Solutions for Assignment #1 : Topics in Mathematical Statistics

Daeyoung Lim
Dept. of Statistics, Korea University

Problem 1

Suppose that k events $B_1, \ldots, B_k$ form a partition of the sample space Ω. For $i = 1, \ldots, k$, let $P(B_i)$ denote the prior probability of B_i. Also, for each event A such that $P(A) > 0$, let $P(B_i|A)$ denote the posterior probability of B_i given that the event A has occurred. Prove that if $P(B_1|A) < P(B_1)$, then $P(B_i|A) > P(B_i)$ for at least one value of i ($i = 2, \ldots, k$).

Recall that

$$\sum_{i=1}^{k} P(B_i) = 1 \quad \text{and} \quad \sum_{i=1}^{k} P(B_i \mid A) = 1 \tag{1}$$

(Proof by contradiction) Suppose that $P(B_i \mid A) < P(B_i), \forall i = 1, \ldots, k$. Then

$$\sum_{i=1}^{k} P(B_i \mid A) < \sum_{i=1}^{k} P(B_i) \qquad (= 1 < 1) \tag{2}$$

Thus, contradiction.

Problem 2

Suppose that a box contains five coins, and that for each coin there is a different probability that a head will be obtained when the coin is tossed. Let p_i denote the probability of a head when the ith coin is tossed ($i = 1, \ldots, 5$), and suppose that $p_1 = 0$, $p_2 = 1/4$, $p_3 = 1/2$, $p_4 = 3/4$, and $p_5 = 1$.

(a) Suppose that one coin is selected at random from the box and when it is tossed once, a head is obtained. What is the posterior probability that the ith coin was selected ($i = 1, \ldots, 5$)?

Let H be the event of obtaining a head and S_i the event of selecting the ith coin. Then

$$P(S_i \mid H) = \frac{P(H \mid S_i)P(S_i)}{\sum_{i=1}^{5} P(H \mid S_i)P(S_i)} = \frac{2}{5}p_i \quad \left(\because p_i = P(H \mid S_i) \text{ and } P(S_i) = \frac{1}{5}\right) \tag{3}$$

(b) If the same coin were tossed again, what would be the probability of obtaining another head?

$$P(H \mid S_i)P(S_i \mid H) = p_i \cdot \frac{2}{5}p_i = \frac{2}{5}p_i^2 \quad (\because P(H \mid S_i, H) = P(H \mid S_i) \text{ by independent trials}) \tag{4}$$

(c) If a tail had been obtained on the first toss of the selected coin and the same coin were tossed again, what would be the probability of obtaining a head on the second toss?

$$P(H \mid S_i)P(S_i \mid H^c) = p_i \cdot \frac{P(H^c \mid S_i)P(S_i)}{\sum_{i=1}^{6} P(H^c \mid S_i)P(S_i)} = \frac{2}{5}p_i(1 - p_i) \tag{5}$$

Remark. *(b) and (c) are related to* ***the posterior predictive probability*** *where we incorporate the previously observed information.*

Problem 3

The *skewness* of a random variable X can be defined as $\gamma_1 = \mu_3/(\mu_2)^{\frac{3}{2}}$ where

$$\mu_n = \mathrm{E}(X - \mathrm{E}(X))^n$$

Find the skewness of a random variable X with a binomial distribution $B(n, \pi)$ of index n and parameter π.

(Method 1: Direct calculation)
Expanding,

$$\begin{aligned} \mathrm{E}\left\{(X - \mathrm{E}(X))^3\right\} &= \mathrm{E}\left(X^3 - 3n\pi X^2 + 3(n\pi)^2 X - (n\pi)^3\right) \\ &= \mathrm{E}(X^3) - 3n\pi\,\mathrm{E}(X^2) + 3n^2\pi^2\,\mathrm{E}(X) - n^3\pi^3 \end{aligned} \tag{6}$$

Since $\mathrm{E}(X) = n\pi$ and $\mathrm{E}(X^2) = n\pi(1 - \pi + n\pi)$, $\mathrm{E}(X^3)$ should be computed through MGF (or PGF).

$$M_X(t) = \mathrm{E}(e^{tX}) = \sum_{x=0}^{n} e^{tx}\binom{n}{x}\pi^x(1-\pi)^{n-x} = (1 - \pi + e^t\pi)^n \quad (\because \text{Binominal theorem}) \tag{7}$$

