Homework 3 solution

Trong Dat Do*

October 14, 2020

Problem 1 As $T \sim \exp(\beta)$, the CDF of T is

$$F_T(t) = (1 - e^{-\beta t})I(t \geq 0) \tag{1}$$

Thus,

$$\begin{aligned}
P(X \leq x) &= P(X \leq x, W = 1) + P(X \leq x, W = -1) &(2)\\
&= P(T \leq x, W = 1) + P(T \geq -x, W = -1) &(3)\\
&= P(T \leq x)P(W = 1) + P(T \geq -x)P(W = -1) &(4)\\
&= \frac{2}{3}P(T \leq x) + \frac{1}{3}P(T \geq -x) &(5)
\end{aligned}$$

$\square$

If $x \geq 0$, then $P(T \leq x) = 1 - e^{-\beta x}$ and $P(T \geq -x) = 1$, thus

$$P(X \leq x) = \frac{2}{3}(1 - e^{-\beta x}) + \frac{1}{3}$$

and the corresponding density function is $f_X(x) = \frac{2\beta}{3}e^{-\beta x}$.

If $x \leq 0$, then $P(T \leq x) = 0$ and $P(T \geq -x) = e^{\beta x}$, thus

$$P(X \leq x) = \frac{1}{3}e^{\beta x}$$

and the corresponding density function is $f_X(x) = \frac{\beta}{3}e^{\beta x}$.

Therefore, the density function is

$$f_X(x) = \frac{\beta e^{\beta x}I(x < 0) + 2\beta e^{-\beta x}I(x \geq 0)}{3}$$

*Ph.D. Student, Department of Statistics, University of Michigan

Problem 2 (a) If we know the distribution is symmetric, then from the definition of symmetric distribution in the question, we have

$$f_V(t) = f_{-V}(t)$$

Using changing variable theorem, by changing $U = -V$, we have $f_{-V}(t) = f_U(t) = f_V(-t)|-1| = f_V(-t)$ and thus $f_V(-t) = f_V(t)$

On the other hand, if we know $f_V(t) = f_V(-t)$, then we could prove $f_V(t) = f_{-V}(t)$ in the same way.

(b)

$$\begin{aligned}
\mathbb{E}[X] &= \int_R tf(t)dt && (6)\\
&= \int_{-\infty}^{0} tf(t)dt + \int_0^{\infty} tf(t)dt && (7)\\
&= \int_0^{\infty} -tf(-t)dt + \int_0^{\infty} tf(t)dt && (8)\\
&= \int_0^{\infty} -tf(t)dt + \int_0^{\infty} tf(t)dt && (9)\\
&= 0 && (10)
\end{aligned}$$

(c) We need to assume that B is independent from V.

$$\begin{aligned}
\mathbb{P}(\bar{V} < t) &= \mathbb{P}(\bar{V} < t, (2B-1) = 1) + \mathbb{P}(\bar{V} < t, (2B-1) = -1) && (11)\\
&= \mathbb{P}(V < t, B = 1) + \mathbb{P}(-V < t, B = 0) && (12)\\
&= \mathbb{P}(V < t)\theta + \mathbb{P}(V < t)(1-\theta) && (13)\\
&= \mathbb{P}(V < t) && (14)
\end{aligned}$$

Take derivative of t on both sides and we will have

$$f_{\bar{V}}(t) = f(t) \qquad (15)$$

□

Problem 3 (a)

$$
\begin{aligned}
\mathbb{P}(W=m) &= \sum_{i=0}^{m} \mathbb{P}(P_1=i, P_2=m-i) && (16)\\
&= \sum_{i=0}^{m} \frac{e^{-\lambda_1}\lambda_1^i}{i!}\frac{e^{-\lambda_2}\lambda_2^{m-i}}{(m-i)!} && (17)\\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{m!}\sum_{i=0}^{m}\binom{m}{i}\lambda_1^i\lambda^{m-i} && (18)\\
&= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^m}{m!} && (19)
\end{aligned}
$$

Thus, $W \sim Poi(\lambda_1+\lambda_2)$

(b) Through the conclusion in (a), we know that this result holds for the summation of two Poisson random variables.

Now, assume this result holds for the summation of $n-1$ Poisson random variables $X_1, \ldots, X_{n-1}$ with parameters $\lambda_1, \ldots, \lambda_{n-1}$, that is $\sum_{i=1}^{n-1} X_i \sim Poi(\sum_{i=1}^{n-1}\lambda_i)$.

