STAT 543 Homework 5 Solution

1.
Problem 2.2.30
In the multinomial Example 2.2.8, suppose some of the n_j are zero. Show that the MLE of θ_j is $\hat{\theta}$ with $\hat{\theta}_j = \frac{n_j}{n}$, $j = 1, ..., k$.
Hint: Suppose without loss of generality that $n_1 = n_2 = \ldots = n_q = 0, n_{q+1} > 0, ..., n_k > 0$.
Then

$$p(x,\theta) = \prod_{j=q+1}^{k} \theta_j^{n_j}$$

, which vanishes if $\theta_j = 0$ for any $j = q+1, ..., k$.

<u>Solution:</u>
Suppose without loss of generality that

$$n_1 = n_2 = \cdots = n_q = 0, n_{q+1} > 0, ..., n_k > 0.$$

Then,

$$p(x,\theta) = \prod_{j=q+1}^{k} \theta_j^{n_j} \ .$$

Then,

$$l_X(\theta) = \sum_{j=q+1}^{k} n_j \log \theta_j, \quad \theta \in \Theta = \{\theta : \theta_j > 0, \sum_{j=q+1}^{k} \theta_j = 1\}$$

, with

$$\theta_k = 1 - \sum_{j=q+1}^{k-1} \theta_j \tag{1}$$

By setting

$$\frac{\partial}{\partial \theta_j} l_X(\theta) = \frac{\partial}{\partial \theta_j} (\sum_{j=q+1}^{k} n_l \log \theta_l) = \sum_{j=q+1}^{k} \frac{n_l}{\theta_l} \cdot \frac{\partial \theta_l}{\partial \theta_j} = 0, \ j = q+1, ..., k-1$$

$$\Rightarrow \frac{n_j}{\theta_j} - \frac{n_k}{\theta_k} = 0 \Rightarrow \frac{\hat{\theta}_j}{\hat{\theta}_k} = \frac{n_j}{n_k}, \ j = q+1, ..., k-1$$

$$\Rightarrow \sum_{j=q+1}^{k-1} \frac{\hat{\theta}_j}{\hat{\theta}_k} = \frac{n}{n_k} - 1 \tag{2}$$

Then, (1) & (2) $\Rightarrow \hat{\theta}_k = \frac{n_k}{n} \Rightarrow \hat{\theta}_j = \frac{n_j}{n_k} \hat{\theta}_k = \frac{n_j}{n}, \ j = q+1, ..., k$.

Problem 2.2.35

Let $g(x)=\dfrac{1}{\pi(1+x^2)}, x\in\mathbb{R}$, be the Cauchy density, let X_1 and X_2 be *i.i.d.* with density $g(x-\theta),\theta\in\mathbb{R}$. Let x_1 and x_2 be the observations and set $\Delta=\frac{1}{2}(x_1-x_2)$.

Let $\hat\theta=\arg\max L_X(\theta)$ be "the" MLE.

(a) Show that if $|\Delta|\le 1$, and then the MLE exists and is unique. Give the MLE when $|\Delta|\le 1$.

(b) Show that if $|\Delta|>1$, and then the MLE is not unique. Find the values of θ that maximize the likelihood $L_X(\theta)$ when $|\Delta|>1$.

Hint: Factor out $(\bar{x}-\theta)$ in the likelihood equation.

<u>Solution:</u>

Since X_1,X_2 *iid* $\sim g(x-\theta)=\dfrac{1}{\pi\left(1+(x-\theta)^2\right)}$, then

$$L_X(\theta)=\frac{1}{\pi\left(1+(x_1-\theta)^2\right)}\cdot\frac{1}{\pi\left(1+(x_2-\theta)^2\right)}$$

$$l_X(\theta)=-\log\left(1+(x_1-\theta)^2\right)-\log\left(1+(x_2-\theta)^2\right)-2\log(\pi)$$

$$l_X'(\theta)=\frac{2(x_1-\theta)}{1+(x_1-\theta)^2}+\frac{2(x_2-\theta)}{1+(x_2-\theta)^2}=0$$

$$\Rightarrow\quad 2(x_1-\theta)+2(x_1-\theta)(x_2-\theta)^2+2(x_2-\theta)+2(x_2-\theta)(x_1-\theta)^2=0$$

$$2(x_1+x_2-2\theta)+2(x_1-\theta)(x_2-\theta)(x_1+x_2-2\theta)=0$$

$$2\left(1+(x_1-\theta)(x_2-\theta)\right)(x_1+x_2-2\theta)=0 \qquad (1)$$

$$\Rightarrow\quad 4\left\{(\theta-\bar{x})^2-\frac{1}{4}\tilde\Delta\right\}(\bar{x}-\theta)=0$$

