4|y)$ , and by  $d^* = 2$  otherwise (please feel free to use this fact if you cannot prove it or if you simply do not want to spend time to prove it! No penalty here!)

In addition, note that  $\pi(0|y) = \infty$  if  $\alpha' = q + \alpha < 1$  and  $\pi(0|y) = 0$  if  $\alpha' > 1$ . Hence,  $\pi(0|y) \geq \pi(4|y)$  if and only if  $q + \alpha \leq 1$ . What happens if  $q + \alpha = 1$ ?