Econ 240A: Problem Set 6
Solutions to Selected Problems from Chapter 6

Chuan Goh
28 February 2001

1.

a.
The likelihood is just the joint density of the observations, i.e.,

$$f(x; \mu) = (2\pi)^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^{n} (x_i - \mu)^2 \right\}.$$ 

b.
We have \( p(\mu|x) \propto f(x; \mu)p(\mu) \), where \( p(\mu) = (2\pi)^{\frac{k}{2}} \exp\{ -\frac{k\mu^2}{2} \} \). So

\[
p(\mu|x) \propto \exp \left\{ -\frac{n}{2}(\mu - \bar{x})^2 - \frac{k}{2}\mu^2 \right\}
= \exp \left\{ -\frac{1}{2} \left[ (n + k)\mu^2 - 2n\bar{x}\mu + \frac{n^2 \bar{x}^2}{n + k} - \frac{n^2 \bar{x}^2}{n + k} + n\bar{x}^2 \right] \right\}
= \exp \left\{ -\frac{1}{2} \left( \frac{n + k}{\sqrt{n + k}} \mu - \frac{n\bar{x}}{\sqrt{n + k}} \right)^2 \right\}
= \exp \left\{ -\frac{n + k}{2} \left( \mu - \frac{n\bar{x}}{n + k} \right)^2 \right\},
\]
which shows that \( \mu|x \sim N(\frac{n\bar{x}}{n + k}, \frac{1}{n + k}) \).

c.
The Bayes risk \( R(T, \mu) \) of an estimate \( T \) is its expected posterior loss, i.e.,
\( R(T, \mu) = E_{\mu|x} L(T, \mu) \). Here \( L(T, \mu) = (T - \mu)^2 \), so

\[
R(T, \mu) = T^2(x) - 2T(x) \frac{n\bar{x}}{n + k} + \left[ \frac{1}{n + k} + \left( \frac{n\bar{x}}{n + k} \right)^2 \right].
\]
d. 

From part c, it should be fairly clear that the procedure $T$ that minimizes Bayes risk is just the posterior mean of $\mu$, i.e., $T^*(x) = \frac{\sum x_i}{n+k}$. 

2. 

a. 

$$f(x; \mu) = \lambda^n \exp(-\lambda \sum_{i=1}^{n} x_i)$$

b. 

Differentiating $\log f(x; \lambda)$ with respect to $\lambda$ and setting the derivative to zero we find that the maximum likelihood estimator to be $\lambda_{ML} = \frac{\sum x_i}{\sum x_i} = \bar{x}^{-1}$. 

c. 

We have an exponential prior density $\pi(\lambda) = \alpha \exp(-\alpha \lambda)$. Let $t = \sum_{i=1}^{n} x_i$ and $u = \lambda(t + \alpha)$. Note that 

\[
\int_{0}^{\infty} \pi(\lambda) f(x; \lambda) d\lambda = \int_{0}^{\infty} \alpha \lambda^n \exp(-\lambda(t + \alpha)) d\lambda = \alpha \int_{0}^{\infty} \left( \frac{u}{t + \alpha} \right)^n \cdot \frac{1}{t + \alpha} e^{-u} du = \frac{\alpha}{(t + \alpha)^{n+1}} \Gamma(n + 1) = \frac{\alpha n!}{(t + \alpha)^{n+1}}.
\]

Therefore the posterior density of $\lambda$ is given by $\pi(\lambda|x) = \frac{1}{\Gamma(n+1)} \lambda^n (t+\alpha)^{n+1} e^{-\lambda(t+\alpha)}$. The Bayes estimate of $\lambda$, that is, the estimate that minimizes posterior loss is the posterior mean of $\lambda$ if we have a quadratic loss function as in Exercise 1 above: 

\[
E[\lambda|x] = \frac{1}{n!} \int_{0}^{\infty} \lambda^{n+1} e^{-\lambda(t+\alpha)} d\lambda = \frac{(t+\alpha)^{n+1}}{n!} \cdot \frac{1}{(t+\alpha)^{n+2}} \int_{0}^{\infty} u^{n+1} e^{-u} du = \frac{\Gamma(n+2)}{n!(t+\alpha)} = \frac{n+1}{t+\alpha}
\]
d.

