M2S1 : EXERCISE SHEET 5 : SOLUTIONS

1. We have the marginal of X given in the usual way from the joint density by

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dy = \int_0^{\infty} f_{X|Y}(x|y) f_Y(y)\ dy \qquad x > 0$$

as we have the density being positive only when arguments x and y are positive. Hence

$$\begin{aligned}
f_X(x) &= \int_0^{\infty} f_{X|Y}(x|y) f_Y(y)\ dy & &= \int_0^{\infty} y e^{-yx}\ \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}\ dy \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\infty} y^{(\alpha+1)-1} e^{-(\beta+x)y}\ dy \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{(\beta+x)^{\alpha+1}} & & \text{as integrand is prop. to a } Gamma(\alpha+1, \beta+x) \text{ pdf} \\
&= \frac{\alpha\beta^\alpha}{(\beta+x)^{\alpha+1}} \qquad x > 0 & & \text{as } \Gamma(\alpha+1) = \alpha\Gamma(\alpha)
\end{aligned}$$

and hence $X \sim Pareto(\alpha, \beta)$

2. To compute the joint density f_{Y_1,Y_2}, use the multivariate transformation theorem; we have

$$\left.\begin{aligned}
Y_1 &= \mu_1 + \sigma_1\sqrt{1-\rho^2}X_1 + \sigma_1\rho X_2 \\
Y_2 &= \mu_2 + \sigma_2 X_2
\end{aligned}\right\} \Leftrightarrow \left\{\begin{aligned}
X_1 &= (Y_1-\mu_1)/(\sigma_1\sqrt{1-\rho^2}) - \rho(Y_2-\mu_2)/\left(\sigma_2\sqrt{1-\rho^2}\right) \\
X_2 &= (Y_2-\mu_2)/\sigma_2
\end{aligned}\right.$$

and hence the Jacobian $J(y_1, y_2)$ is the modulus of the determinant of the matrix of partial derivatives;

$$J(y_1,y_2) = \left|\left| \begin{bmatrix} \dfrac{1}{\sigma_1\sqrt{1-\rho^2}} & \dfrac{-\rho}{\sigma_2\sqrt{1-\rho^2}} \\ 0 & \dfrac{1}{\sigma_2} \end{bmatrix} \right|\right| = \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

as σ_1, σ_2, and $\sqrt{1-\rho^2}$ are all positive quantities. Hence the joint pdf f_{Y_1,Y_2} is given in terms of the joint pdf $f_{X_1,X_2}$by

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(x_1,x_2) J(y_1,y_2) = \frac{1}{2\pi}\exp\left\{-\frac{1}{2}\left(x_1^2+x_2^2\right)\right\} J(y_1,y_2)$$

where

$$x_1 = (y_1-\mu_1)/(\sigma_1\sqrt{1-\rho^2}) - \rho(y_2-\mu_2)/(\sigma_2\sqrt{1-\rho^2}) \qquad x_2 = (y_2-\mu_2)/\sigma_2$$

for fixed (y_1, y_2) define the inverse transformations. Now,

$$\begin{aligned}
x_1^2 + x_2^2 &= \left((y_1-\mu_1)/(\sigma_1\sqrt{1-\rho^2}) - \rho(y_2-\mu_2)/(\sigma_2\sqrt{1-\rho^2})\right)^2 + ((y_2-\mu_2)/\sigma_2)^2 \\
&= \frac{(y_1-\mu_1)^2}{\sigma_1^2(1-\rho^2)} + \frac{\rho^2(y_2-\mu_2)^2}{\sigma_2^2(1-\rho^2)} - \frac{2\rho(y_1-\mu_1)(y_2-\mu_2)}{\sigma_1\sigma_2(1-\rho^2)} + \frac{(y_2-\mu_2)^2}{\sigma_2^2} \\
&= \frac{1}{(1-\rho^2)}\left[\frac{(y_1-\mu_1)^2}{\sigma_1^2} + \frac{(y_2-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(y_1-\mu_1)(y_2-\mu_2)}{\sigma_1\sigma_2}\right]
\end{aligned}$$

and hence using the transformation result we have

$$
\begin{aligned}
f_{Y_1,Y_2}(y_1,y_2) &= f_{X_1,X_2}((y_1-\mu_1)/(\sigma_1\sqrt{1-\rho^2}) - \rho(Y_2-\mu_2)/(\sigma_2\sqrt{1-\rho^2}), (y_2-\mu_2)/\sigma_2) J(y_1,y_2) \\
&= \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(y_1-\mu_1)^2}{\sigma_1^2} + \frac{(y_2-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(y_1-\mu_1)(y_2-\mu_2)}{\sigma_1\sigma_2}\right]\right\}
\end{aligned}
$$

This is the Bivariate Normal pdf; we say that (Y_1, Y_2) have a bivariate normal distribution. Note that this function is symmetric in form; we can exchange the triples (y_1, μ_1, σ_1) and (y_2, μ_2, σ_2) without changing the functional form. Note finally that, in vector form we have the pdf in the form

$$
f_{Y_1,Y_2}(y_1,y_2) = \frac{1}{2\pi}\frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{y}-\boldsymbol{\mu})\right\} \qquad \mathbf{y} \in \mathbb{R}^2
$$

where

$$
\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \qquad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \qquad \Sigma^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix}
$$

