M2S1 - EXERCISES 7: SOLUTIONS

1. The key is to find IID random variables $X_1, ..., X_n$ such that

$$X = \sum_{i=1}^{n} X_i,$$

and then to use the Central Limit Theorem result for large n:

$$Z_n = \frac{\sum\limits_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \to Z \sim Normal(0,1), \quad \text{so that } X = \sum_{i=1}^{n} X_i \sim Normal(n\mu, n\sigma^2), \text{ approximately,}$$

where $\mu = \mathrm{E}_{f_X}[X_i]$ and $\sigma^2 = \mathrm{Var}_{f_X}[X_i]$.

(i) $X \sim Binomial(n, \theta) \implies X = \sum\limits_{i=1}^{n} X_i$ where $X_i \sim Bernoulli(\theta)$, so that $\mu = \mathrm{E}_{f_X}[X_i] = \theta$ and $\sigma^2 = \mathrm{Var}_{f_X}[X_i] = \theta(1-\theta)$, and hence

$$Z_n = \frac{\sum\limits_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \sim Normal(0,1) \implies X \sim Normal(n\theta, n\theta(1-\theta)), \text{ approximately.}$$

(ii) $X \sim Poisson(\lambda) \implies X = \sum\limits_{i=1}^{n} X_i$ where $X_i \sim Poisson(\lambda/n)$, so that $\mu = \mathrm{E}_{f_X}[X_i] = \lambda/n$ and $\sigma^2 = \mathrm{Var}_{f_X}[X_i] = \lambda/n$, and hence

$$Z_n = \frac{\sum\limits_{i=1}^{n} X_i - n\frac{\lambda}{n}}{\sqrt{n(\lambda/n)}} = \frac{\sum\limits_{i=1}^{n} X_i - \lambda}{\sqrt{\lambda}} \sim Normal(0,1) \implies X \sim Normal(\lambda, \lambda), \text{ approximately.}$$

Note that this uses the result that the sum of independent Poisson variables also has a Poisson distribution (proved using mgfs).

(iii) $X \sim NegBinomial(n, \theta) \implies X = \sum\limits_{i=1}^{n} X_i$ where $X_i \sim Geometric(\theta)$, so that $\mu = \mathrm{E}_{f_X}[X_i] = 1/\theta$ and $\sigma^2 = \mathrm{Var}_{f_X}[X_i] = (1-\theta)/\theta^2$, and hence

$$Z_n = \frac{\sum\limits_{i=1}^{n} X_i - n\frac{1}{\theta}}{\sqrt{n\left((1-\theta)/\theta^2\right)}} \sim Normal(0,1) \implies X \sim Normal\left(\frac{n}{\theta}, \frac{n(1-\theta)}{\theta^2}\right), \text{ approximately.}$$

2. $Y_n = \max\{X_1, ..., X_n\}$, so in the limit as $n \to \infty$ we have the limit for *fixed* y as

$$F_{Y_n}(y) = \{F_X(y)\}^n = y^n \to \begin{cases} 0, & y < 1, \\ 1, & y \geq 1, \end{cases}$$

that is, a step function with single step of size 1 at $y = 1$. Hence the limiting random variable Y is a discrete variable with $P[Y = 1] = 1$, that is, the limiting distribution is *degenerate* at 1. Also, $Z_n = \min\{X_1, ..., X_n\}$ so in the limit as $n \to \infty$ we have the limit for *fixed* z as

$$F_{Z_n}(z) = 1 - \{1 - F_X(z)\}^n = 1 - (1-z)^n \to \begin{cases} 0, & z \leq 0, \\ 1, & z > 0, \end{cases}$$

that is, a step function with single step of size 1 at $z = 0$. Hence the limiting random variable Z is a discrete variable with $P[Z = 0] = 1$, that is, the limiting distribution is *degenerate* at 0. Note here that the limiting function is **not** a cdf as it is not **right-continuous**, but that the limiting distribution does still exist - the familiar definition of convergence in distribution only refers to pointwise convergence **at points of continuity of the limit function**, and here is limit function is not continuous at zero.

