Stat 346/446: Theoretical Statistics II: Homework 7 Solutions

Textbook Exercises

7.62 Let $X_1, \ldots, X_n$ be a random sample from a $\mathcal{N}(\theta, \sigma^2)$ population, σ^2 known. Consider estimating θ using squared error loss. Let $\pi(\theta)$ be a $\mathcal{N}(\mu, \tau^2)$ prior distribution on θ and let δ^π be the Bayes estimator of θ. Verify the following formulas for the risk function and Bayes risk.

(a) (346 : 2 pts, 446 : 1 pts.) For any constants a and b, the estimator $\delta(x) = a\bar{X} + b$ has risk function

$$R(\theta, \delta) = a^2 \frac{\sigma^2}{n} + (b - (1-a)\theta)^2.$$

Let $\delta(\mathbf{X}) = a\bar{X} + b$, where $\bar{X}$ is the sample mean. We compute the risk under squared error loss:

$$R(\theta, \delta) = \mathbb{E}_\theta[(\delta(\mathbf{X}) - \theta)^2].$$

Since $\bar{X} \sim \mathcal{N}(\theta, \sigma^2/n)$, we have:

$$\mathbb{E}[\delta(\mathbf{X})] = a\theta + b, \quad \mathrm{Var}[\delta(\mathbf{X})] = a^2 \cdot \frac{\sigma^2}{n}.$$

By the bias-variance decomposition:

$$R(\theta, \delta) = \mathrm{Var}(\delta(\mathbf{X})) + (\mathbb{E}[\delta(\mathbf{X})] - \theta)^2 = a^2 \frac{\sigma^2}{n} + (a\theta + b - \theta)^2.$$

Simplifying the squared bias term:

$$(a\theta + b - \theta)^2 = (b - (1-a)\theta)^2.$$

Therefore, the risk is:

$$\boxed{R(\theta, \delta) = a^2 \frac{\sigma^2}{n} + (b - (1-a)\theta)^2}.$$

(b) (346 : 2 pts, 446 : 1 pts.) Let $\eta = \frac{\sigma^2}{n\tau^2 + \sigma^2}$. The risk function for the Bayes estimator is

$$R(\theta, \delta^\pi) = (1-\eta)^2 \frac{\sigma^2}{n} + \eta^2 (\theta - \mu)^2.$$

The Bayes estimator under squared error loss is the posterior mean. Posterior:

$$\theta \mid \bar{X} \sim \mathcal{N}\left(\frac{n\tau^2 \bar{X} + \sigma^2 \mu}{n\tau^2 + \sigma^2}, \quad \frac{\sigma^2 \tau^2}{n\tau^2 + \sigma^2} \right)$$

So the Bayes estimator $\delta^\pi(X)$ is:

$$\delta^\pi(X) = \mathbb{E}[\theta \mid \bar{X}] = (1-\eta)\bar{X} + \eta\mu$$

where

$$\eta = \frac{\sigma^2}{n\tau^2 + \sigma^2}$$

This is of the form $a\bar{X} + b$ with: $a = 1 - \eta$ and $b = \eta\mu$.
Using the result from part (a), we get:

$$R(\theta, \delta^\pi) = (1-\eta)^2 \frac{\sigma^2}{n} + (\eta\mu - \eta\theta)^2 = (1-\eta)^2 \frac{\sigma^2}{n} + \eta^2(\theta - \mu)^2.$$

So the risk function is:

$$\boxed{R(\theta, \delta^\pi) = (1-\eta)^2 \frac{\sigma^2}{n} + \eta^2(\theta - \mu)^2}.$$

(c) (346 & 446 : 1 pts.) The Bayes risk for the Bayes estimator is

$$B(\pi, \delta^\pi) = \tau^2\eta.$$

The Bayes risk is the expected risk under the prior:

$$B(\pi, \delta^\pi) = \mathbb{E}_\pi[R(\theta, \delta^\pi)].$$

Using the expression from (b):

$$B(\pi, \delta^\pi) = (1-\eta)^2 \frac{\sigma^2}{n} + \eta^2 \mathbb{E}_\pi[(\theta - \mu)^2].$$

Since $\theta \sim \mathcal{N}(\mu, \tau^2)$, we have $\mathbb{E}_\pi[(\theta - \mu)^2] = \tau^2$, so:

$$B(\pi, \delta^\pi) = (1-\eta)^2 \frac{\sigma^2}{n} + \eta^2\tau^2.$$

Now recall:

$$\eta = \frac{\sigma^2}{n\tau^2 + \sigma^2} \Rightarrow (1-\eta)^2 \frac{\sigma^2}{n} + \eta^2\tau^2 = \tau^2\eta.$$

Therefore, the Bayes risk is:

$$\boxed{B(\pi, \delta^\pi) = \tau^2\eta}.$$

7.65 A loss function investigated by Zellner (1986) is the LINEX (LINear-EXponential) loss, a loss function that can handle asymmetries in a smooth way. The LINEX loss is given by

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

where c is a positive constant. As the constant c varies, the loss function varies from very asymmetric to almost symmetric.