Where,

$$\tilde\Delta=(x_1+x_2)^2-4(x_1x_2+1)$$
$$=(x_1-x_2)^2-4$$

a) If $|\Delta|\le 1,$ i.e. $|x_1-x_2|\le 2,$ then $\tilde\Delta=(x_1-x_2)^2-4\le 0$.

$$\begin{cases}\tilde\Delta<0 & \Rightarrow 1+(x_1-\theta)(x_2-\theta)=0 \text{ has no root in } \mathbb{R}\\ \tilde\Delta=0 & \Rightarrow 1+(x_1-\theta)(x_2-\theta)=0 \text{ has only root } \hat\theta=\bar{x}\end{cases} \qquad (2)$$

(1) and (2) $\Rightarrow l_x'(\theta)=0$ has only one root $\hat\theta=\frac{1}{2}(x_1+x_2)=\bar{x}$

Check that

$$l''_X(\theta) = \frac{-2\left[1+(x_1-\theta)^2\right]+4(x_1-\theta)^2}{\left[1+(x_1-\theta)^2\right]^2} + \frac{-2\left[1+(x_2-\theta)^2\right]+4(x_2-\theta)^2}{\left[1+(x_2-\theta)^2\right]^2}$$

$$= \frac{2(x_1-\theta)^2-2}{\left[1+(x_1-\theta)^2\right]^2} + \frac{2(x_2-\theta)^2-2}{\left[1+(x_2-\theta)^2\right]^2}$$

And

$$l''_X(\hat{\theta}) = \frac{2\left(x_1-\hat{\theta}\right)^2-2}{\left[1+\left(x_1-\hat{\theta}\right)^2\right]^2} + \frac{2\left(x_2-\hat{\theta}\right)^2-2}{\left[1+\left(x_2-\hat{\theta}\right)^2\right]^2} = \frac{(x_1-x_2)^2-4}{\left[1+\dfrac{(x_1-x_2)^2}{4}\right]^4} \le 0$$

Hence, $\hat{\theta} = \overline{x}$ is the *MLE* of θ when $|\Delta| \le 1$.

b)

If $|\Delta| > 1$ i.e. $|x_1 - x_2| > 2$, then $\tilde{\Delta} = (x_1-x_2)^2 - 4 > 0$.

$\Rightarrow \quad (\theta - \overline{x})^2 - \frac{1}{4}\tilde{\Delta} = 0$ has 2 roots.

$\hat{\theta}_1 = \overline{x} + \frac{1}{2}\sqrt{\tilde{\Delta}}$, $\hat{\theta}_2 = \overline{x} - \frac{1}{2}\sqrt{\tilde{\Delta}}$

So, $l'_X(\theta) = 0$ has 3 roots.

$\hat{\theta}_1 = \overline{x} + \frac{1}{2}\sqrt{\tilde{\Delta}}$, $\hat{\theta}_2 = \overline{x} - \frac{1}{2}\sqrt{\tilde{\Delta}}$, $\hat{\theta}_3 = \overline{x}$