Then for n random variables $X_1, \ldots, X_n$ with parameters $\lambda_1, \ldots, \lambda_n$, denote $T = \sum_{i=1}^{n-1} X_i$. According to the assumption, $T \sim Poi(\sum_{i=1}^{n-1}\lambda_i)$. As the result holds for two random variables, $\sum_{i=1}^{n} X_i = T + X_n \sim Poi(\sum_{i=1}^{n-1}\lambda_i + \lambda_n) = Poi(\sum_{i=1}^{n}\lambda_i)$.

(c) By checking the m.g.f. of Poisson distribution, we have

$$\mathbb{E}[e^{tS_n}] = e^{\Lambda_n(e^t-1)}, t \in \mathbb{R}. \quad (20)$$

Let $X = \frac{S_n-\Lambda_n}{\sqrt{\Lambda_n}}$, we have

$$
\begin{aligned}
\mathbb{E}[e^{tX}] &= \mathbb{E}[e^{t\frac{S_n-\Lambda_n}{\sqrt{\Lambda_n}}}] && (21)\\
&= e^{-t\sqrt{\Lambda}}\mathbb{E}[e^{t\frac{S_n}{\sqrt{\Lambda_n}}}] && (22)\\
&= e^{-t\sqrt{\Lambda}}\exp\{\Lambda_n(e^{t/\sqrt{\Lambda_n}}-1)\}. && (23)
\end{aligned}
$$

By Taylor theorem,

$$e^{t/\sqrt{\Lambda_n}} = 1 + \frac{t}{\sqrt{n}} + \frac{t^2}{2\Lambda_n} + o(\Lambda_n^{-1}). \quad (24)$$

Plugging this equation into $\mathbb{E}[e^{tX}]$ yields:

$$
\begin{aligned}
\mathbb{E}[e^{tX}] &= e^{-t\sqrt{\Lambda}}\exp\{\Lambda_n(e^{t/\sqrt{\Lambda_n}}-1)\} && (25)\\
&= e^{-t\sqrt{\Lambda}}\exp\{\Lambda_n(1+\frac{t}{\sqrt{n}}+\frac{t^2}{2\Lambda_n}+o(\Lambda_n^{-1})-1)\} && (26)\\
&= \exp\{\frac{t^2}{2}+o(1)\} \to \exp\{\frac{t^2}{2}\}, && (27)
\end{aligned}
$$

which is the m.g.f. of $N(0,1)$ distribution. As the convergence is uniform in a neighborhood around 0, we have the limiting distribution to be $N(0,1)$. □

Problem 4 We can easily calculate the pmf to be:

$$P(Y=y) = \binom{y+r-1}{y}(1-p)^y p^r, y \in \mathbb{N}. \tag{28}$$

According to the property of pmf, we have

$$\sum_{y=0}^{\infty} \binom{y+r-1}{y}(1-p)^y p^r = p^r \sum_{y=0}^{\infty} \binom{y+r-1}{y}(1-p)^y = 1, \tag{29}$$

thus, $p^{-r} = \sum_{y=0}^{\infty} \binom{y+r-1}{y}(1-p)^y$. Then we calculate the pmf of Y to be

$$\begin{aligned} \mathbb{E}[e^{tY}] &= \sum_{y=0}^{\infty} e^{ty}\binom{y+r-1}{y}(1-p)^y p^r && (30) \\ &= p^r \sum_{y=0}^{\infty} \binom{y+r-1}{y}(e^t(1-p))^y && (31) \\ &= p^r(1-e^t(1-p))^{-r}, && (32) \end{aligned}$$

where t satisfying $|e^t(1-p)| < 1$, which contains a neighborhood around 0. To calculate the limit for $r \to \infty$, $p \to 1$ and $r(1-p) \to \lambda$, we have

$$\frac{p^r}{(1-e^t(1-p))^r} = \frac{1}{(\frac{1-e^t}{p}+e^t)^r}. \tag{33}$$

As $\lim_{n\to\infty}(1+\frac{a}{n})^n = e^a$, we can check the limit of

$$r\left(\frac{1-e^t}{p}+e^t-1\right) = r(1-e^t)\left(\frac{1-p}{p}\right) = (1-e^t)\frac{r(1-p)}{p} \to \lambda(1-e^t). \tag{34}$$

Thus,

$$\frac{1}{(\frac{1-e^t}{p}+e^t)^r} \to e^{-\lambda(1-e^t)},$$

which is the m.g.f. of Poisson random variable. Notice that the convergence is uniform for all points in a neighborhood of 0, so we have the limiting distribution is a Poisson distribution.