(a) (446 : 1 pts.) For $c = 0.2,\ 0.5,\ 1$, plot $L(\theta, a)$ as a function of $a - \theta$.
The plot is in Figure 1

(b) (446 : 2 pts.) If $X \sim F(x|\theta)$, show that the Bayes estimator of θ, using a prior π, is given by

$$\delta^\pi(X) = -\frac{1}{c} \log \mathbb{E}(e^{-c\theta} \mid X).$$

We want to minimize the posterior expected loss:

$$\delta^\pi(X) = \arg\min_a \mathbb{E}[L(\theta, a) \mid X] = \arg\min_a \left\{ \mathbb{E}[e^{c(a-\theta)} \mid X] - c(a - \mathbb{E}[\theta \mid X]) - 1 \right\}.$$

Note that:

$$\mathbb{E}[L(\theta, a) \mid X] = e^{ca} \cdot \mathbb{E}[e^{-c\theta} \mid X] - ca + c\mathbb{E}[\theta \mid X] - 1.$$

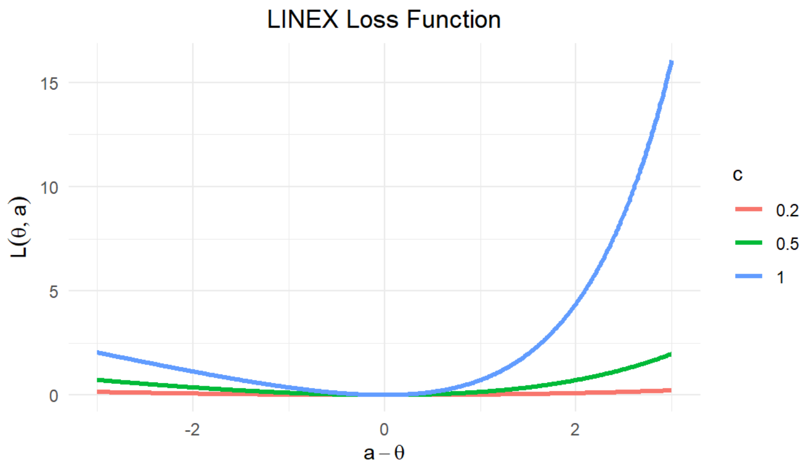

Figure 1: LINEX Loss Function

Taking derivative with respect to a:

$$\frac{d}{da}\mathbb{E}[L(\theta,a) \mid X] = ce^{ca}\mathbb{E}[e^{-c\theta} \mid X] - c.$$

Setting derivative to zero:

$$e^{ca}\mathbb{E}[e^{-c\theta} \mid X] = 1 \quad \Rightarrow \quad ca = -\log \mathbb{E}[e^{-c\theta} \mid X].$$

Thus, the Bayes estimator under LINEX loss is:

$$\boxed{\delta^{\pi}(X) = -\frac{1}{c}\log \mathbb{E}[e^{-c\theta} \mid X]}.$$

(c) (446 : 1 pts.) Let $X_1, \ldots, X_n$ be iid $\mathcal{N}(\theta, \sigma^2)$, where σ^2 is known, and suppose that θ has the noninformative prior $\pi(\theta) = 1$. Show that the Bayes estimator versus LINEX loss is given by

$$\delta^B(\bar{X}) = \bar{X} - \left(\frac{c\sigma^2}{2n}\right).$$

Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\theta, \sigma^2)$ and $\pi(\theta) = 1$ be the noninformative prior. Then the posterior distribution is:

$$\theta \mid \bar{X} \sim \mathcal{N}\left(\bar{X}, \frac{\sigma^2}{n}\right).$$

We want to compute:

$$\delta^B(\bar{X}) = -\frac{1}{c}\log \mathbb{E}[e^{-c\theta} \mid \bar{X}].$$

Since $\theta \mid \bar{X} \sim \mathcal{N}(\bar{X}, \sigma^2/n)$, we use the moment-generating function:

$$\mathbb{E}[e^{-c\theta} \mid \bar{X}] = \exp\left(-c\bar{X} + \frac{c^2\sigma^2}{2n}\right).$$

Taking log and negating:

$$\delta^B(\bar{X}) = -\frac{1}{c}\left(-c\bar{X} + \frac{c^2\sigma^2}{2n}\right) = \bar{X} - \frac{c\sigma^2}{2n}.$$

Thus, the Bayes estimator is:

$$\boxed{\delta^B(\bar{X}) = \bar{X} - \frac{c\sigma^2}{2n}}.$$

8.55 Let X have a $n(\theta, 1)$ distribution, and consider testing $H_0\colon \theta \geq \theta_0$ versus $H_1\colon \theta < \theta_0$. Use the loss function (8.3.13) and investigate the three tests that reject H_0 if $X < -z_\alpha + \theta_0$ for $\alpha = 0.1, 0.3,$ and 0.5.

