STAT2602 Assignment 1

Cheng Ho Ming, Eric (3036216734) [Section 1A, 2024]

September 30, 2024

1. (i) The cumulative density function of X is:

$$\begin{aligned}
F(x) = Pr(X \leq x) &= \sum_{i=1}^{x} f(i) \\
&= \sum_{i=1}^{x} 2 * (\frac{1}{3})^x \\
&= 2 * (\frac{1}{3})^1 + 2 * (\frac{1}{3})^2 + ... + 2 * (\frac{1}{3})^x \\
&= 2 * (\frac{\frac{1}{3} * (1 - (\frac{1}{3})^x)}{1 - \frac{1}{3}}) \\
&= \frac{2 * \frac{1}{3}}{\frac{2}{3}} * (1 - (\frac{1}{3})^x) \\
&= 1 - \frac{1}{3^x} \text{ for } x = 1,2,3,\ldots
\end{aligned}$$

(ii) The moment generating function (m.g.f.) of X is:

$$\begin{aligned}
M_X(t) = E(e^{tX}) &= \sum_{x=1}^{\infty} e^{tx} f(x) \\
&= \sum_{x=1}^{\infty} e^{tx} * 2 * (\frac{1}{3})^x \\
&= 2 * \sum_{x=1}^{\infty} (\frac{e^t}{3})^x \text{ for } e^t < 3 \\
&= 2 * \frac{\frac{e^t}{3}}{1 - \frac{e^t}{3}} \text{ for } t < \ln 3 \text{ (Geometric series)} \\
\therefore M_X(t) &= \frac{2e^t}{3 - e^t} \text{ for } t < \ln 3
\end{aligned}$$

(iii) Using the moment generating function,

$$
\begin{aligned}
E(X) &= M'_X(0) \\
&= \frac{d}{dt}(\frac{2e^t}{3-e^t})|_{t=0} \\
&= \frac{(3-e^t)(2e^t)-(2e^t)(-e^t)}{(3-e^t)^2}|_{t=0} \\
&= \frac{(3-1)(2)-(2)(-1)}{(3-1)^2} \\
&= \frac{6}{4} \\
&= \frac{3}{2} \\
Var(X) &= M''_X(0)-(M'_X(0))^2 \\
&= \frac{d^2}{dt^2}(\frac{2e^t}{3-e^t})|_{t=0}-(\frac{3}{2})^2 \\
&= \frac{d}{dt}(\frac{(3-e^t)(2e^t)-(2e^t)(-e^t)}{(3-e^t)^2})|_{t=0}-\frac{9}{4} \\
&= \frac{d}{dt}(\frac{6e^t-2e^{2t}+2e^{2t}}{(3-e^t)^2})|_{t=0}-\frac{9}{4} \\
&= \frac{d}{dt}(\frac{6e^t}{(3-e^t)^2})|_{t=0}-\frac{9}{4} \\
&= (\frac{(3-e^t)^2*6e^t-6e^t*2(3-e^t)(-e^t)}{(3-e^t)^4})|_{t=0}-\frac{9}{4} \\
&= \frac{(3-1)^2*6-6*2(3-1)(-1)}{(3-1)^4}-\frac{9}{4} \\
&= \frac{4*6+6*2*2}{(3-1)^4}-\frac{9}{4} \\
&= \frac{24+24}{16}-\frac{9}{4} \\
&= 3-\frac{9}{4} \\
&= \frac{3}{4}
\end{aligned}
$$

2. (i) Since $X_1, X_2, ..., X_n$ are $\sim_{\text{i.i.d.}} Gamma(3,\theta)$, the moment generating function (m.g.f.) of Y is:

$$
\begin{aligned}
M_Y(t) &= M_{X_1}(t)*M_{X_2}(t)*...*M_{X_n}(t) \\
&= (\frac{\theta}{\theta-t})^3*(\frac{\theta}{\theta-t})^3*...*(\frac{\theta}{\theta-t})^3 \\
&= (\frac{\theta}{\theta-t})^{3n} \\
&= \text{m.g.f. of } Gamma(3n,\theta)
\end{aligned}
$$

$\therefore Y \sim Gamma(3n,\theta)$ where $\theta > 0$.