Then,

$$L_X\left(\hat{\theta}_1\right) = \frac{1}{\pi\left(1+\left(x_1-\hat{\theta}_1\right)^2\right)} \cdot \frac{1}{\pi\left(1+\left(x_2-\hat{\theta}_1\right)^2\right)} = \cdot \frac{1}{\pi^2 T_1(x_1,x_2)}$$

$$T_1(x_1,x_2) = \left(\frac{(x_1-x_2)^2}{4} + 1 + \frac{\tilde{\Delta}}{4}\right)^2 - \frac{1}{4}(x_1-x_2)^2 \cdot \tilde{\Delta}$$

$$= \left(\frac{1}{2}\tilde{\Delta}+2\right)^2 - \frac{1}{4}\tilde{\Delta}^2 - \tilde{\Delta} = \frac{1}{4}\tilde{\Delta}^2 + 2\tilde{\Delta} + 4 - \frac{1}{4}\tilde{\Delta}^2 - \tilde{\Delta} = \tilde{\Delta} + 4$$

$$L_X\left(\hat{\theta}_2\right) = \frac{1}{\pi\left(1+\left(x_1-\hat{\theta}_2\right)^2\right)} \cdot \frac{1}{\pi\left(1+\left(x_2-\hat{\theta}_2\right)^2\right)} = \cdot \frac{1}{\pi^2 T_2(x_1,x_2)}$$

$$T_2(x_1,x_2) = \left(\frac{(x_1-x_2)^2}{4} + 1 + \frac{\tilde{\Delta}}{4}\right)^2 - \frac{1}{4}(x_1-x_2)^2 \cdot \tilde{\Delta}$$

$$= \left(\frac{1}{2}\tilde{\Delta}+2\right)^2 - \frac{1}{4}\tilde{\Delta}^2 - \tilde{\Delta} = \frac{1}{4}\tilde{\Delta}^2 + 2\tilde{\Delta} + 4 - \frac{1}{4}\tilde{\Delta}^2 - \tilde{\Delta} = \tilde{\Delta} + 4$$

$$L_X\left(\hat{\theta}_3\right) = \frac{1}{\pi\left(1+\left(x_1-\hat{\theta}_3\right)^2\right)} \cdot \frac{1}{\pi\left(1+\left(x_2-\hat{\theta}_3\right)^2\right)} = \cdot \frac{1}{\pi^2 T_3\left(x_1,x_2\right)}$$

$$T_3\left(x_1,x_2\right) = \left(1+\frac{\left(x_1-x_2\right)^2}{4}\right)^2 = \left(\frac{5}{4}\tilde{\Delta}+2\right)^2$$

Then,

$$\mathrm{T}_1\left(x_1,x_2\right) = \mathrm{T}_2\left(x_1,x_2\right) < \mathrm{T}_3\left(x_1,x_2\right)$$

$$\Rightarrow L_X\left(\hat{\theta}_1\right) = L_X\left(\hat{\theta}_2\right) > L_X\left(\hat{\theta}_3\right)$$

So, the MLE is not unique, and $\hat{\theta}_{MLE} = \bar{x} + \frac{1}{2}\sqrt{\tilde{\Delta}}$ and $\bar{x} - \frac{1}{2}\sqrt{\tilde{\Delta}}$.

Problem 2.2.39

Let X_i denote the number of hits at a certain Web site on day $i, i = 1,...,n$. Assume that $S = \sum_{i=1}^{n} X_i$ has a Poisson, $\mathrm{P}(n\lambda)$, distribution. On day $n+1$ the Web Master decides to keep track of two types of hits (money making and not money making). Let V_j and W_j denote the number of hits of type 1 and 2 on day j, $j = n+1,...,n+m$. Assume that $S_1 = \sum_{j=n+1}^{n+m} V_j$ and $S_2 = \sum_{j=n+1}^{n+m} W_j$ have $\mathrm{P}(m\lambda_1)$ and $\mathrm{P}(m\lambda_2)$ distributions, where $\lambda_1 + \lambda_2 = \lambda$. Also assume that $S, S_1,$ and S_2 are independent. Find the MLEs of λ_1 and λ_2 based on $S, S_1,$ and S_2 .

