ECO 520 - Fall 2019
Final Exam

Haixiang Zhu
SBU ID#113029589

December 6, 2019

1. *Proof.*

$$\because X_n \xrightarrow{p} X, Y_n \xrightarrow{p} Y, X = Y$$
$$\therefore \lim_{n\to\infty} P(|X_n - X| > \epsilon_1) \to 0 \quad \forall \epsilon_1 > 0$$
$$\lim_{n\to\infty} P(|Y_n - Y| > \epsilon_2) \to 0 \quad \forall \epsilon_2 > 0$$
$$\lim_{n\to\infty} P(|X_n - Y_n| = |X_n - X - (Y_n - Y)| \geq |X_n - X| - |Y_n - Y| > \epsilon) \to 0$$
where $\epsilon = |\epsilon_1 - \epsilon_2|$

$\square$

2. (a) MGF

$$\begin{aligned}
M_{XY}(s,t) &= E[e^{sX+tY}] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{sx+ty} f_{XY}(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int_0^1\int_0^1 e^{sx+ty}(x+y)\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{s}\int_0^1 e^{ty}[\frac{s-1}{s}e^s + \frac{1}{s} + y(e^s-1)]\mathrm{d}y \\
&= \frac{1}{st}[(e^t-1)(\frac{s-1}{s}e^s+\frac{1}{s}) + (e^s-1)(\frac{t-1}{t}e^t+\frac{1}{t})] \\
&= \frac{2st-s-t}{s^2t^2}e^{s+t} + \frac{s+t-st}{s^2t^2}(e^s+e^t) - \frac{s+t}{s^2t^2} \qquad (1)
\end{aligned}$$

CF

$$\begin{aligned}
\phi_{XY}(u,v) &= E[e^{iuX+ivY}] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{iux+ivy} f_{XY}(x,y)\mathrm{d}x\mathrm{d}y
\end{aligned}$$
$$\xlongequal{\text{plug } s=iu, t=iv \text{ into (1)}} -\frac{2uv+i(u+v)}{u^2v^2}e^{i(u+v)} + \frac{uv+i(u+v)}{u^2v^2}(e^{iu}+e^{iv}) - \frac{i(u+v)}{u^2v^2}$$

(b) The marginal distribution of X is

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{XY}(x,y)\mathrm{d}y \\
&= \int_0^1 (x+y)\mathrm{d}y \\
&= x + \frac{1}{2}
\end{aligned}
$$

MGF

$$
\begin{aligned}
M_X(t) &= E[e^{tX}] \\
&= \int_{-\infty}^{\infty} e^{tx} f_X(x)\mathrm{d}x \\
&= \int_0^1 e^{tx}(x+\frac{1}{2})\mathrm{d}x \\
&= \frac{1}{t}[\frac{1}{2}(e^t-1) + (\frac{t-1}{t}e^t + \frac{1}{t})] \\
&= \frac{3t-2}{2t^2}e^t + \frac{2-t}{2t^2} \qquad (2)
\end{aligned}
$$

CF

$$
\begin{aligned}
\phi_X(u) &= E[e^{iuX}] \\
&= \int_{-\infty}^{\infty} e^{iux} f_X(x)\mathrm{d}x \\
&\overset{\text{plug } t=iu \text{ into (2)}}{=\!=\!=\!=\!=} -\frac{3iu-2}{2u^2}e^{iu} + \frac{iu-2}{2u^2}
\end{aligned}
$$

3. The MGF of a binomial distribution converges to the MGF of a Poisson distribution.

Proof.

$$
\begin{aligned}
\because M_X(t) &= (pe^t + 1 - p)^n \\
\therefore \lim_{n\to\infty, p\to 0} M_X(t) &= \lim_{n\to\infty, p\to 0}[1 + p(e^t-1)]^n \\
&= \lim_{n\to\infty, p\to 0}[1 + p(e^t-1)]^{[\frac{1}{p(e^t-1)}\cdot p(e^t-1)n]} \\
&= e^{\lim_{np\to\lambda} p(e^t-1)n} \\
&= e^{\lambda(e^t-1)}
\end{aligned}
$$

□

4. MLE of the bivariate normal distribution

(a) MLE with known parameters

Note that $f_{X_1X_2}(-x_1,-x_2) = f_{X_1X_2}(x_1,x_2)$, applying MOM

$$\begin{cases} \mu_1 = E[X_1] = 0 \\ \mu_2 = E[X_2] = 0 \\ \sigma_1^2 = E[X_1^2] = \dfrac{1}{n}\displaystyle\sum_i x_{1i}^2 \\ \sigma_2^2 = E[X_2^2] = \dfrac{1}{n}\displaystyle\sum_i x_{2i}^2 \end{cases} \tag{3}$$

The log-likelihood function (ignoring any constant term)

$$l_{X_1X_2}(\rho) = -\frac{n}{2}\log(1-\rho^2) - \sum_{i=1}^{n} \frac{\sigma_2^2 x_{1i}^2 - 2\rho\sigma_1\sigma_2 x_{1i}x_{2i} + \sigma_1^2 x_{2i}^2}{2(1-\rho^2)\sigma_1^2\sigma_2^2}$$