Now, could compute the marginal pdf of Y_1 and Y_2 by integrating out from the joint pdf, for example

$$
f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1,y_2)\, dy_2
$$

where the integral can be computed as follows;

$$
\begin{aligned}
f_{Y_1}(y_1) &= \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1,y_2)\, dy_2 \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(y_1-\mu_1)^2}{\sigma_1^2} + \frac{(y_2-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(y_1-\mu_1)(y_2-\mu_2)}{\sigma_1\sigma_2}\right]\right\} dy_2
\end{aligned}
$$

Setting $t_1 = (y_1-\mu_1)/\sigma_1$ (*a constant*), and substituting $t_2 = (y_2-\mu_2)/\sigma_2$ in this integral we obtain

$$
f_{Y_1}(y_1) = \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1^2(1-\rho^2)}}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[t_1^2 + t_2^2 - 2\rho t_1 t_2\right]\right\} dt_2
$$

and can complete the square in the exponent as

$$
t_1^2 + t_2^2 - 2\rho t_1 t_2 = (t_2 - \rho t_1)^2 + t_1^2(1-\rho^2)
$$

so that

$$
\begin{aligned}
f_{Y_1}(y_1) &= \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1^2(1-\rho^2)}}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[(t_2-\rho t_1)^2 + t_1^2(1-\rho^2)\right]\right\} dt_2 \\
&= \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1^2(1-\rho^2)}} \exp\left\{-\frac{1}{2}t_1^2\right\}\int_{-\infty}^{\infty} \exp\left\{-\frac{(t_2-\rho t_1)^2}{2(1-\rho^2)}\right\} dt_2 \\
&= \frac{1}{2\pi}\frac{1}{\sqrt{\sigma_1^2(1-\rho^2)}} \exp\left\{-\frac{1}{2}t_1^2\right\}\sqrt{2\pi(1-\rho^2)}
\end{aligned}
$$

as the integrand is proportional to a Normal pdf with expectation ρt_1 and variance $(1-\rho^2)$. Therefore, cancelling terms and substituting back in for $y_1$we have

$$
f_{Y_1}(y_1) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{-\frac{1}{2}\frac{(y_1-\mu_1)^2}{\sigma_1^2}\right\}
$$

so that $Y_1 \sim Normal(\mu_1, \sigma_1^2)$. By symmetry, or by following an identical argument, we have that $Y_2 \sim Normal(\mu_2, \sigma_2^2)$.

Note also that we have that $Y_2 = \mu_2 + \sigma_2 X_2$ implies $Y_2 \sim Normal(\mu_2, \sigma_2^2)$ from elementary properties (location/scale transformations) of Normal random variables. For the conditional distributions, can use elementary properties of Normal random variables again, that is, *given* $Y_2 = y_2$ so that $X_2 = (y_2 - \mu_2)/\sigma_2$

$$Y_1 = \mu_1 + \sigma_1\sqrt{1-\rho^2}X_1 + \sigma_1\rho(y_2-\mu_2)/\sigma_2 \sim Normal(\mu_1 + \sigma_1\rho(y_2-\mu_2)/\sigma_2, \sigma_1^2(1-\rho^2))$$

that is, via a location/scale transformation $Y_1 = a + bX_1$with $a = \mu_1 + \sigma_1\rho(y_2-\mu_2)/\sigma_2$ and $b = \sigma_1\sqrt{1-\rho^2}$, and similarly for the conditional for Y_2 given $Y_1 = y_1$. Note that the conditional densities can also be computed from the definition

$$f_{Y_1|Y_2}(y_1|y_2) = \frac{f_{Y_1,Y_2}(y_1,y_2)}{f_{Y_2}(y_2)}$$

To compute the correlation, first compute the covariance using the Iterated Expectation result; we calculate

$$\mathrm{Cov}_{f_{Y_1},f_{Y_2}}[Y_1,Y_2] = \mathrm{E}_{f_{Y_1},f_{Y_2}}[Y_1Y_2] - \mathrm{E}_{f_{Y_1}}[Y_1]\,\mathrm{E}_{f_{Y_2}}[Y_2]$$

where, by the Law of Iterated Expectation

$$\mathrm{E}_{f_{Y_1},f_{Y_2}}[Y_1Y_2] = \mathrm{E}_{f_{Y_2}}\left[\mathrm{E}_{f_{Y_1|Y_2}}[Y_1Y_2|Y_2=y_2]\right]$$

and as $Y_1|Y_2 = y_2 \sim Normal(\mu_1 + \sigma_1\rho(y_2-\mu_2)/\sigma_2, \sigma_1^2(1-\rho^2))$

$$\mathrm{E}_{f_{Y_1|Y_2}}[Y_1Y_2|Y_2=y_2] = (\mu_1 + \sigma_1\rho(y_2-\mu_2)/\sigma_2)\,y_2$$