Then

$$M_X^{(3)}(t) = \pi n e^t(\pi(e^t - 1) + 1)^{n-3}(\pi^2(n^2e^{2t} + (1-3n)e^t + 1) + \pi((3n-1)e^t - 2) + 1) \tag{8}$$

and

$$\mathrm{E}(X^3) = M_X^{(3)}(0) = \pi n(\pi^2(n-2)(n-1) + 3\pi(n-1) + 1) \tag{9}$$

which yields

$$\mu_3 = n\pi(1-\pi)(1-2\pi) \tag{10}$$

Therefore,

$$\gamma_1 = \frac{1 - 2\pi}{\sqrt{n\pi(1-\pi)}} \tag{11}$$

(Method 2: Change of variable)
Consider $X \overset{D}{=} \sum_{i=1}^{n} Y_i$ where $Y_i \sim \text{Bernoulli}(\pi)$. Thus

$$X - n\pi \overset{D}{=} \sum_{i=1}^{n} Y_i - n\pi = \sum_{i=1}^{n}(Y_i - \pi) \tag{12}$$

If we again let $Z_i = Y_i - \pi$,

$$X - n\pi \overset{D}{=} \sum_{i=1}^{n} Z_i, \quad \text{where } Z_i = \begin{cases} 1 - \pi, & \text{with } \pi \\ -\pi, & \text{with } 1 - \pi \end{cases} \tag{13}$$

Thus,

$$\begin{aligned} (X - n\pi)^3 &= \left(\sum_{i=1}^{n} Z_i\right)^3 = \sum_{i=1}^{n} Z_i^3 + \sum_{i \neq j} Z_i^2 Z_j + \sum_{i \neq j \neq k} Z_i Z_j Z_k \\ \mathrm{E}\left\{(X - n\pi)^3\right\} &= \sum_{i=1}^{n} \mathrm{E}(Z_i^3) = n\,\mathrm{E}(Z_1^3) \qquad (\because Z_i \perp\!\!\!\perp Z_j \ \forall i \neq j \ \text{ and } \ \mathrm{E}(Z_i) = 0) \\ &= n\pi(1-\pi)(1-2\pi) \implies \gamma_1 = \frac{1-2\pi}{\sqrt{n\pi(1-\pi)}} \end{aligned} \tag{14}$$

Problem 4

Define

$$I = \int_0^\infty \exp\left(-\frac{1}{2}z^2\right) \mathrm{d}z$$

and show that

$$I = \int_0^\infty \exp\left(-\frac{1}{2}(xy)^2\right) y \,\mathrm{d}x = \int_0^\infty \exp\left(-\frac{1}{2}(zx)^2\right) z \,\mathrm{d}x.$$

Deduce that

$$I^2 = \int_0^\infty \int_0^\infty \exp\left\{-\frac{1}{2}(x^2+1)z^2\right\} z \,\mathrm{d}z \,\mathrm{d}x.$$

and show that $I = \sqrt{\pi/2}$. (This method is due to Laplace, 1812, Section 24.)

Using the transformation $z = yx$,

$$\mathrm{d}z = y\,\mathrm{d}x \implies \int_0^\infty \exp\left(-\frac{1}{2}z^2\right) \mathrm{d}z = \int_0^\infty \exp\left(-\frac{1}{2}(xy)^2\right) y\,\mathrm{d}x \tag{15}$$

Since y is just a dummy variable, we can replace it with z:

$$I = \int_0^\infty \exp\left(-\frac{1}{2}(zx)^2\right) z\,\mathrm{d}x \tag{16}$$

Then

$$\begin{aligned} I^2 &= \left(\int_0^\infty \exp\left(-\frac{1}{2}z^2\right) \mathrm{d}z\right)\left(\int_0^\infty \exp\left(-\frac{1}{2}(zx)^2\right) z\mathrm{d}x\right) \\ &= \int_0^\infty \int_0^\infty \exp\left(-\frac{1}{2}(x^2+1)z^2\right) z\,\mathrm{d}z\,\mathrm{d}x \end{aligned} \tag{17}$$

To evaluate the double integral, let

$$\begin{gathered} u = -\frac{1}{2}(x^2+1)z^2 \implies \mathrm{d}u = -(x^2+1)z\,\mathrm{d}z \\ \int_0^\infty \exp\left(-\frac{1}{2}(x^2+1)z^2\right) z\,\mathrm{d}z = -\frac{1}{x^2+1}\int_0^{-\infty} e^u\,\mathrm{d}u = \frac{1}{x^2+1} \end{gathered} \tag{18}$$