FOC

$$\begin{aligned} \frac{\partial l_{X_1X_2}(\rho)}{\partial\rho} &= \frac{n\rho}{1-\rho^2} + \sum_{i=1}^{n}\frac{x_{1i}x_{2i}}{(1-\rho^2)\sigma_1\sigma_2} - \rho\sum_{i=1}^{n}\frac{\sigma_2^2 x_{1i}^2 - 2\rho\sigma_1\sigma_2 x_{1i}x_{2i} + \sigma_1^2 x_{2i}^2}{(1-\rho^2)^2\sigma_1^2\sigma_2^2} \\ &\overset{\text{plug (3)}}{=\!=} \frac{n\rho}{1-\rho^2} - \frac{2n\rho}{(1-\rho^2)^2} + \sum_{i=1}^{n}\frac{x_{1i}x_{2i}}{(1-\rho^2)\sigma_1\sigma_2} + 2\rho^2\sum_{i=1}^{n}\frac{x_{1i}x_{2i}}{(1-\rho^2)^2\sigma_1\sigma_2} \\ &= \frac{1+\rho^2}{(1-\rho^2)^2}\left(\frac{\sum_{i=1}^{n} x_{1i}x_{2i}}{\sigma_1\sigma_2} - n\rho\right) = 0 \end{aligned}$$

FOC is satisfied when the term inside parenthesis is equal to 0. Thus

$$\hat{\rho}_n = \frac{\sum_{i=1}^{n} x_{1i}x_{2i}}{n\sigma_1\sigma_2}$$

(b) Asymptotic distribution

The second derivative is given by

$$\begin{aligned} H_n(\rho) &= \frac{\partial^2 l_{X_1X_2}(\rho)}{\partial\rho^2} \\ &= \left(\frac{1+\rho^2}{(1-\rho^2)^2}\right)'\left(\frac{\sum_{i=1}^{n} x_{1i}x_{2i}}{\sigma_1\sigma_2} - n\rho\right) - n\frac{1+\rho^2}{(1-\rho^2)^2} \end{aligned}$$

$$H_n(\hat{\rho}_n) = -n\frac{1+\hat{\rho}_n^2}{(1-\hat{\rho}_n^2)^2} < 0$$

which again justify the MLE above.
The information matrix is

$$I_n(\hat{\rho}_n) = -E[H_n(\hat{\rho}_n)] = \frac{1+\hat{\rho}_n^2}{(1-\hat{\rho}_n^2)^2}$$

Directly using ML theory

$$\sqrt{n}(\hat{\rho}_n - \rho) \xrightarrow{d} N\left(0, \frac{(1-\rho^2)^2}{1+\rho^2}\right)$$

(c) MLE with unknown parameters (see statistical inference 7.18)
The log-likelihood function (ignoring any constant term)

$$\begin{aligned} l(\rho|x_1, x_2) &= l(\rho|x_1)l(\rho, x_2|x_1) \\ &= -\frac{n}{2}[\log(\sigma_1^2) + \log(\sigma_2^2)] - \frac{\sum_{i=1}^n (x_{1i}-\mu_1)^2}{2\sigma_1^2} - \frac{\sum_{i=1}^n \left[x_{2i} - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(x_{1i}-\mu_1)\right]^2}{2\sigma_2^2(1-\rho^2)} \end{aligned}$$

FOC

$$\frac{\partial l(\rho|x_1,x_2)}{\partial \mu_2} = \frac{\sum_{i=1}^n x_{2i} - \mu_2 - \rho\frac{\sigma_2}{\sigma_1}(x_{1i}-\mu_1)}{\sigma_2^2(1-\rho^2)} = 0$$

$$\Rightarrow \sum_{i=1}^n (x_{2i} - \mu_2) = \sum_{i=1}^n \rho\frac{\sigma_2}{\sigma_1}(x_{1i}-\mu_1) \tag{4}$$

$$\frac{\partial l(\rho|x_1,x_2)}{\partial \sigma_2^2} = -\frac{n}{2\sigma_2^2} + \frac{1}{2\sigma_2^3}\sum_{i=1}^n (x_{2i}-\mu_2)\left[\frac{x_{2i}-\mu_2}{\sigma_2} - \rho\frac{x_{1i}-\mu_1}{\sigma_1}\right] = 0$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^n \left[\frac{(x_{2i}-\mu_2)^2}{\sigma_2^2} - \rho\frac{(x_{1i}-\mu_1)(x_{2i}-\mu_2)}{\sigma_1\sigma_2}\right] = 1 \tag{5}$$