Note that these results are intuitively reasonable for, as the sample size gets increasingly large, we will obtain a random variable arbitrarily close to each end of the range. Note also that these results describe *convergence in distribution*, but also we have for $1 > \varepsilon > 0$

$$\begin{array}{l} \mathrm{P}\left[|Y_n - 1| < \varepsilon\right] = \mathrm{P}\left[1 - Y_n < \varepsilon\right] = \mathrm{P}\left[1 - \varepsilon < Y_n\right] = 1 - \mathrm{P}\left[Y_n < 1 - \varepsilon\right] = 1 - \varepsilon^n \to 1 \\ \mathrm{P}\left[|Z_n - 0| < \varepsilon\right] = \mathrm{P}\left[Z_n < \varepsilon\right] = 1 - (1 - \varepsilon)^n \to 1 \end{array} \quad \text{as } n \to \infty,$$

so we also have *convergence in probability* of Y_n to 1 and of Z_n to 0.

3. $Z_n = \min\{X_1, ..., X_n\}$, so

$$F_{Z_n}(z) = 1 - \{1 - F_X(z)\}^n = 1 - \left(1 - \left(1 - \frac{1}{z}\right)\right)^n = 1 - \frac{1}{z^n}, \qquad z > 1,$$

and so, in the limit as $n \to \infty$ we have the limit for *fixed* z as

$$F_{Z_n}(z) \to \left\{ \begin{array}{ll} 0, & z \leq 1, \\ 1, & z > 1, \end{array} \right.$$

that is, a step function with single step of size 1 at $z = 1$. Hence the limiting random variable Z is a discrete variable with

$$P[Z = 1] = 1,$$

that is, the limiting distribution is *degenerate* at 1. Again, the limiting function is not a cdf as it is not right continuous, but this does not affect our conclusion, as the limit function is not continuous at 1.

Now if $U_n = Z_n^n$, we have from first principles that, for $u > 1$,

$$F_{U_n}(u) = \mathrm{P}\left[U_n \leq u\right] = \mathrm{P}\left[Z_n^n \leq u\right] = \mathrm{P}\left[Z_n \leq u^{1/n}\right] = 1 - \frac{1}{\left(u^{1/n}\right)^n} = 1 - \frac{1}{u},$$

which is a valid cdf, but which does not depend on n. Hence the limiting distribution of U_n is precisely

$$F_U(u) = 1 - \frac{1}{u}, \quad u > 1.$$

For $u \leq 1$, $F_{U_n}(u) = 0$ for all n, so clearly $F_{U_n}(u) \longrightarrow 0$ for u in this range. Hence the limiting distribution function is continuous at $u = 1$ (indeed, at all u).

4. $Y_n = \max\{X_1, ..., X_n\}$, so

$$F_{Y_n}(y) = \{F_X(y)\}^n = \left(\frac{1}{1 + e^{-y}}\right)^n, \qquad y \in \mathbb{R},$$

and so, in the limit as $n \to \infty$, we have the limit for *fixed* y as

$$F_{Y_n}(y) \to 0, \quad \text{for all } y.$$

Hence there is *no limiting distribution.*

If $U_n = Y_n - \log n$, we have from first principles that, for $u > -\log n$,

$$F_{U_n}(u) = \mathrm{P}\left[U_n \leq u\right] = \mathrm{P}\left[Y_n - \log n \leq u\right] = \mathrm{P}\left[Y_n \leq u + \log n\right] = F_{Y_n}(u + \log n) = \left(\frac{1}{1 + e^{-u - \log n}}\right)^n,$$

so that

$$F_{U_n}(u) = \left(\frac{1}{1 + \dfrac{e^{-u}}{n}} \right)^n = \left(1 + \frac{e^{-u}}{n}\right)^{-n} \to \exp\left\{-e^{-u}\right\}, \quad \text{as } n \to \infty,$$

which is a valid cdf. Hence the limiting distribution is

$$F_U(u) = \exp\left\{-e^{-u}\right\}, \quad u \in \mathbb{R}.$$

5. $Y_n = \max\{X_1, ..., X_n\}$, so

$$F_{Y_n}(y) = \{F_X(y)\}^n = \left(\frac{\lambda y}{1 + \lambda y}\right)^n, \qquad y > 0,$$

and so, in the limit as $n \to \infty$, we have the limit for *fixed* y as

$$F_{Y_n}(y) \to 0, \quad \text{for all } y.$$

Hence there is *no limiting distribution*.