Thus,

$$I^2 = \int_0^\infty \frac{1}{x^2+1}\,\mathrm{d}x = \arctan x\Big|_0^\infty = \frac{\pi}{2} \implies I = \sqrt{\frac{\pi}{2}} \tag{19}$$

Problem 5

Let X_1, X_2 be two independent random variables each with p.d.f. $f_1(x) = e^{-x}$ for $x > 0$ and $f_1(x) = 0$ for $x \leq 0$. Let $Z = X_1 - X_2$ and $W = X_1/X_2$.
(Notice that $\{Z = 0\} = \{W = 1\}$, but the conditional distribution of X_1 given $Z = 0$ is not the same as the conditional distribution of given $W = 1$. This discrepancy is known as the *Borel paradox* : The conditional p.d.f.'s are not like conditioning on events of probability 0. We can see that "Z is very close to 0" is not the same as "W is very close to 1".)
(a) Prove that the conditional p.d.f of X_1 given $Z = 0$ is

$$g_1(x_1|0) = \begin{cases} 2e^{-2x_1} & x_1 > 0, \\ 0 & \text{otherwise.} \end{cases}$$

By the change of variable,

$$\begin{aligned} &\begin{cases} X_1 = X_1 \\ X_2 = X_1 - Z \end{cases} \implies |J| = \left|\det\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}\right| = 1 \\ &f_{X_1,Z}(x_1, z) = e^{-2x_1+z} \end{aligned} \tag{20}$$

For Z,

$$\begin{aligned} P(Z \leq z) &= P(X_1 - X_2 \leq z) = P(X_2 \geq X_1 - z) \\ &= \begin{cases} \displaystyle\int_0^\infty \int_{x_1-z}^\infty e^{-x_1-x_2}\,\mathrm{d}x_2\,\mathrm{d}x_1, & \text{if } z \leq 0 \\ \displaystyle 1 - \int_z^\infty \int_0^{x_1-z} e^{-x_1-x_2}\,\mathrm{d}x_2\,\mathrm{d}x_1, & \text{if } z > 0 \end{cases} \\ &= \begin{cases} \dfrac{e^z}{2}, & \text{if } z \leq 0 \\ 1 - \dfrac{e^{-z}}{2}, & \text{if } z > 0 \end{cases} \\ f_Z(z) &= \frac{1}{2}e^{-|z|}, \quad -\infty < z < \infty \end{aligned} \tag{21}$$

Thus,

$$g_1(x_1 \mid z) = 2e^{-2x_1+z+|z|} \implies g_1(x_1 \mid 0) = 2e^{-2x_1}, \quad x_1 > 0 \tag{22}$$

Problem 5 continued on next page. . .

(b) Prove that the conditional p.d.f. of X_1 given $W = 1$ is

$$h_1(x_1|1) = \begin{cases} 4x_1e^{-2x_1} & x_1 > 0, \\ 0 & \text{otherwise.} \end{cases}$$

By the change of variable,

$$\begin{aligned} &\begin{cases} X_1 = X_1 \\ X_2 = \dfrac{X_1}{W} \end{cases} \implies |J| = \left|\det\begin{pmatrix} 1 & 0 \\ 1/w & -x_1/w^2 \end{pmatrix}\right| = \frac{x_1}{w^2} \\ &f_{X_1,W}(x_1, w) = e^{-(1+1/w)x_1}\frac{x_1}{w^2} \end{aligned} \tag{23}$$

and

$$P\left(\frac{X_1}{X_2} \le w\right) = P\left(X_2 \ge \frac{X_1}{w}\right) = \int_0^\infty \int_{x_1/w}^\infty e^{-x_1-x_2}\,\mathrm{d}x_2\,\mathrm{d}x_1 = \frac{w}{w+1} \implies f_W(w) = \frac{1}{(1+w)^2} \tag{24}$$

Thus,

$$h_1(x_1 \mid w) = \left(\frac{w+1}{w}\right)^2 x_1e^{-(1+1/w)x_1} \implies h_1(x_1 \mid 1) = 4x_1e^{-2x_1}, \quad x_1 > 0 \tag{25}$$

Problem 6

Suppose that the random variable K has a logarithmic series distribution with parameter θ (where $0 < \theta < 1$), so that

$$P(K = k) = \frac{\alpha\theta^k}{k} \qquad (k = 1, 2, \ldots)$$

where $\alpha = -[\log(1-\theta)]^{-1}$. Find the mean and variance of K.

