PSM Fall 2016, Exercise Sheet 5

Return on Thursday October 20 in class

Note. You are encouraged to work in teams of two – but no larger. If you work in a team, submit only a single sheet with both names indicated on it. Nicely type-set solutions are highly appreciated.

Problem 1 (50 points total) Let $(X_1, \ldots, X_n)$ be i.i.d. random variables with known variance σ^2 and unknown mean $\mu \in \mathbb{R}$, which is to be estimated. Assume quadratic loss.

1. (5 points) Formally specify the sample space S, the statistical model $\mathcal{P}$, the decision space D and the loss function L.
   $X_i : (\omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}, \mathcal{P})$ with $\mathcal{P} = \{P \text{ with} E_P[X_i] = \mu, VAR[X_i] = \sigma^2, \mu \in \mathbb{R}\}$, $D = \mathbb{R}$ (our estimate for μ), and $L(\theta, d) = (\mu - d)^2$.
2. (20 points) Determine the risk function for a general procedure of the type

$$t(x_1, \ldots, x_n) = a + b\bar{x}_n, \qquad \text{with } a, b \in \mathbb{R}, \bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

   Hint: Note that $E[\big(\mu - (a + b\bar{X}_n)\big)^2] = E[\big((b - (b-1))\mu - a - b\bar{X}_n\big)^2]$
   For squared error loss $L(\mu, t) = (\mu - t)^2$ and $t(x) = a + b\bar{x}_n$ we get:

$$\begin{aligned}
R_t(\mu) &= E_\mu[(\mu - (a + b\bar{x}_n))^2] \\
&= E_\mu[(b\mu - (b-1)\mu - a - b\bar{x}_n)^2] \\
&= E_\mu[(b(\mu - \bar{x}_n) - ((b-1)\mu + a))^2] \\
&= b^2 E_\mu[(\bar{x}_n - \mu)^2] - ((b-1)\mu + a)^2 + 2E_\mu[(b(\bar{x}_n - \mu)\,((b-1)\mu + a)] \\
&= b^2 VAR[\bar{x}_n] + ((b-1)\mu + a)^2 \\
&= \frac{b^2\sigma^2}{n} + ((b-1)\mu + a)^2 .
\end{aligned}$$

3. (10 points) Derive the specific risk functions for the six procedures $t_{1,n}, t_{2,n}, t_{3,n}, ..., t_{6,n}$ that are defined as follows:

$$\begin{aligned}
t_{1,n} &= \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}_n \\
t_{2,n} &= 2 \cdot \bar{x}_n \\
t_{3,n} &= \frac{\bar{x}_n + \frac{1}{n}}{1 + \frac{1}{n}} \\
t_{4,n} &= -n\bar{x}_n \\
t_{5,n} &= \bar{x}_n + 1 \\
t_{6,n} &= 1.
\end{aligned}$$

$$
\begin{aligned}
R_{t_{1,n}}(\mu) &= \frac{\sigma^2}{n} \\
R_{t_{2,n}}(\mu) &= \frac{4\sigma^2}{n} + \mu^2 \\
R_{t_{3,n}}(\mu) &= \frac{n\sigma^2}{(n+1)^2} + \left(\frac{1-\mu}{n+1}\right)^2 \\
R_{t_{4,n}}(\mu) &= n\sigma^2 + (n+1)^2\mu^2 \\
R_{t_{5,n}}(\mu) &= \frac{\sigma^2}{n} + 1 \\
R_{t_{6,n}}(\mu) &= (-\mu+1)^2
\end{aligned}
$$