<u>Solution:</u>

$$s = \sum_{i=1}^{n} x_i \sim P(n\lambda)$$

$$s_1 = \sum_{j=n+1}^{n+m} v_j \sim P\left(m\lambda_1\right) \qquad s_2 = \sum_{j=n+1}^{n+m} w_j \sim P\left(m\lambda_2\right)$$

$$L_n(\lambda) = \frac{e^{-n\lambda}(n\lambda)^s}{s!} \cdot \frac{e^{-m\lambda_1}\left(m\lambda_1\right)^{s_1}}{s_1!} \cdot \frac{e^{-m\lambda_2}\left(m\lambda_2\right)^{s_2}}{s_2!}$$

$$= \frac{e^{-(n+m)\lambda} \cdot (n\lambda)^s \cdot \left(m\lambda_1\right)^{s_1} \cdot \left(m\lambda_2\right)^{s_2}}{s!\,s_1!\,s_2!}$$

$$l_n(\lambda) = -(n+m)\lambda + s\log(n\lambda) + s_1\log\left(m\lambda_1\right) + s_2\log\left(m\lambda_2\right) - \log(s!) - \log\left(s_1!\right) - \log\left(s_2!\right)$$

$$\frac{\partial l_n}{\partial \lambda_1} = -(n+m) + \frac{s}{\lambda_1+\lambda_2} + \frac{s_1}{\lambda_1} = 0$$

$$\frac{\partial l_n}{\partial \lambda_2} = -(n+m) + \frac{s}{\lambda_1+\lambda_2} + \frac{s_2}{\lambda_2} = 0$$

$$Let\,\frac{s_1}{\lambda_1}=k=\frac{s_2}{\lambda_2} \qquad \Rightarrow s_1=\lambda_1 k,\quad s_2=\lambda_2 k$$

$$\Rightarrow \lambda_1+\lambda_2=\frac{s_1+s_2}{k}$$

$$\Rightarrow -(n+m)+\frac{ks}{s_1+s_2}=-k \qquad \Rightarrow \frac{s+s_1+s_2}{s_1+s_2}k=n+m$$

$$\Rightarrow k=\frac{(s_1+s_2)(n+m)}{s+s_1+s_2}$$

$$\Rightarrow \hat{\lambda}_1=\frac{s_1(s+s_1+s_2)}{(s_1+s_2)(n+m)}$$

$$\hat{\lambda}_2=\frac{s_2(s+s_1+s_2)}{(s_1+s_2)(n+m)}$$

Check that

$$\begin{pmatrix}\frac{\partial^2 l_n}{\partial\lambda_1^2} & \frac{\partial^2 l_n}{\partial\lambda_1\partial\lambda_2}\\ \frac{\partial^2 l_n}{\partial\lambda_1\partial\lambda_2} & \frac{\partial^2 l_n}{\partial\lambda_2^2}\end{pmatrix}=\begin{pmatrix}-\frac{s}{(\lambda_1+\lambda_2)^2}-\frac{s_1}{\lambda_1^2} & -\frac{s}{(\lambda_1+\lambda_2)^2}\\ -\frac{s}{(\lambda_1+\lambda_2)^2} & -\frac{s}{(\lambda_1+\lambda_2)^2}-\frac{s_2}{\lambda_2^2}\end{pmatrix}$$

, satisfies:

(1)

$$\frac{\partial^2 l_n}{\partial\lambda_1^2}=-\frac{s}{(\lambda_1+\lambda_2)^2}-\frac{s_1}{\lambda_1^2}<0$$

$$\frac{\partial^2 l_n}{\partial\lambda_2^2}=-\frac{s}{(\lambda_1+\lambda_2)^2}-\frac{s_2}{\lambda_2^2}<0$$

(2)

$$\begin{vmatrix}\frac{\partial^2 l_n}{\partial\lambda_1^2} & \frac{\partial^2 l_n}{\partial\lambda_1\partial\lambda_2}\\ \frac{\partial^2 l_n}{\partial\lambda_1\partial\lambda_2} & \frac{\partial^2 l_n}{\partial\lambda_2^2}\end{vmatrix}=\begin{vmatrix}-\frac{s}{(\lambda_1+\lambda_2)^2}-\frac{s_1}{\lambda_1^2} & -\frac{s}{(\lambda_1+\lambda_2)^2}\\ -\frac{s}{(\lambda_1+\lambda_2)^2} & -\frac{s}{(\lambda_1+\lambda_2)^2}-\frac{s_2}{\lambda_2^2}\end{vmatrix}$$

$$=\left(\frac{s}{(\lambda_1+\lambda_2)^2}+\frac{s_1}{\lambda_1^2}\right)\cdot\left(\frac{s}{(\lambda_1+\lambda_2)^2}+\frac{s_2}{\lambda_2^2}\right)-\frac{s^2}{(\lambda_1+\lambda_2)^4}$$

$$=\frac{s_1}{\lambda_1^2}\cdot\frac{s}{(\lambda_1+\lambda_2)^2}+\frac{s_2}{\lambda_2^2}\cdot\frac{s}{(\lambda_1+\lambda_2)^2}+\frac{s_1 s_2}{\lambda_1^2\lambda_2^2}>0$$