$Z_n = \min\{X_1, ..., X_n\}$ so, in the limit as $n \to \infty$, we have the limit for *fixed* $z > 0$ as

$$F_{Z_n}(z) = 1 - \{1 - F_X(z)\}^n = 1 - \left(1 - \left(1 - \frac{1}{1 + \lambda z}\right)\right)^n = 1 - \frac{1}{(1 + \lambda z)^n} \to \begin{cases} 0, & z \le 0, \\ 1, & z > 0, \end{cases}$$

that is, a step function with single step of size 1 at $z = 0$. Hence the limiting random variable Z is a discrete variable with $\mathrm{P}[Z = 0] = 1$: the limiting distribution is *degenerate* at 0. Again, the limiting function is not a cdf as it not right continuous, but this does not affect our conclusion, as the limit function is not continuous at 0.

If $U_n = Y_n/n$, we have from first principles that, for $u > 0$,

$$F_{U_n}(u) = \mathrm{P}[U_n \le u] = \mathrm{P}[Y_n/n \le u] = \mathrm{P}[Y_n \le nu] = F_{Y_n}(nu) = \left(\frac{\lambda nu}{1 + \lambda nu}\right)^n,$$

so that

$$F_{U_n}(u) = \left(\frac{\lambda nu}{1 + \lambda nu}\right)^n = \left(1 + \frac{1}{n\lambda u}\right)^{-n} \to \exp\left\{-\frac{1}{\lambda u}\right\}, \quad \text{as } n \to \infty,$$

which is a valid cdf. Hence the limiting distribution is

$$F_U(u) = \exp\left\{-\frac{1}{\lambda u}\right\}, \quad u > 0.$$

If $V_n = nZ_n$, we have from first principles that, for $u > 0$,

$$F_{V_n}(v) = \mathrm{P}[V_n \le v] = \mathrm{P}[nZ_n \le v] = \mathrm{P}[Z_n \le v/n] = F_{Z_n}(v/n) = 1 - \left(\frac{1}{1 + \dfrac{\lambda v}{n}} \right)^n$$

so that

$$F_{V_n}(v) = 1 - \left(1 + \frac{\lambda v}{n}\right)^{-n} \to 1 - \exp\{-\lambda v\}, \quad \text{as } n \to \infty,$$

which is a valid cdf. Hence the limiting distribution is

$$F_V(v) = 1 - \exp\{-\lambda v\}, \quad v > 0.$$

Hence the limiting random variable $V \sim Exponential(\lambda)$.

6. $X_i \sim Poisson(\lambda)$, so $\sum_{i=1}^{n} X_i \sim Poisson(n\lambda)$ by mgfs and hence (by Q1 result) using the Central Limit Theorem,

$$\sum_{i=1}^{n} X_i \sim Normal(n\lambda, n\lambda) \quad \text{approximately,}$$

and hence

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim Normal\left(\lambda, \frac{\lambda}{n}\right) \quad \text{approximately.}$$

Hence, for $\varepsilon > 0$,

$$\mathrm{P}\left[\left|\overline{X} - \lambda\right| < \varepsilon\right] = \mathrm{P}\left[\lambda - \varepsilon < \overline{X} < \lambda + \varepsilon\right] \approx \Phi\left(\frac{\varepsilon}{\sqrt{\lambda/n}}\right) - \Phi\left(\frac{-\varepsilon}{\sqrt{\lambda/n}}\right) \to 1$$

as $n \to \infty$. Hence, $\overline{X}$ *converges in probability* to λ

$$\overline{X} \xrightarrow{p} \lambda.$$

Now, if $T_n = \exp\{-M_n\}$, then for $\varepsilon > 0$ we have

$$P\left[\left|T_n - e^{-\lambda}\right| < \varepsilon\right] = P\left[e^{-\lambda} - \varepsilon < T_n < e^{-\lambda} + \varepsilon\right] = P\left[-\log(e^{-\lambda} + \varepsilon) < M_n < -\log(e^{-\lambda} - \varepsilon)\right]$$

and hence

$$P\left[\left|T_n - e^{-\lambda}\right| < \varepsilon\right] \approx \Phi\left(\frac{-\log(e^{-\lambda} - \varepsilon) - \lambda}{\sqrt{\lambda/n}}\right) - \Phi\left(\frac{-\log(e^{-\lambda} + \varepsilon) - \lambda}{\sqrt{\lambda/n}}\right) \to 1,$$

as $n \to \infty$. Hence, T_n *converges in probability* to $e^{-\lambda}$.