We have \( W = 2n \lambda \bar{x} = 2 \lambda \sum_{i=1}^{n} x_i \), so the characteristic function of \( W \) is given by

\[
c_W(s) = E[e^{iWt}] = \prod_{i=1}^{n} e^{i(2\lambda s)} = \prod_{i=1}^{n} \int_{0}^{\infty} \lambda \exp\{i 2 \lambda x_i s - \lambda x_i \} dx_i = \prod_{i=1}^{n} (1 - 2i s)^{-1} = (1 - 2i s)^n = (1 - 2i s) \frac{n!}{(\alpha - 1)^n},
\]

which is the characteristic function of a \( \chi^2_{2n} \) random variable.

We have \( \lambda_{ML} = \frac{2n \lambda}{\nu} \), which is distributed as \( 2n \lambda \chi^{-2}_{\nu} \), where \( \chi^{-2}_{\nu} \) refers to an inverse chi-square distribution with \( \nu \) degrees of freedom.\(^1\) The pdf of the MLE of \( \lambda \) is therefore given by

\[
p(\lambda_{ML}) = \frac{(2n \lambda)^n}{2^{\nu} \Gamma(n)} \frac{\lambda_{ML}^{n-1}}{\chi^{-2}_{\nu} \exp \left\{ - \frac{1}{2} \cdot \frac{2n \lambda}{\lambda_{ML}} \right\} }.
\]

The derivation of \( p(\lambda_{ML}) \) is left as an exercise.

4.

We have \( k_1, \ldots, k_n \) iid Bernoulli(\( p \)), so the likelihood function is given by

\[
f(k_1, \ldots, k_n) = p^{\sum_{i=1}^{n} k_i} \frac{(1 - p)^{n - \sum_{i=1}^{n} k_i}}{n!}.
\]

Let \( K = \sum_{i=1}^{n} k_i \). \( K \) is clearly sufficient for \( p \) by the factorization criterion. It is also minimal sufficient.\(^2\) Note that if there exists a function \( h \) and a statistic \( U \) such that \( K = h(U) \), then \( K \) cannot contain more information about \( p \) than \( U \), which (after a moment’s thought) indicates that \( U \) is also sufficient for \( p \). (Alternatively, sufficiency can be shown by substituting \( h(U) \) for \( K \) in the likelihood function for the sample.)

a.

We have \( U = (k_1, \ldots, k_n) \). Then \( K = \sum_{i=1}^{n} U_i \), so \( U \) is sufficient.

---

\(^1\) Suppose \( X \sim \chi^2_{\nu} \). Then \( X^{-1} \sim \chi^{-2}_{\nu} \).

\(^2\) See handout on sufficiency. The Bernoulli class of distributions can be shown to be a member of the exponential family of distributions.
b. 
We have $U = (k_1^2, (k_2 + \cdots + k_n)^2)$. Then $K = \sqrt{U_1} + \sqrt{U_2}$ so $U$ is sufficient.

c. 
We have $U = \frac{K}{n}$. Then $K = nU$ so $U$ is sufficient.

d. 
We have $U = (\frac{K}{n}, k_2^2 + \cdots + k_n^2)$. Then $K = nU_1$ so $U$ is sufficient.

e. 
We have $U = k_1^2 + \cdots + k_n^2$. Here $K = U$ so $U$ is sufficient.

6. We have an estimator $T(X)$ that has finite variance $V_T$ and is unbiased for $\theta$, so its mean squared error is $V_T$. Denote a Stein shrinkage estimator by $S_\lambda(X) = (1 - \lambda)T(X) + \lambda \theta$. Note that in general $S_\lambda(X)$ will be biased for $\theta$. Denote the mean squared error of $S_\lambda(X)$ by $M(\lambda)$:

$$M(\lambda) = (1 - \lambda)^2 V_T + \lambda^2 (17 - \theta)^2.$$ 

Its derivative is given by

$$M'(\lambda) = -2(1 - \lambda)V_T + 2\lambda(17 - \theta)^2.$$ 

For any value of $\theta$ $M'(\lambda) < 0$ whenever $\lambda < \frac{V_T}{(17 - \theta)^2 + V_T} \leq 1$. Since $S_0(X) = T(X)$, this shows that for $\lambda$ strictly between zero and a small number less than or equal to 1 the MSE of $S_\lambda(X)$ will be uniformly smaller than that of $T(X)$. Whether $S_\lambda(X)$ or $T(X)$ is the better estimator naturally depends on the utility function of the investigator for the particular application at hand.