(ii) Using the moment generating function,

$$
\begin{aligned}
E(cY) &= cE(Y) = cM'_Y(0) \\
\theta^{-1} &= c(\frac{d}{dt}(\frac{\theta}{\theta-t})^{3n})|_{t=0} \\
\theta^{-1} &= c(3n(\frac{\theta}{\theta-t})^{3n-1}\frac{\theta}{(\theta-t)^2})|_{t=0} \\
\theta^{-1} &= c*3n*1*\frac{1}{\theta} \\
1 &= 3nc \\
c &= \frac{1}{3n}
\end{aligned}
$$

(iii) The moment generating function (m.g.f.) of $3\theta Y + 1$ is:

$$\begin{aligned} M_{3\theta Y+1}(t) &= E(e^{t(3\theta Y+1)}) \\ &= e^t E(e^{3\theta Y t}) \\ &= e^t M_Y(3\theta t) \\ &= e^t (\frac{\theta}{\theta - 3\theta t})^{3n} \\ &= e^t (\frac{1}{1-3t})^{3n} \text{ for } t < \frac{1}{3} \end{aligned}$$

3. (i) The mean of X is:

$$\begin{aligned} E(X) &= M'_X(0) \\ &= (-\frac{3}{4}e^{-3t} + \frac{e^t}{4})|_{t=0} \\ &= -\frac{3}{4} + \frac{1}{4} \\ &= -\frac{1}{2} \end{aligned}$$

The variance of X is:

$$\begin{aligned} Var(X) &= M''_X(0) - (M'_X(0))^2 \\ &= (\frac{9}{4}e^{-3t} + \frac{e^t}{4})|_{t=0} - (-\frac{1}{2})^2 \\ &= \frac{9}{4} + \frac{1}{4} - \frac{1}{4} \\ &= \frac{9}{4} \end{aligned}$$

(ii) By the moment generating function,

$$\begin{aligned} E(e^{tX}) &= M_X(t) \\ \sum_{x \in X(\Omega)} e^{tx} f(x) &= \frac{1}{4}e^{-3t} + \frac{1}{2} + \frac{1}{4}e^t \\ &= \frac{1}{4}e^{-3t} + \frac{1}{2}e^{0t} + \frac{1}{4}e^{1t} \end{aligned}$$

Comparing the coefficients of e^{-3t}, e^{0t}, e^{1t} on both sides, we have:

$$\therefore f(x) = \begin{cases} \frac{1}{4} & \text{if } x = -3 \\ \frac{1}{2} & \text{if } x = 0 \\ \frac{1}{4} & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases} \text{, which is the pmf of } X.$$

Checking the expression of the pmf by (i), we have:

$$\begin{aligned} E(X) &= -3 * \frac{1}{4} + 0 * \frac{1}{2} + 1 * \frac{1}{4} \\ &= -\frac{1}{2} \\ Var(X) &= E(X^2) - (E(X))^2 \\ &= ((-3)^2 * \frac{1}{4} + 0^2 * \frac{1}{2} + 1^2 * \frac{1}{4}) - (-\frac{1}{2})^2 \\ &= (\frac{9}{4} + \frac{1}{4}) - \frac{1}{4} \\ &= \frac{9}{4} \end{aligned}$$

Which matches the results in (i).