Problem 2

As in the first problem on Assignment 1, suppose that $X_1, X_2, \ldots, X_{10}$ are $i.i.d.$ exponential with mean θ. That is, suppose that the marginal pdf is

$$f(x\mid\theta) = I[x \ge 0]\frac{1}{\theta}\exp(-\frac{x}{\theta})$$

Suppose further that what is observed is not the X_i, but rather $Y_1, Y_2, \ldots, Y_{10}$ for

$Y_i = X_i$ rounded to the nearest integer

Write out an E-M algorithm for finding the MLE of θ based on the available data.

<u>Solution:</u>

Let $y_i \equiv \left[x_i\right] \equiv$ nearest integer number to x_i

$$L(\theta \mid \underline{x}) = f(\underline{x} \mid \theta) = I(\min X(i) \ge 0)\frac{1}{\theta^n}\exp\left(-\frac{\sum x_i}{\theta}\right)$$

$$l(\theta) = -n\log(\theta) - \frac{\sum x_i}{\theta} + \log\left(I(\min X(i) \ge 0)\right)$$

Let $\dfrac{\partial l(\theta)}{\partial\theta} = -\dfrac{n}{\theta} + \dfrac{\Sigma x_i}{\theta^2} = 0$

$$\Rightarrow \hat{\theta} = \frac{\Sigma x_i}{n}$$

(E) Step:

$$E_{\theta^{(r)}}(x_i \mid y_i) = \frac{\displaystyle\int_{\min(y_i-0.5,0)}^{y_i+0.5} x\frac{1}{\theta^{(r)}}\exp\left(-\frac{x}{\theta^{(r)}}\right)dx}{\displaystyle\int_{\min(y_i-0.5,0)}^{y_i+0.5} \frac{1}{\theta^{(r)}}\exp\left(-\frac{x}{\theta^{(r)}}\right)dx}$$

(M) Setp:

$$\theta^{(r+1)} = \frac{\sum_{i=1}^{n} E_{\hat{\theta}^{(r)}}(x_i \mid y_i)}{n}$$

So the EM Algorithm is that choose initial value θ_0 first, and then by repeating (E) step and (M) step, we get a sequence of estimates $\theta^{(r)}$, $r = 1, 2, \cdots$ until $\theta^{(r)}$ satisfies some convergence criterion, like $\left|\theta^{(r)} - \theta^{(r+1)}\right| < 10^{-10}$.

Problem 3

Write out an *E-M* algorithm for finding the *MLE* of ***p*** based on all data collected in the three studies mentioned in the 3rd problem of Assignment 1.

<u>Solution:</u>

Let

$n_1 =$ the # coutcome ++, $n_{1i} =$ # coutcome ++ in study i, $i = 1,2,3$

$n_2 =$ the # coutcome +- $n_{2i} =$ # coutcome +- in study i, $i = 1,2,3$

$n_3 =$ the # coutcome -+ $n_{3i} =$ # coutcome -+ in study i, $i = 1,2,3$

$n_4 =$ the # coutcome -- $n_{4i} =$ # coutcome -- in study i, $i = 1,2,3$

$n_1 = \sum n_{1i}, n_2 = \sum n_{2i}, n_3 = \sum n_{3i}, n_4 = \sum n_{4i}$

Let $p = (p_{++}, p_{+-}, p_{-+}, p_{--})$, then from example 2.2.8, and by *pdf* of multinomial,

$$\hat{p}_{++} = \frac{n_1}{n},\ \hat{p}_{+-} = \frac{n_2}{n},\ \hat{p}_{-+} = \frac{n_3}{n},\ \hat{p}_{--} = \frac{n_4}{n}.$$

