Homework #2

Ruixin Guo

October 8, 2021

1.

$$\begin{aligned}
\text{MSE}(T) &= E(T-\theta)^2 \\
&= E(T^2 - 2T\theta + \theta^2) \\
&= E(T^2) - E(2T\theta) + E(\theta^2) \\
&= E(T^2) - 2\theta E(T) + \theta^2 \qquad [\theta \text{ is a constant.}] \\
&= (E(T^2) - E(T)^2) + (E(T)^2 - 2\theta E(T) + \theta^2) \\
&= \text{Var}(T) + (b(T))^2 \qquad [\text{Var}(T) = (E(T^2) - E(T)^2);\ b(T) = E(T-\theta)^2]
\end{aligned}$$

2. Let $F_Z(x) = P(Z < x)$ be the CDF of $\mathcal{N}(0,1)$.

$$\begin{aligned}
F_{SZ}(x) &= P(SZ < x) \\
&= P(Z < x|S=1)P(S=1) + P(-Z < x|S=-1)P(S=-1) \\
&= \frac{1}{2}P(Z < x) + \frac{1}{2}P(Z > -x) \\
&= P(Z < x) \\
&= F_Z(x)
\end{aligned}$$

The fourth equality is because the PDF of $\mathcal{N}(0,1)$ is symmetric about the y-axis, $P(Z < x) = P(Z > -x)$. Thus Z and SZ has the same distribution. $SZ \sim \mathcal{N}(0,1)$.

3.

a. The PDF of $\mathcal{N}(\mu,\sigma)$ is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$$

When μ is known, $f(x)$ can be expressed as a conditional probability indexed by σ:

$$f(x|\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{(x-\mu)^2}{2\sigma^2}\} \tag{1}$$

Let $h(x) = \frac{1}{\sqrt{2\pi}}$, $c(\sigma) = \frac{1}{\sigma}$, $w_1(\sigma) = \frac{1}{\sigma^2}$ and $t_1(x) = -\frac{(x-\mu)^2}{2}$. Equation (1) can be expressed as

$$f(x|\sigma) = h(x)c(\sigma)\exp\{w_1(\sigma)t_1(x)\}$$

Since $h(x) \geq 0, c(\sigma) \geq 0$, and $t_1(x)$ does not depend on σ, $f(x|\sigma)$ belongs to exponential family.

When σ is known, $f(x)$ can be expressed as a conditional probability indexed by μ:

$$f(x|\mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2\sigma^2}(x^2 - 2\mu x + \mu^2)\} \tag{2}$$

Let $h(x) = \frac{1}{\sqrt{2\pi}\sigma}$, $c(\sigma) = 1$, $w_1(\mu) = 1$, $t_1(x) = x^2$, $w_2(\mu) = 2\mu$, $t_2(x) = x$, $w_3(\mu) = \mu^2$ and $t_3(x) = 1$. Equation (1) can be expressed as

$$f(x|\sigma) = h(x)c(\sigma) \exp\{\sum_{i=1}^{3} w_i(\mu)t_i(x)\}$$

Since $h(x) \geq 0, c(\mu) \geq 0$, and $t_1(x), t_2(x), t_3(x)$ does not depend on μ, $f(x|\mu)$ belongs to exponential family.

b. The PDF of Beta(a, b) is

$$f(x) = \frac{1}{\beta(a, b)} x^{a-1}(1-x)^{b-1}$$

where $\beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ and $0 < x < 1$.

When a is known, $f(x)$ can be expressed as a conditional probability indexed by b:

$$f(x|b) = \frac{1}{\beta(a, b)} x^{a-1} \exp\{(b-1)\log(1-x)\} \tag{3}$$

Let $h(x) = x^{a-1}, c(b) = \frac{1}{\beta(a,b)}, w_1(b) = b - 1$ and $t_1(x) = \log(1-x)$. Equation (3) can be expressed as

$$f(x|b) = h(x)c(b) \exp\{w_1(b)t_1(x)\}$$

Thus $f(x|b)$ belongs to exponential family.