4. (5 points) From these calculations can you assert that any of these four procedures is inadmissible? Which ones?
   $t_{2,n}$ and $t_{5,n}$: since both are dominated by $t_{1,n}$.
5. (5 points) If n is large, which of the six procedures would you use, and why?
   For μ very close to 1, $t_{3,n}$ for very large n, otherwise just $t_{1,n}$.
6. (15 points) Show that $t(x_1, \ldots, x_n) = a + b\bar{x}_n$, with $a, b \in \mathbb{R}, \bar{x}_n = \frac{1}{n}\sum x_i$ is inadmissible for
   (a) $b > 1$
   Let $t'(x) = a + b\bar{x}_n$ with $b > 1$. Then

$$
\begin{aligned}
R_{t'}(\mu) &= \frac{b^2\sigma^2}{n} + ((b-1)\mu + a)^2 \\
&> \frac{\sigma^2}{n} \\
&= R_{t_{1,n}}(\mu).
\end{aligned}
$$

   (b) $b < 0$
   Let $t'(x) = a + b\bar{x}_n$ with $b < 0$ and $t_0 = -\frac{a}{b-1}$ a corresponding constant estimator for each choice of a and b. Then

$$
\begin{aligned}
R_{t'}(\mu) &= \frac{b^2\sigma^2}{n} + ((b-1)\mu + a)^2 \\
&\geq ((b-1)\mu + a)^2 \\
&= (b-1)^2\left(\mu + \frac{a}{b-1}\right)^2 \\
&\geq R_{t_0}(\mu).
\end{aligned}
$$

   (c) $b = 1$ and $a \neq 0$.
   Let $t'(x) = a + \bar{x}_n$. Then

$$
\begin{aligned}
R_{t'}(\mu) &= \frac{\sigma^2}{n} + ((1-1)\mu + a)^2 \\
&= \frac{\sigma^2}{n} + a^2 \\
&> \frac{\sigma^2}{n} \\
&= R_{t_{1,n}}(\mu).
\end{aligned}
$$

7. (10 points) Suppose the statistician decides to restrict considerations to procedures $t_{a,b,n} = a + b\bar{x}_n$. He or she is concerned about the behaviour of the risk function when μ is large. Show that the risk function remains limited (and actually approaches 0 for large n) as $|\mu| \to \infty$ if and only if $b = 1$; and that *among procedures* with $b = 1$, the choice $a = 0$ gives uniformly smallest risk function.

   From above we know that:

$$R_t(\mu) = \frac{b^2\sigma^2}{n} + ((b-1)\mu + a)^2 .$$

   For large μ the second part dominates the risk function and hence the risk function only remains limited for $|\mu| \to \infty$ if this part remains limited, i.e. $b = 1$.

   Among procedures with $b = 1$ we then have

$$R_t(\mu) = \frac{\sigma^2}{n} + a^2,$$

   which is obviously minimized for $a = 0$.

Problem 2 (30 points)

Instead of a final exam your stats professor offers you a game. He has two coins, one of which is fair, so head and tail occur with probability $\frac{1}{2}$. The other coin is fabricated and comes up heads with probability $\frac{1}{5}$. Your professor selects one coin and you are allowed to make one single toss. You have to come up with a guess whether the coin is fair or fabricated. Assume zero-one loss.

1. (5 points) Provide a formal description of the sample space, the statistical model, its default parameter space, the decision space and the loss function.

   The setting can be modelled by the sample space $S = \{0, 1\}$ where 0 means 'tail occurs' and 1 that 'head occurs'. The statistical model is described by the set $\mathcal{P} = \{P_{\frac{1}{2}}, P_{\frac{1}{5}}\}$ with $P_\theta(1) = \theta$. The default parameter space is $\Theta = \{\frac{1}{2}, \frac{1}{3}\}$, the decision space $D = \{d_0, d_1\}$ with $d_i = \frac{1}{2+3*i}, i = 0, 1$.

$$L(\theta, t) = \begin{cases} 0 & for\ \theta = t \\ 1 & for\ \theta \neq t. \end{cases}$$

2. (5 points) In this context, compute the general form of the risk function for any statistical procedure t.