Observations we have:

study 1: $n_{11} + n_{21},\ \ n_{31} + n_{41}$

study 2: $n_{12} + n_{22},\ \ n_{32} + n_{42}$

study 3: $n_{13} + n_{23},\ \ n_{33} + n_{43}$

(E) Step

For study 1:

$$E_p^{(r)}\left(n_{11} \mid s(n) = n_{11} + n_{21}\right) = (n_{11} + n_{21})\frac{p_{++}^{(r)}}{p_{++}^{(r)} + p_{+-}^{(r)}}$$

$$E_p^{(r)}\left(n_{21} \mid s(n) = n_{11} + n_{21}\right) = (n_{11} + n_{21})\frac{p_{+-}^{(r)}}{p_{++}^{(r)} + p_{+-}^{(r)}}$$

$$E_p^{(r)}\left(n_{31} \mid s(n) = n_{31} + n_{41}\right) = (n_{31} + n_{41})\frac{p_{-+}^{(r)}}{p_{-+}^{(r)} + p_{--}^{(r)}}$$

$$E_p^{(r)}\left(n_{41} \mid s(n) = n_{31} + n_{41}\right) = (n_{31} + n_{41})\frac{p_{--}^{(r)}}{p_{-+}^{(r)} + p_{--}^{(r)}}$$

For study 2:

$$E_p^{(r)}\left(n_{12} \mid s(n) = n_{12} + n_{32}\right) = (n_{12} + n_{32})\frac{p_{++}^{(r)}}{p_{++}^{(r)} + p_{-+}^{(r)}}$$

$$E_p^{(r)}\left(n_{32} \mid s(n) = n_{12} + n_{32}\right) = (n_{12} + n_{32})\frac{p_{-+}^{(r)}}{p_{++}^{(r)} + p_{-+}^{(r)}}$$

$$E_p^{(r)}\left(n_{22} \mid s(n) = n_{22} + n_{42}\right) = (n_{22} + n_{42})\frac{p_{+-}^{(r)}}{p_{+-}^{(r)} + p_{--}^{(r)}}$$

$$E_p^{(r)}\left(n_{42} \mid s(n) = n_{22} + n_{42}\right) = (n_{22} + n_{42})\frac{p_{--}^{(r)}}{p_{+-}^{(r)} + p_{--}^{(r)}}$$

(M) Step:

$$\hat{p}_{++}^{(r+1)} = \frac{E_p^{(r)}(n_1)}{n}, \qquad \hat{p}_{+-}^{(r+1)} = \frac{E_p^{(r)}(n_2)}{n}$$

$$\hat{p}_{-+}^{(r+1)} = \frac{E_p^{(r)}(n_3)}{n}, \qquad \hat{p}_{--}^{(r+1)} = \frac{E_p^{(r)}(n_4)}{n}$$

So, the EM Algorithm is that choose an initial value $p^{(0)}$ first, and then by repeating (E) step and (M) step, we get a sequence of $p^{(r)}, \quad r = 1,2,\ldots$ until $p^{(r)}$ satisfies some convergence criterion, like $\left|p^{(r)} - p^{(r+1)}\right| < 10^{-6}$.

4.
Problem 3.2.1
Show that if $X_1,...,X_n$ is a $\mathrm{N}(\theta,\sigma^2)$ sample and π is the improper prior $\pi(\theta)=1, \theta \in \Theta = \mathbb{R}$, then the improper Bayes rule for squared error loss is $\delta^*(X) = \overline{x}$.