When b is known, $f(x)$ can be expressed as a conditional probability indexed by a:

$$f(x|a) = \frac{1}{\beta(a, b)} x^{b-1} \exp\{(a-1)\log(1-x)\} \tag{4}$$

Let $h(x) = x^{b-1}, c(a) = \frac{1}{\beta(a,b)}, w_1(a) = a - 1$ and $t_1(x) = \log(1-x)$. Equation (4) can be expressed as

$$f(x|a) = h(x)c(a) \exp\{w_1(a)t_1(x)\}$$

Thus $f(x|a)$ belongs to exponential family.

c. The PMF of Pois(λ) is

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

which can be regarded as a conditional probability indexed by λ:

$$P(X = x|\lambda) = \frac{1}{x!} e^{-\lambda} \exp\{x \log \lambda\} \tag{5}$$

Let $h(x) = \frac{1}{x!}, c(\lambda) = e^{-\lambda}, w_1(\lambda) = \log \lambda$ and $t_1(x) = x$. Equation (5) can be expressed as

$$P(X = x|\lambda) = h(x)c(\lambda) \exp\{w_1(\lambda)t_1(x)\}$$

Thus $P(X = x|\lambda)$ belongs to exponential family.

d. The PMF of NBin(r, p) is

$$P(X = x) = \binom{x+r-1}{r-1} p^r (1-p)^x$$

When r is known, $P(X = x)$ can be regarded as a conditional probability indexed by p:

$$P(X = x|p) = \binom{x+r-1}{r-1} \exp\{r \log p + x \log(1-p)\} \tag{6}$$

Let $h(x) = \binom{x+r-1}{r-1}, c(p) = 1, w_1(p) = r \log p, t_1(x) = 1, w_2(p) = \log(1-p)$ and $t_2(x) = x$. Equation (6) can be expressed as

$$P(X = x|p) = h(x)c(p) \exp\{\sum_{i=0}^{2} w_i(p)t_i(x)\}$$

Thus $P(X = x|p)$ belongs to exponential family.

4.

Given that $E(X) = 0, Var(X) = 1, E(Y) = 0, Var(Y) = 1, Cov(X, Y) = \rho$, we have

$$E(Z) = E(aX + bY) = aE(X) + bE(Y) = 0 \tag{7}$$
$$E(W) = E(cX + dY) = cE(X) + dE(Y) = 0 \tag{8}$$

$$\begin{aligned} Var(Z) = Var(aX + bY) &= Var(aX) + Var(bY) + 2Cov(aX, bY) \\ &= a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y) \\ &= a^2 + b^2 + 2ab\rho = 1 \end{aligned} \tag{9}$$

$$\begin{aligned} Var(W) = Var(cX + dY) &= Var(cX) + Var(dY) + 2Cov(cX, dY) \\ &= c^2 Var(X) + d^2 Var(Y) + 2cdCov(X, Y) \\ &= c^2 + d^2 + 2cd\rho = 1 \end{aligned} \tag{10}$$

$$\begin{aligned} Cov(Z, W) &= Cov(aX + bY, cX + dY) \\ &= Cov(aX, cX) + Cov(aX, dY) + Cov(bY, cX) + Cov(bY, dY) \\ &= acVar(X) + adCov(X, Y) + bcCov(Y, X) + bdVar(Y) \\ &= ac + bd + \rho(ad + bc) = 0 \end{aligned} \tag{11}$$

Equations (7) and (8) hold for any a, b, c, d.

Equations (9), (10) and (11) hold only when a, b, c, d satisfy the following relationship (in terms of ρ):

$$\begin{cases} a^2 + b^2 + 2ab\rho = 1 \\ c^2 + d^2 + 2cd\rho = 1 \\ ac + bd + \rho(ad + bc) = 0 \end{cases}$$

5.

$$\begin{aligned} E((Y - E(Y|X))^2|X) &= \sum_y (y - E(Y|X))^2 P(Y = y|X) \\ &= \sum_y y^2 P(Y = y|X) - \sum_y 2yE(Y|X)P(Y = y|X) + \sum_y E(Y|X)^2 P(Y = y|X) \\ &= E(Y^2|X) - 2E(Y|X)E(Y|X) + E(Y|X)^2 \\ &= E(Y^2|X) - E(Y|X)^2 \end{aligned}$$