Note: This is called the Law of rare event. Because p is small, is it rare to have a successful trial. As n is large, to compute the CDF of this Binomial distribution is cumbersome, and this Law allow us to approximate it by distribution of Poisson. See also: `https://en.wikipedia.org/wiki/Poisson_limit_theorem`. □

Problem 5 (1) In total we need to choose $\sum_{j=1}^{l} \epsilon_{i_j}$ R's and $l - \sum_{j=1}^{l} \epsilon_{i_j}$ D's. Therefore the probability is

$$\frac{N_D(N_D-1)\dots(N_D - \sum_{j=1}^{l} \epsilon_{i_j} + 1)N_R(N_R-1)\dots(N_D - (l - \sum_{j=1}^{l} \epsilon_{i_j}) + 1)}{N(N-1)\dots(N-l+1)}. \tag{35}$$

(2) Because there are $\binom{n}{m}$ way to choose m number to take the value 1 from n number, the probability is

$$\binom{n}{m} \frac{N_D(N_D-1)\dots(N_D-m+1)N_R(N_R-1)\dots(N_D-(n-m)+1)}{N(N-1)\dots(N-n+1)} = \frac{\binom{N_D}{m}\binom{N_R}{n-m}}{\binom{N}{n}}, \tag{36}$$

which follows the Hypergeometric distribution.

(3) When $n = N$, because we choose all the people, we would have $S_N = N_D$ is a deterministic quantity. Hence

$$E(S_N/N) = N_D/N = p_D, \quad \mathrm{Var}(S_N/N) = 0. \tag{37}$$

□

Problem 6 From the question, we have f is continuous, $f(x) = f(-x) \quad \forall x \in \mathbb{R}$, and it is monotonically increasing on $(-\infty, 0]$ and monotonically decreasing on $[0, \infty)$. Write I_h as $(a - h, a + h)$. Now we need to find a to minimize $P(X \in (a - h, a + h))$. We have

$$P(X \in (a-h, a+h)) = \int_{a-h}^{a+h} f(x - \theta_0)dx = \int_{a-\theta_0-h}^{a-\theta_0+h} f(x)dx \tag{38}$$

By Leibniz's derivative rule, we have

$$\frac{\partial}{\partial a} \int_{a-\theta_0-h}^{a-\theta_0+h} f(x)dx = f(a - \theta_0 + h) - f(a - \theta_0 - h), \tag{39}$$

which equals 0 when $a = \theta_0$. Whenever $a > \theta_0$, we have

$$f(a-\theta_0+h) - f(a-\theta_0-h) = \begin{cases} f(a-\theta_0+h) - f(h-(a-\theta_0)) < 0, & \text{if } a - \theta_0 \leq h \\ f(a-\theta_0+h) - f(a-\theta_0-h) < 0, & \text{if } a - \theta_0 > h, \end{cases} \tag{40}$$

by the symmetry and the monotonically decreasing of f on $(0, \infty)$. Similarly, whenever $a < \theta_0$, we have

$$f(a-\theta_0+h) - f(a-\theta_0-h) = \begin{cases} f(-h-(a-\theta_0)) - f(-h+(a-\theta_0)) > 0, & \text{if } a - \theta_0 > -h \\ f(a-\theta_0+h) - f(a-\theta_0-h) > 0, & \text{if } a - \theta_0 \leq -h. \end{cases} \tag{41}$$

Hence

$$\frac{\partial}{\partial a} P(X \in (a-h, a+h)) \begin{cases} > 0 & \text{if } a < \theta_0, \\ = 0 & \text{if } a = \theta_0, \\ < 0 & \text{if } a > \theta_0. \end{cases} \tag{42}$$

□

Hence the optimal a we need to choose is θ_0, and $I_h = (\theta_0 - h, \theta_0 + h)$.

Note: This is similar to Theorem 9.3.2. in the textbook, where we try to find the thinnest confident interval. For more information, see chapter 9.3. Methods of Evaluating Interval Estimations.