(Method 1: Direct calculation)

$$\begin{aligned} \mathrm{E}(K) &= \alpha\sum_{k=1}^\infty \theta^k = \frac{\alpha\theta}{1-\theta} \\ \mathrm{E}(K^2) &= \alpha\sum_{k=1}^\infty k\theta^k = \alpha\theta\sum_{k=1}^\infty k\theta^{k-1} = \alpha\theta\sum_{k=1}^\infty \frac{d}{d\theta}\theta^k = \alpha\theta\cdot\frac{d}{d\theta}\frac{\theta}{1-\theta} = \frac{\alpha\theta}{(1-\theta)^2} \\ \mathrm{Var}(K) &= \frac{\alpha\theta(1-\alpha\theta)}{(1-\theta)^2} \end{aligned} \tag{26}$$

(Method 2: Probability generating function)
Recall that

$$\log(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k} \tag{27}$$

It follows that

$$\begin{aligned}
\mathrm{E}(s^X) &= G_K(s) = \alpha \sum_{k=1}^{\infty} \frac{(s\theta)^k}{k} = -\alpha \log(1-s\theta) \\
\mathrm{E}(K) &= G_K'(1) = -\left.\frac{\alpha\theta}{s\theta-1}\right|_{s=1} = \frac{\alpha\theta}{1-\theta} \\
\mathrm{Var}(K) &= G_K''(1) + G_K'(1) - \left\{G_K'(1)\right\}^2 = \left.\frac{\alpha\theta^2}{(1-s\theta)^2}\right|_{s=1} + \frac{\alpha\theta}{1-\theta} - \left(\frac{\alpha\theta}{1-\theta}\right)^2 \\
&= \frac{\alpha\theta(1-\alpha\theta)}{(1-\theta)^2} \\
&(\because G_K^{(m)}(1) = \mathrm{E}\left\{K(K-1)\cdots(K-m+1)\right\} \quad (\text{“}m\text{th factorial moment of } K\text{”}))
\end{aligned} \tag{28}$$

(Method 3: Moment generating function)

$$\begin{aligned}
\mathrm{E}(e^{tK}) &= M_K(t) = \alpha \sum_{k=1}^{\infty} \frac{(e^t\theta)^k}{k} = -\alpha \log(1-\theta e^t) \\
\mathrm{E}(K) &= M_K'(0) = -\left.\frac{\alpha\theta e^t}{\theta e^t - 1}\right|_{t=0} = \frac{\alpha\theta}{1-\theta} \\
\mathrm{E}(K^2) &= M_K''(0) = \left.\frac{\alpha\theta e^t}{(1-\theta e^t)^2}\right|_{t=0} = \frac{\alpha\theta}{(1-\theta)^2} \\
\mathrm{Var}(K) &= \frac{\alpha\theta(1-\alpha\theta)}{(1-\theta)^2}
\end{aligned} \tag{29}$$

(Method 4: Cumulant generating function)

$$\begin{aligned}
K(t) &= \log M_K(t) = \log(-\alpha \log(1-\theta e^t)) \\
\mathrm{E}(K) &= K'(0) = \left.\frac{\theta e^t}{(\theta e^t - 1)\log(1-\theta e^t)}\right|_{t=0} = \frac{-\theta}{(1-\theta)\log(1-\theta)} = \frac{\alpha\theta}{1-\theta} \quad \left(\because \alpha = -\frac{1}{\log(1-\theta)}\right) \\
\mathrm{Var}(K) &= K''(0) = \left.\frac{\theta e^t(\theta e^t + \log(1-\theta e^t))}{(\theta e^t - 1)^2 \left\{\log(1-\theta e^t)\right\}^2}\right|_{t=0} = -\frac{\theta(\theta + \log(1-\theta))}{(\theta-1)^2\left\{\log(1-\theta)\right\}^2} \\
&= -\frac{\alpha^2\theta(\theta - 1/\alpha)}{(1-\theta)^2} = \frac{\alpha\theta(1-\alpha\theta)}{(1-\theta)^2}
\end{aligned} \tag{30}$$

Problem 7

A random variable X is sub-Gaussian if there is some $c > 0$ such that

$$\mathrm{E}\left(e^{tX}\right) \leq e^{c^2t^2/2}$$

for all t.

(a) Suppose that X is sub-Gaussian. Show that $\mathrm{E}(X) = 0$ and $\mathrm{Var}(X) \leq c^2$.