4. (i) The empirical distribution function is:

$$F_{10}(x) = \begin{cases} 0 & \text{for } x < 0 \\ 0.1 & \text{for } 0 \le x < 1 \\ 0.2 & \text{for } 1 \le x < 2 \\ 0.4 & \text{for } 2 \le x < 3 \\ 0.6 & \text{for } 3 \le x < 4 \\ 0.7 & \text{for } 4 \le x < 6 \\ 0.9 & \text{for } 6 \le x < 7 \\ 1 & \text{for } x >= 7 \end{cases}$$

(ii) Using the empirical distribution, we have:

$$\begin{aligned} Pr(X <= 4) &= F(4) \approx F_{10}(4) = 0.7 \\ Pr(4 < X < 7) &= Pr(4 < X <= 6) \\ &\approx F_{10}(6) - F_{10}(4) \\ &= 0.9 - 0.7 \\ &= 0.2 \end{aligned}$$

5. (i) $\because \xi_1$ and ξ_2 are independent,

$\therefore$ the moment generating function of X is:

$$\begin{aligned} M_X(t) &= M_{\xi_1}(t) * M_{\xi_2}(t) \\ &= \exp(\theta t + \frac{1}{2}t^2) * \exp(\lambda\theta t + \frac{1}{2}\lambda^2 t^2) \\ &= \exp(\theta t + \frac{1}{2}t^2 + \lambda\theta t + \frac{1}{2}\lambda^2 t^2) \\ &= \exp((\theta + \lambda\theta)t + \frac{1}{2}(1 + \lambda^2)t^2) \text{ for } t \in \mathbb{R} \end{aligned}$$

(ii)

$$\begin{aligned} M_X(t) &= \exp((\theta + \lambda\theta)t + \frac{1}{2}(1 + \lambda^2)t^2) \\ M'_X(t) &= (\theta + \lambda\theta + (1 + \lambda^2)t)M_X(t) \\ M''_X(t) &= (\theta + \lambda\theta + (1 + \lambda^2)t)^2 M_X(t) + M_X(t) * (1 + \lambda^2) \\ M'''_X(t) &= (\theta + \lambda\theta + (1 + \lambda^2)t)^3 * M_X(t) + M_X(t) * 2(\theta + \lambda\theta + (1 + \lambda^2)t) * (1 + \lambda^2) \\ &\quad + (1 + \lambda^2)(\theta + \lambda\theta + (1 + \lambda^2)t)M_X(t) \\ \therefore E(X^3) &= M'''_X(0) \\ &= (\theta + \lambda\theta)^3 + 2 * (\theta + \lambda\theta)(1 + \lambda^2) + (1 + \lambda^2)(\theta + \lambda\theta) \\ &= \theta^3(1 + \lambda)^3 + 2 * \theta(1 + \lambda)(1 + \lambda^2) + \theta(1 + \lambda^2)(1 + \lambda) \\ &= \theta(1 + \lambda)(\theta^2(1 + \lambda)^2 + 2(1 + \lambda^2) + (1 + \lambda^2)) \\ &= \theta(1 + \lambda)(\theta^2(1 + \lambda)^2 + 3(1 + \lambda^2)) \end{aligned}$$

(iii) Given the moment generating function of X is:
$M_X(t) = \exp((\theta + \lambda\theta)t + \frac{1}{2}(1 + \lambda^2)t^2)$
$\therefore X \sim N(\theta + \lambda\theta, 1 + \lambda^2)$

6. Since X is a continuous random variable,

$$\begin{aligned}
\text{Mean of } X = E(X) = \mu &= \int_0^2 x\frac{x^3}{4}dx \\
&= \frac{1}{4}\int_0^2 x^4 dx \\
&= \frac{1}{4}[\frac{1}{5}x^5]_0^2 \\
&= \frac{1}{4} * \frac{1}{5} * 2^5 \\
&= \frac{32}{20} \\
&= \frac{8}{5} \\
\text{Variance of } X = Var(X) = \sigma^2 &= E(X^2) - (E(X))^2 \\
&= \int_0^2 x^2\frac{x^3}{4}dx - (\frac{8}{5})^2 \\
&= \frac{1}{4}\int_0^2 x^5 dx - \frac{64}{25} \\
&= \frac{1}{4}[\frac{1}{6}x^6]_0^2 - \frac{64}{25} \\
&= \frac{1}{4} * \frac{1}{6} * 2^6 - \frac{64}{25} \\
&= \frac{64}{24} - \frac{64}{25} \\
&= \frac{8}{75}
\end{aligned}$$