6.

a. Since $p \sim \text{Beta}(a, b)$, we first compute $E(p), Var(p)$ and $E(p(p-1))$, which will be used in computing $E(X)$ and $Var(X)$ below.

$$\begin{aligned} E(p) &= \int_0^1 p \frac{1}{\beta(a,b)} p^{a-1}(1-p)^{b-1} dp \\ &= \frac{1}{\beta(a,b)} \int_0^1 p^a (1-p)^{b-1} dp \\ &= \frac{\beta(a+1,b)}{\beta(a,b)} = \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} = \frac{a}{a+b} \end{aligned}$$

$$\begin{aligned} Var(p) &= E(p^2) - E(p)^2 \\ &= \int_0^1 p^2 \frac{1}{\beta(a,b)} p^{a-1}(1-p)^{b-1} dp - E(p)^2 \\ &= \frac{\beta(a+2,b)}{\beta(a,b)} - E(p)^2 \\ &= \frac{(a+1)a}{(a+b+1)(a+b)} - \frac{a^2}{(a+b)^2} = \frac{ab}{(a+b)^2(a+b+1)} \end{aligned}$$

$$\begin{aligned} E(p(1-p)) &= \int_0^1 p(1-p) \frac{1}{\beta(a,b)} p^{a-1}(1-p)^{b-1} dp \\ &= \frac{\beta(a+1,b+1)}{\beta(a,b)} = \frac{ab}{(a+b)(a+b+1)} \end{aligned}$$

By the law of total expectation,

$$\begin{aligned} E(X) &= E(E(X|P)) \\ &= E(np) \\ &= nE(p) \\ &= \frac{na}{a+b} \end{aligned}$$

By the law of total variance

$$\begin{aligned}
Var(X) &= E(Var(X|P)) + Var(E(X|P)) \\
&= E(np(1-p)) + Var(np) \\
&= nE(p(1-p)) + n^2 Var(p) \\
&= n\frac{ab}{(a+b)(a+b+1)} + n^2\frac{ab}{(a+b)^2(a+b+1)}
\end{aligned}$$

b.

$$\begin{aligned}
P(X=x) &= \int_0^1 P(X=x|P=p)P(P=p)dp \\
&= \int_0^1 \binom{n}{x} p^x(1-p)^{n-x} \cdot \frac{1}{\beta(a,b)} p^{a-1}(1-p)^{b-1}dp \\
&= \binom{n}{x} \cdot \frac{1}{\beta(a,b)} \int_0^1 p^{x+a-1}(1-p)^{n+b-x-1}dp \\
&= \binom{n}{x} \cdot \frac{1}{\beta(a,b)} \cdot \beta(x+a, n+b-x) \\
&= \binom{n}{x} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(x+a)\Gamma(n+b-x)}{\Gamma(a+b+n)}
\end{aligned}$$

c. Since $E(P) = \frac{a}{a+b}$ and $Var(P) = \frac{ab}{(a+b)^2(a+b+1)}$,

$$\begin{aligned}
Var(X) &= nE(P)(1-E(P)) + n(n-1)Var(P) \\
&= n\frac{a}{a+b}(1 - \frac{a}{a+b}) + n(n-1)\frac{ab}{(a+b)^2(a+b+1)} \\
&= n\frac{ab}{(a+b)^2} + n^2\frac{ab}{(a+b)^2(a+b+1)} - n\frac{ab}{(a+b)^2(a+b+1)} \\
&= n[\frac{ab}{(a+b)^2} - \frac{ab}{(a+b)^2(a+b+1)}] + n^2\frac{ab}{(a+b)^2(a+b+1)} \\
&= n\frac{ab}{(a+b)(a+b+1)} + n^2\frac{ab}{(a+b)^2(a+b+1)}
\end{aligned}$$

Which is the same result we get in **a.**. Thus the equation is proved.