Similarly, do the same procedure for $l(\rho|x_1, x_2) = l(\rho|x_2)l(\rho, x_1|x_2)$

$$\sum_{i=1}^n (x_{1i} - \mu_1) = \sum_{i=1}^n \rho\frac{\sigma_1}{\sigma_2}(x_{2i}-\mu_2) \tag{6}$$

$$\frac{1}{n}\sum_{i=1}^n \left[\frac{(x_{1i}-\mu_1)^2}{\sigma_1^2} - \rho\frac{(x_{1i}-\mu_1)(x_{2i}-\mu_2)}{\sigma_1\sigma_2}\right] = 1 \tag{7}$$

From (4) and (6), $\rho^2 = 1$, a contracdition. Therefore, it is must be the case that

$$\sum_{i=1}^n (x_{1i} - \mu_1) = \sum_{i=1}^n (x_{1i} - \mu_1) = 0$$

$$\Rightarrow \hat{\mu}_1 = \bar{x}_1, \hat{\mu}_2 = \bar{x}_2 \tag{8}$$

From (5), (7) and (8)

$$\frac{\frac{1}{n}\sum_{i=1}^{n}(x_{1i}-\bar{x}_1)^2}{\sigma_1^2} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_{2i}-\bar{x}_2)^2}{\sigma_2^2}$$

$$\Rightarrow \hat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{1i}-\bar{x}_1)^2, \hat{\sigma}_2^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{2i}-\bar{x}_2)^2 \tag{9}$$

Plugging (8) and (9) into joint distribution

$$l(\rho|x_1,x_2) = -\frac{n}{2}\log(1-\rho^2) - \frac{1}{2(1-\rho^2)}\left[2n - 2\rho\sum_{i=1}^{n}\frac{(x_{1i}-\bar{x}_1)(x_{2i}-\bar{x}_2)}{\hat{\sigma}_1\hat{\sigma}_2}\right]$$

FOC

$$\begin{aligned}\frac{\partial l(\rho|x_1,x_2)}{\partial\rho} &= \frac{n\rho}{1-\rho^2} - \frac{2n\rho}{(1-\rho^2)^2} + \frac{1+\rho^2}{(1-\rho^2)^2}\sum_{i=1}^{n}\frac{(x_{1i}-\bar{x}_1)(x_{2i}-\bar{x}_2)}{\hat{\sigma}_1\hat{\sigma}_2} \\ &= \frac{1+\rho^2}{(1-\rho^2)^2}\left[\sum_{i=1}^{n}\frac{(x_{1i}-\bar{x}_1)(x_{2i}-\bar{x}_2)}{\hat{\sigma}_1\hat{\sigma}_2} - n\rho\right] = 0\end{aligned}$$

MLE is

$$\hat{\rho} = \frac{1}{n}\sum_{i=1}^{n}\frac{(x_{1i}-\bar{x}_1)(x_{2i}-\bar{x}_2)}{\hat{\sigma}_1\hat{\sigma}_2}$$

5. Poisson distribution

   (a) Posterior distribution

$$\begin{aligned}\pi_{\Theta|X}(\theta|x) &= \frac{\pi_\Theta(\theta)f_{X|\Theta}(x|\theta)}{f_X(x)} \\ &= \frac{\frac{\lambda e^{-(\lambda+1)\theta}\theta^x}{x!}}{\int_0^\infty \frac{\lambda e^{-(\lambda+1)\theta}\theta^x}{x!}\mathrm{d}\theta} \\ &= \frac{(-1)^{x+2}(\lambda+1)^{x+1}e^{-(\lambda+1)\theta}\theta^x}{x!}\end{aligned}$$

   Unfortunately, the Poisson and exponential do not form a conjugate family of distributions.

   (b) The likelihood and log-likelihood are

$$\begin{aligned}L_X(\theta) &= \prod_{i=1}^{n}\frac{e^{-\theta}\theta^{x_i}}{x_i!} \\ l_X(\theta) &= -n\theta - \sum_{i=1}^{n}\log(x_i!) + \log(\theta)\sum_{i=1}^{n}x_i\end{aligned}$$

FOC

$$-n + \frac{1}{\theta}\sum_{i=1}^{n} x_i = 0$$

Hence, the MLE is

$$\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$$

(c)

$$\begin{aligned}
E[X] &= \int_0^{\infty} \frac{(-1)^{x+2}(\lambda+1)^{x+1}e^{-(\lambda+1)\theta}\theta^x}{(x-1)!}\mathrm{d}\theta \\
&= (-1)^2(\lambda+1)x\int_0^{\infty} e^{-(\lambda+1)\theta}\mathrm{d}\theta \\
&= x
\end{aligned}$$

(d) MAP

$$\begin{aligned}
\hat{\theta}_{MAP} &= \arg\max_{\theta}[l_X(\theta) + \log \pi_{\Theta}(\theta)] \\
&= \arg\max_{\theta}[-n\theta + \log(\theta)\sum_{i=1}^{n} x_i - \lambda\theta] \\
&\Rightarrow -n + \frac{1}{\theta}\sum_{i=1}^{n} x_i - \lambda = 0 \\
\hat{\theta}_{MAP} &= \frac{1}{n+\lambda}\sum_{i=1}^{n} x_i
\end{aligned}$$