Using central limit theorem, we have:

$$\begin{aligned}
Pr(1.2 \leq \bar{X} \leq 1.6) &= Pr(\frac{\sqrt{n}(1.2-\mu)}{\sigma} \leq \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \leq \frac{\sqrt{n}(1.6-\mu)}{\sigma}) \\
&= Pr(\frac{\sqrt{25}(1.2-\frac{8}{5})}{\sqrt{8/75}} \leq Z \leq \frac{\sqrt{25}(1.6-\frac{8}{5})}{\sqrt{8/75}}) \\
&\because Z \sim N(0,1), \\
\therefore Pr(1.2 \leq \bar{X} \leq 1.6) &= Pr(-6.123724 \leq Z \leq 0) \\
&\approx 0.5
\end{aligned}$$

7. (i) Since X is a continuous random variable, the cumulative distribution function of Y is:

$F_Y(y) = Pr(Y \leq y)$

$= Pr(X_1 <= y, X_2 <= y, X_3 <= y, \ldots, X_{12} <= y)$

(as $Y = X_{(12)}$which is the 12th smallest value in the sample)

$= [Pr(X_1 <= y)]^{12}$ (as $X_1, X_2, \ldots, X_{12}$ are independent)

$= [F_X(y)]^{12}$

as $X \sim U(0,1)$

$= y^{12}$ for $y \in (0,1)$

Therefore, the probability density function of Y is:

$$\begin{aligned}
f_y(y) &= \frac{d}{dy}F_Y(y) \\
&= \frac{d}{dy}y^{12} \\
&= 12y^{11} \text{ for } y \in (0,1)
\end{aligned}$$

(ii) Given $Z = (\sum_{i=1}^{12} X_i) - 6,$ the moment generating function of Z is:

$$\begin{aligned}
M_Z(t) &= E(e^{tZ}) \\
&= E(e^{t(\sum_{i=1}^{12} X_i)-6t}) \\
&= e^{-6t} * E(e^{t\sum_{i=1}^{12} X_i}) \\
&= e^{-6t} * E(e^{tX_1} * e^{tX_2} * \cdots * e^{tX_{12}}) \\
&= e^{-6t} * [E(e^{tX_1})]^{12} \text{ (as } X_1, X_2, \ldots, X_{12} \text{ are independent)} \\
&= e^{-6t} * [M_X(t)]^{12}
\end{aligned}$$

$$\begin{aligned}
\because M_X(t) &= E(e^{tX}) \\
&= \int_0^1 e^{tx} dx \\
&= \begin{cases} \frac{e^t-1}{t} & \text{if } t \neq 0 \\ 1 & \text{if } t = 0 \end{cases}
\end{aligned}$$

$$\begin{aligned}
\therefore M_Z(t) &= \begin{cases} e^{-6t}\frac{(e^t-1)^{12}}{t^{12}} & \text{if } t \neq 0 \\ e^{-6t} & \text{if } t = 0 \end{cases} \\
&= \begin{cases} e^{-6t}\frac{(e^t-1)^{12}}{t^{12}} & \text{if } t \neq 0 \\ 1 & \text{if } t = 0 \end{cases}
\end{aligned}$$

(iii) Given $X_1, X_2, \ldots, X_{12} \overset{i.i.d.}{\sim} U(0,1)$,
Let $\bar{X} = (\sum_{i=1}^{12} X_i)/12$, then:

$$\mu = E(X_1) = 0.5$$
$$\sigma = \sqrt{Var(X_1)} = \sqrt{\frac{1}{12}}$$

We have:

$$\begin{aligned}
\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} &= \frac{\sqrt{12}(\bar{X}-0.5)}{\sqrt{\frac{1}{12}}} \\
&= 12(\bar{X}-0.5) \\
&= 12\bar{X} - 6 \\
&= Z
\end{aligned}$$

By central limit theorem, $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \to_d N(0,1)$ as $n \to \infty$.
Since $n = 12$ is large enough, we can say that Z is approximately distributed as standard normal $N(0,1)$.