(a) (346 : 2 pts, 446 : 1 pts.) For $b = c = 1$, graph and compare their risk functions.
Loss function (8.3.13):

$$L(\theta, a_0) = \begin{cases} 0, & \theta \geq \theta_0 \\ b(\theta_0 - \theta), & \theta < \theta_0 \end{cases} \qquad L(\theta, a_1) = \begin{cases} c(\theta - \theta_0)^2, & \theta \geq \theta_0 \\ 0, & \theta < \theta_0 \end{cases}$$

Let $\phi_\alpha(\theta) = \Phi(-z_\alpha + \theta_0 - \theta)$. Then the risk function is

$$R(\theta) = \begin{cases} (1 - \phi_\alpha(\theta)) \cdot b(\theta_0 - \theta), & \theta < \theta_0 \\ \phi_\alpha(\theta) \cdot c(\theta - \theta_0)^2, & \theta \geq \theta_0 \end{cases}$$

The plot is in Figure 2, labeled as "b=1,c=1"

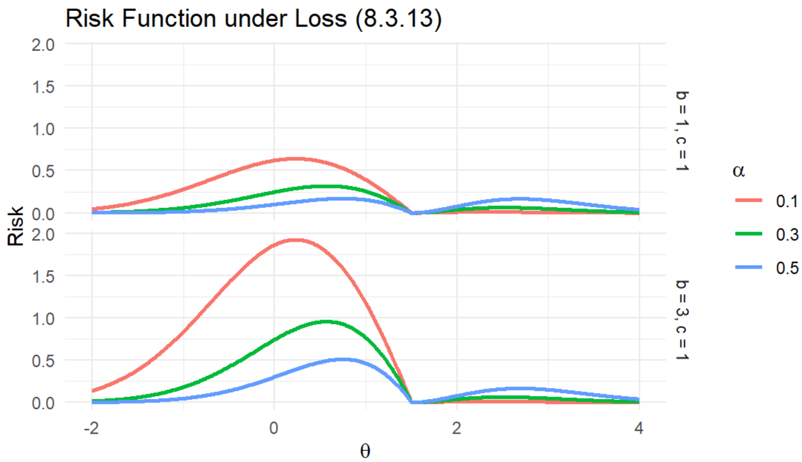

Figure 2: Risk Function

Choosing α affects how risk is distributed:

- Small α: conservative test, low Type I risk, high Type II risk.
- Large α: aggressive test, high Type I risk, low Type II risk.

(b) (346 : 2 pts, 446 : 1 pts.) For $b = 3$, $c = 1$, graph and compare their risk functions.
Use:

$$R(\theta) = \begin{cases} 3(1 - \Phi(-z_\alpha + \theta_0 - \theta))(\theta_0 - \theta), & \theta < \theta_0 \\ \Phi(-z_\alpha + \theta_0 - \theta)(\theta - \theta_0)^2, & \theta \geq \theta_0 \end{cases}$$

The plot is in Figure 2, labeled as "b=3,c=1"
With asymmetric loss ($b = 3,\ c = 1$), increasing α improves power and significantly lowers risk on the left, at the cost of modestly increasing risk on the right.

(c) (346 & 446 : 1 pts.) Graph and compare the power functions of the three tests to the risk functions in parts (a) and (b).
The plot is in Figure 3.
1. Power functions are unaffected by loss

- Power depends on the test procedure (e.g., the critical region defined by α), not the loss function.
- Hence, all curves for different b, c share the same shape and location.

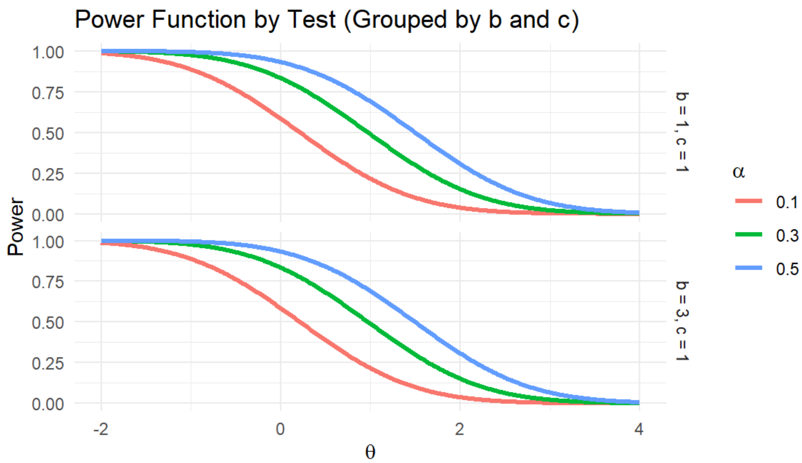

Figure 3: Power Function

2. Risk functions are shaped by the loss
   - With **symmetric loss** ($b = c$), the risk is more balanced.
   - With **asymmetric loss** ($b > c$), the test **prefers avoiding Type II error**, leading to higher risk on the left ($\theta < \theta_0$) and lower on the right.