Solution:
Given $X_1,...,X_n \sim i.i.d. N(\theta,\sigma^2)$, and the improper prior $\pi(\theta)=1, \theta \in \Theta = \mathbb{R}$, then

$$\pi(\theta \mid x_1,...,x_n) = \frac{f(x_1,...,x_n \mid \theta)\pi(\theta)}{\int_0^1 f(x_1,...,x_n \mid \theta)\pi(\theta)d\theta}$$

$$= \frac{(2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2}\sum_i (x_i-\theta)^2\}}{\int_0^\infty (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2}\sum_i (x_i-\theta)^2\}d\theta}$$

$$= \frac{(2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2}\sum_i (x_i-\overline{x})^2 + n(\overline{x}-\theta)^2\}}{\int_0^\infty (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2}\sum_i (x_i-\overline{x})^2 + n(\overline{x}-\theta)^2\}d\theta}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\{-\frac{(\theta-\overline{x})^2}{2\sigma^2/n}\} \quad (\text{By } \overline{x} \sim N(\theta, \frac{\sigma^2}{n}))$$

So,

$$\theta \mid x_1,...,x_n \sim N(\overline{x}, \frac{\sigma^2}{n})$$

, and then the improper Bayes rule for squared error loss is

$$\delta^*(\underline{x}) = E(\theta \mid x_1,...,x_n) = \overline{x} .$$

Problem 3.2.2

Let $X_1,...,X_n$ be the indicators of n Bernoulli trials with success probability θ. Suppose $l(\theta,a)$ is the quadratic loss $(\theta-a)^2$ and that the prior $\pi(\theta)$ is the beta, $\beta(r,s)$, density. Find the Bayes estimated $\hat{\theta}_B$ of θ and write it as a weighted average $\omega\theta_0+(1-\omega)\bar{X}$ of the mean θ_0 of the prior and the sample mean $\bar{X}=\frac{S}{n}$. Show that $\hat{\theta}_B=\frac{S+1}{n+2}$ for the uniform prior.

<u>Solution:</u>

Given $X_1,...,X_n \sim i.i.d.Bernoulli(\theta)$, then

$$f(x_1,...,x_n\mid\theta)=\theta^{\sum_i x_i}(1-\theta)^{n-\sum_i x_i}, 0<\theta<1.$$

And for the prior $\pi(\theta)\sim Beta(r,s)$,

$$\pi(\theta\mid x_1,...,x_n)=\frac{f(x_1,...,x_n\mid\theta)\pi(\theta)}{\int_0^1 f(x_1,...,x_n\mid\theta)\pi(\theta)d\theta}$$

$$=\frac{\theta^{\sum_i x_i}(1-\theta)^{n-\sum_i x_i}\frac{1}{B(r,s)}\theta^{r-1}(1-\theta)^{s-1}}{\int_0^1\theta^{\sum_i x_i}(1-\theta)^{n-\sum_i x_i}\frac{1}{B(r,s)}\theta^{r-1}(1-\theta)^{s-1}d\theta}$$

$$=\frac{1}{B(\sum_i x_i+r,n+s-\sum_i x_i)}\theta^{\sum_i x_i+r-1}(1-\theta)^{n+s-\sum_i x_i-1}$$

So,

$$\theta\mid x_1,...,x_n\sim Beta(\sum_i x_i+r,n+s-\sum_i x_i)$$

Then under the quadratic loss, the Bayes rule is

$$\delta^*(\underline{x})=\hat{\theta}_B=E(\theta\mid\underline{x})=\frac{\sum_i x_i+r}{n+r+s}=\frac{S+r}{n+r+s}$$

Let

$$\theta_0=\frac{r}{r+s}, \quad \text{and} \quad \omega=\frac{r+s}{n+r+s}$$

, then

$$\delta^*(\underline{x})=\omega\theta_0+(1-\omega)\bar{x}.$$

Especially, for uniform prior, since uniform is actually $Beta(1,1)$, then

$$\delta^*(\underline{x})=\hat{\theta}_B=E(\theta\mid\underline{x})=\frac{\sum_i x_i+r}{n+r+s}=\frac{S+1}{n+2}$$

Problem 3.2.3
In Problem 3.2.2 preceeding, five the MLE of the Bernoulli variance $q(\theta)=\theta(1-\theta)$ and five the Bayes estimate of $q(\theta)$. Check whether $q(\hat{\theta}_B)=E(q(\theta)\mid X)$, where $\hat{\theta}_B$ is the Bayes estimate of θ.