8. (i) Let $P(n)$ be the predicate "$\bar{X}_n = \frac{1}{n}(\sum_{i=1}^{n} X_i) \overset{d}{=} X_1$" for $n \in \mathbb{Z}^+$, where $X \overset{d}{=} Y$ means X and Y are in the same distribution.
For $n = 1$, L.H.S. $= \bar{X}_1 = X_1$
, R.H.S. $= X_1$
$\because$ L.H.S. $\overset{d}{=}$ R.H.S. ,
$\therefore P(1)$ is true.

Assume P(k) is true for some $k \in \mathbb{Z}^+$, i.e. "$\bar{X}_k = \frac{1}{k}(\sum_{i=1}^{k} X_i) \overset{d}{=} X_1$", For $n = k+1$,

$$
\begin{aligned}
\text{L.H.S. } &= \bar{X}_{k+1} = \frac{1}{k+1}(\sum_{i=1}^{k+1} X_i) \\
&= \frac{1}{k+1}(X_{k+1} + \sum_{i=1}^{k} X_i) \\
&= \frac{1}{k+1}(X_{k+1} + k\bar{X}_k) \\
&= \frac{1}{k+1}X_{k+1} + \frac{k}{k+1}\bar{X}_k
\end{aligned}
$$

Let $p = \dfrac{1}{k+1}$ and $1 - p = 1 - \dfrac{1}{k+1} = \dfrac{k}{k+1}$,

given that $T = pU + (1-p)V$ is also distributed as Cauchy,

$$
\begin{aligned}
&= pX_{k+1} + (1-p)\bar{X}_k \\
&\stackrel{d}{=} pX_{k+1} + (1-p)X_1 \text{ (by induction hypothesis)} \\
&\stackrel{d}{=} pX_1 + (1-p)X_1 \text{ (given that } X_1 \stackrel{d}{=} X_{k+1}) \\
&\stackrel{d}{=} pX_1 - pX_1 + X_1 \\
&\stackrel{d}{=} X_1
\end{aligned}
$$

$\because P(k+1)$ is true if $P(k)$ is true,

$\therefore$ By the principle of mathematical induction, $P(n)$ is true for all $n \in \mathbb{Z}^+$.

Therefore, $\bar{X}_n = \bar{X} = \frac{1}{n}(\sum_{i=1}^{n} X_i)$ has the same distribution as X_1, which is Cauchy.

$$
\begin{aligned}
E(X_1) &= \int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)}dx \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x}{1+x^2}dx \\
&= \frac{1}{\pi}(\lim_{y\to\infty}\int_{-y}^{0} \frac{x}{1+x^2}dx + \lim_{y\to\infty}\int_{0}^{y} \frac{x}{1+x^2}dx) \\
&= -\infty + \infty
\end{aligned}
$$

$\because$ the integral is undefined,

$\therefore E(X_1)$ does not exist.

Therefore, $\lim_{n\to\infty} \bar{X}_n$ does not exist, so $\lim_{n\to\infty} Pr(|\bar{X}_n - X| \geq \epsilon)$ does not exist for any $\epsilon > 0$.

(ii) For (weak) law of large numbers, it requires the sequence of independent and identically distributed random variables to have *FINITE* mean.
Since the Cauchy distribution does not have a finite mean, the (weak) law of large numbers does not apply to the Cauchy distribution.