The inequality translates to

$$\sum_{k=0}^{\infty} \frac{t^k}{k!} \mathrm{E}(X^k) \leq \sum_{k=0}^{\infty} \frac{c^{2k}t^{2k}}{2^k k!} \tag{31}$$

Observe that up to the quadratic terms,

$$t\,\mathrm{E}(X) + \frac{t^2}{2}\mathrm{E}(X^2) + o(t^2) \leq \frac{c^2t^2}{2} + o(t^2) \quad \text{as } t \to 0 \tag{32}$$

Note that $f(t) \in o(t^2) \iff f(t)/t^2 \to 0$ as $t \to 0$. Thus, for $t > 0$, dividing by t,

$$\lim_{t \downarrow 0}\left(\mathrm{E}(X) + \frac{t}{2}\mathrm{E}(X^2) + \frac{o(t^2)}{t}\right) \leq \lim_{t \downarrow 0}\left(\frac{c^2t}{2} + \frac{o(t^2)}{t}\right) \implies \mathrm{E}(X) \leq 0 \tag{33}$$

Again for $t < 0$, dividing by t,

$$\lim_{t \uparrow 0}\left(\mathrm{E}(X) + \frac{t}{2}\mathrm{E}(X^2) + \frac{o(t^2)}{t}\right) \geq \lim_{t \uparrow 0}\left(\frac{c^2t}{2} + \frac{o(t^2)}{t}\right) \implies \mathrm{E}(X) \geq 0 \tag{34}$$

Thus, $\mathrm{E}(X) = 0$.
Now dividing both sides by t^2,

$$\lim_{t \to 0} \frac{\mathrm{E}(X^2)}{2} \leq \lim_{t \to 0}\left(\frac{c^2}{2} + \frac{o(t^2)}{t^2}\right) \implies \mathrm{Var}(X) \leq c^2 \tag{35}$$

(b) Let $X \sim \mathrm{Unif}(0, 1)$. Show that $X - 1/2$ is sub-Gaussian and check if (a) holds under $X - 1/2$.

$$\mathrm{E}(e^{t(X-1/2)}) = \int_0^1 e^{tx-t/2}\,\mathrm{d}x = \frac{e^{t/2} - e^{-t/2}}{t} \tag{36}$$

Observe that

$$\begin{aligned}
e^{t/2} &= \sum_{k=0}^{\infty} \frac{(t/2)^k}{k!} \\
e^{-t/2} &= \sum_{k=0}^{\infty} \frac{(-1)^k(t/2)^k}{k!} \\
e^{t/2} - e^{-t/2} &= 2\sum_{k=0}^{\infty} \frac{(t/2)^{2k+1}}{(2k+1)!} \\
\mathrm{E}(e^{t(X-1/2)}) &= \sum_{k=0}^{\infty} \frac{(t/2)^{2k}}{(2k+1)!} \leq \sum_{k=0}^{\infty} \frac{(t/2)^{2k}}{k!2^k} = \sum_{k=0}^{\infty} \frac{(t^2/8)^k}{k!} = e^{t^2/8} \qquad (\because (2k+1)! \geq k!2^k)
\end{aligned} \tag{37}$$

Thus, $X - 1/2$ is sub-Gaussian and $c = 1/2$.

Problem 8

Let X denote a random variable with a standard Laplace distribution with p.d.f

$$f(x) = \frac{1}{2}\exp\{-|x|\}, \ -\infty < x < \infty.$$

Find the first, second and the third cumulants, i.e. κ_1, κ_2 and κ_3.

$$\begin{aligned}
M_X(t) &= \frac{1}{2}\int_{-\infty}^{\infty} e^{-|x|+tx}\,dx && (38)\\
&= \frac{1}{2}\left(\int_{-\infty}^{0} e^{(1+t)x}\,dx + \int_{0}^{\infty} e^{(-1+t)x}\,dx\right) && (39)\\
&= \frac{1}{2}\left(\frac{1}{1+t}e^{(1+t)x}\bigg|_{-\infty}^{0} + \frac{1}{-1+t}e^{(-1+t)x}\bigg|_{0}^{\infty}\right) && (40)\\
&= \frac{1}{2}\left(\frac{1}{1+t} - \frac{1}{-1+t}\right) && (41)\\
&= \frac{1}{1-t^2} && (42)
\end{aligned}$$

Then

$$K_X(t) = -\log\left(1-t^2\right) \tag{43}$$

Thus,

$$\begin{aligned}
\kappa_1 &= \frac{d}{dt}K_X(t)\bigg|_{t=0} = 0\\
\kappa_2 &= \frac{2(t^2+1)}{(1-t^2)^2}\bigg|_{t=0} = 2\\
\kappa_3 &= -\frac{4t(t^2+3)}{(t^2-1)^3}\bigg|_{t=0} = 0
\end{aligned} \tag{44}$$