Solution:

$$\hat{\theta}_{MLE}=\overline{x}$$

$$E\left(q(\theta)\mid \underline{x}\right)=\widehat{q(\theta)}_{MLE}=\hat{\theta}_{MLE}\left(1-\hat{\theta}_{MLE}\right)=\overline{x}\left(1-\overline{x}\right)$$

$$q\left(\hat{\theta}_B\right)=\hat{\theta}_B\left(1-\hat{\theta}_B\right) \qquad \hat{\theta}_B=E\left(\theta\mid x\right)$$

Since

$$\theta\mid\underline{x}\sim \mathrm{Beta}\left(\sum x_i+r, n-s-\sum x_i\right)$$

$$E\left(\theta\mid\underline{x}\right)=\hat{\theta}_B=\frac{\sum x_i+r}{n+s+r}$$

$$E\left(\theta^2\mid\underline{x}\right)=\frac{\left(\sum x_i+r\right)\left(\sum x_i+r+1\right)}{(n+s+r)\left(n+s+r+1\right)}$$

$$\begin{aligned}E\left(q(\theta)\mid\underline{x}\right)&=E\left(\theta(1-\theta)\mid\underline{x}\right)\\&=E\left(\theta\mid\underline{x}\right)-E\left(\theta^2\mid\underline{x}\right)\\&=\frac{\left(\sum x_i+r\right)}{n+s+r}-\frac{\left(\sum x_i+r\right)\left(\sum x_i+r+1\right)}{(n+s+r)\left(n+r+s+1\right)}\\&=\frac{\left(\sum x_i+r\right)\left(n+s-\sum x_i\right)}{(n+s+r)\left(n+r+s+1\right)}\end{aligned}$$

$$\begin{aligned}q\left(\hat{\theta}_B\right)&=\left(\frac{\sum x_i+r}{n+s+r}\right)\left(1-\frac{\sum x_i+r}{n+s+r}\right)\\&=\frac{\left(\sum x_i+r\right)\left(n+s-\sum x_i\right)}{(n+s+r)^2}\neq E\left(q(\theta)\mid\underline{x}\right)\end{aligned}$$

Problem 3.2.4
In the Bernoulli Problem 3.2.2 with uniform prior on the probability of success θ, we found that $\frac{S+1}{n+2}$ is the Bayes rule. In some studies (see Section 6.4.3), the parameter $\lambda=\frac{\theta}{1-\theta}$, which is called the *odds ratio* (for success), is preferred to θ. If we put a (improper) uniform prior on λ, under what condition on S does the Bayes rule exist and what is the Bayes rule?

<u>Solution:</u>

By $f_Y(y) = f_X(x)\left|\frac{\partial X}{\partial Y}\right|$, we have

$$\lambda = \frac{\theta}{1-\theta} \ \& \ \pi(\lambda) = 1 \Rightarrow \pi(\theta) = \frac{1}{(1-\theta)^2}, 0 < \theta < 1$$

$$f(x_1,\dots x_n \mid \theta)\pi(\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}\frac{1}{(1-\theta)^2} = \theta^{\sum x_i}(1-\theta)^{n-2-\sum x_i}$$

$$f(\theta \mid x_1,\dots x_n)\pi(\theta) = \frac{f(x_1,\dots x_n \mid \theta)\pi(\theta)}{\int f(x_1,\dots x_n \mid \theta)\pi(\theta)\,d\theta}$$

$$= \frac{\theta^{\sum x_i}(1-\theta)^{n-2-\sum x_i}}{\int \theta^{\sum x_i}(1-\theta)^{n-2-\sum x_i}\,d\theta} = \frac{1}{\beta\left(\sum x_i + 1, n-1-\sum x_i\right)}\cdot\theta^{\sum x_i}(1-\theta)^{n-2-\sum x_i}$$

$$\Rightarrow \theta \mid x_1,\dots x_n \sim \beta\left(\sum x_i + 1, n-1-\sum x_i\right)$$

$$\hat{\theta}_B = E(\theta \mid x_1,\dots x_n) = \frac{s+1}{n}$$

That requires $\begin{cases} \sum x_i > 0 \\ \sum x_i + 1 > 0 \quad \Rightarrow 0 < s < n-1, s = \sum_{i=1}^{n} x_i \\ n-1-\sum x_i > 0 \end{cases}$