   This yields general risk function

$$\begin{aligned} R_t(\theta) = E_\theta[L(\theta, t)] &= 0 \cdot P_\theta(\theta = t) + 1 \cdot P_\theta(\theta \neq t) \\ &= P_\theta(\theta \neq t) \\ &= P_\theta(\text{'wrong decision'}) \\ &= \begin{cases} P_{\frac{1}{2}}(d_1) & for\ \theta = \frac{1}{2} \\ P_{\frac{1}{5}}(d_0) & for\ \theta = \frac{1}{5} \end{cases} \end{aligned}$$

3. (10 points) Student A now comes up with the idea the professor will always use the fair coin, Student B is convinced that he always uses the fabricated

one, Student C and D make their decision dependent on the outcome of the coin toss. So their four statistical procedures are given by

$$
\begin{aligned}
t_A : S \to D, &\qquad x \mapsto d_0, \forall x \in S \\
t_B : S \to D, &\qquad x \mapsto d_1, \forall x \in S \\
t_C : S \to D, &\qquad x \mapsto \begin{cases} d_1, & \text{for } x = 0 \\ d_0, & \text{for } x = 1 \end{cases} \\
t_D : S \to D, &\qquad x \mapsto \begin{cases} d_0, & \text{for } x = 0 \\ d_1, & \text{for } x = 1 \end{cases}
\end{aligned}
$$

Compute the risk functions for these four procedures.

The corresponding risk functions are given by:

$$
\begin{aligned}
R_{t_A}(\theta) &= \begin{cases} 0 & \text{for } \theta = \frac{1}{2} \\ 1 & \text{for } \theta = \frac{1}{5} \end{cases} \\
R_{t_B}(\theta) &= \begin{cases} 1 & \text{for } \theta = \frac{1}{2} \\ 0 & \text{for } \theta = \frac{1}{5} \end{cases} \\
R_{t_C}(\theta) &= \begin{cases} \frac{1}{2} & \text{for } \theta = \frac{1}{2} \\ \frac{1}{5} & \text{for } \theta = \frac{1}{5} \end{cases} \\
R_{t_D}(\theta) &= \begin{cases} \frac{1}{2} & \text{for } \theta = \frac{1}{2} \\ \frac{4}{5} & \text{for } \theta = \frac{1}{5} \end{cases}
\end{aligned}
$$

4. (10 points)
Compute the Bayes risk of the four statistical procedures for the a priori distribution given by density

$$
p_\zeta(\theta) = \begin{cases} 0.4 & \textit{for } \theta = \frac{1}{2} \\ 0.6 & \textit{for } \theta = \frac{1}{5}. \end{cases}
$$

The Bayes risk for the four procedures is then given by

$$
\begin{aligned}
r_{t_A}(\zeta) &= p_\zeta(P_{\frac{1}{2}}) R_{t_A}(\frac{1}{2}) + p_\zeta(P_{\frac{1}{5}}) R_{t_A}(\frac{1}{5}) \\
&= 0.4 \cdot 0 + 0.6 \cdot 1 \\
&= 0.6 \\
r_{t_B}(\zeta) &= p_\zeta(P_{\frac{1}{2}}) R_{t_B}(\frac{1}{2}) + p_\zeta(P_{\frac{1}{5}}) R_{t_B}(\frac{1}{5}) \\
&= 0.4 \cdot 1 + 0.6 \cdot 0 \\
&= 0.4 \\
r_{t_C}(\zeta) &= p_\zeta(P_{\frac{1}{2}}) R_{t_C}(\frac{1}{2}) + p_\zeta(P_{\frac{1}{5}}) R_{t_C}(\frac{1}{5}) \\
&= 0.4 \cdot \frac{1}{2} + 0.6 \cdot \frac{1}{5} \\
&= 0.32 \\
r_{t_D}(\zeta) &= p_\zeta(P_{\frac{1}{2}}) R_{t_D}(\frac{1}{2}) + p_\zeta(P_{\frac{1}{5}}) R_{t_D}(\frac{1}{5}) \\
&= 0.4 \cdot \frac{1}{2} + 0.6 \cdot \frac{4}{5} \\
&= 0.68
\end{aligned}
$$