and hence

$$\begin{aligned}\mathrm{E}_{f_{Y_2}}\left[\mathrm{E}_{f_{Y_1|Y_2}}[Y_1Y_2|Y_2=y_2]\right] &= \mathrm{E}_{f_{Y_2}}[(\mu_1 + \sigma_1\rho(Y_2-\mu_2)/\sigma_2)\,Y_2] \\ &= (\mu_1 - \sigma_1\rho\mu_2/\sigma_2)\mathrm{E}_{f_{Y_2}}[Y_2] + \sigma_1\rho\mathrm{E}_{f_{Y_2}}\left[Y_2^2\right]/\sigma_2 \\ &= (\mu_1 - \sigma_1\rho\mu_2/\sigma_2)\mu_2 + \sigma_1\rho(\mu_2^2+\sigma_2^2)/\sigma_2 \\ &= \mu_1\mu_2 - \sigma_1\rho\mu_2^2/\sigma_2 + \sigma_1\rho\mu_2^2/\sigma_2 + \sigma_1\sigma_2\rho\end{aligned}$$

and hence

$$\mathrm{E}_{f_{Y_1},f_{Y_2}}[Y_1Y_2] = \mu_1\mu_2 + \sigma_1\sigma_2\rho$$

$$\mathrm{Cov}_{f_{Y_1},f_{Y_2}}[Y_1,Y_2] = \mathrm{E}_{f_{Y_1},f_{Y_2}}[Y_1Y_2] - \mathrm{E}_{f_{Y_1}}[Y_1]\,\mathrm{E}_{f_{Y_2}}[Y_2] = \mu_1\mu_2 + \sigma_1\sigma_2\rho - \mu_1\mu_2 = \sigma_1\sigma_2\rho$$

so that, finally,

$$\mathrm{Corr}_{f_{Y_1},f_{Y_2}}[Y_1,Y_2] = \frac{\mathrm{Cov}_{f_{Y_1},f_{Y_2}}[Y_1,Y_2]}{\sqrt{\mathrm{Var}_{f_{Y_1}}[Y_1]\,\mathrm{Var}_{f_{Y_2}}[Y_2]}} = \frac{\sigma_1\sigma_2\rho}{\sqrt{\sigma_1^2\sigma_2^2}} = \rho$$

3. Key is to find the i.i.d 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

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

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

(i) $X \sim Binomial(n,\theta) \Longrightarrow X = \sum\limits_{i=1}^{n} X_i$ 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

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

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

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

Note that this uses the result that the sum of independent Poisson variables also has a Poisson distribution (proved using mgfs), and also note that this is in agreement with the mgf limit result for the "standardized" Poisson example given in lectures.

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

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

(iv) $X \sim Gamma(\alpha,\beta) \Longrightarrow X = \sum\limits_{i=1}^{n} X_i$ where $X_i \sim Gamma\left(\dfrac{\alpha}{n}, \beta\right)$ so that $\mu = \mathrm{E}_{f_X}[X_i] = \dfrac{\alpha}{n\beta}$ and $\sigma^2 = \mathrm{Var}_{f_X}[X_i] = \dfrac{\alpha}{n\beta^2}$ and hence

$$Z_n = \frac{\sum\limits_{i=1}^{n} X_i - n\dfrac{\alpha}{n\beta}}{\sqrt{n\dfrac{\alpha}{n\beta^2}}} = \frac{\sum\limits_{i=1}^{n} X_i - \dfrac{\alpha}{\beta}}{\sqrt{\dfrac{\alpha}{\beta^2}}} \sim Normal(0,1) \Longrightarrow X \sim Normal\left(\frac{\alpha}{\beta}, \frac{\alpha}{\beta^2}\right) \text{ approximately}$$

Note that this uses the result that the sum of independent Gamma variables with the same second parameter also has a Gamma distribution (proved using mgfs), and also note that this is in agreement with the mgf limit result for the "standardized" Gamma example given in lectures in the special case $\beta = 1$.

.

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

$$F_{Y_n}(y) = \{F_X(y)\}^n = y^n \qquad 0 \leq y \leq 1$$

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

$$F_{Y_n}(y) \to F_Y(y) = \begin{cases} 0 & y < 1 \\ 1 & y \geq 1 \end{cases}$$

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

$$\mathrm{P}[Y = 1] = 1$$

that is, the limiting distribution is *degenerate* at 1.

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

$$F_{Z_n}(z) = 1 - \{1 - F_X(z)\}^n = 1 - (1 - z)^n \qquad 0 \le z \le 1$$

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

$$F_{Z_n}(z) \to F_Z(z) = \begin{cases} 0 & z < 0 \\ 1 & z \ge 0 \end{cases}$$

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

$$\mathrm{P}[Z = 0] = 1$$

that is, the limiting distribution is *degenerate* at 0 (as given in lectures)

Note that these results are intuitively reasonable as, 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 $1 > \varepsilon > 0$

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

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

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

$$F_{Y_n}(y) = \{F_X(y)\}^n = (1 - e^{-\lambda y})^n \qquad y > 0$$

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

$$F_{Y_n}(y) \to 0 \qquad \text{for all } y$$

Hence there is *no limiting distribution*..

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

$$F_{Z_n}(z) = 1 - \{1 - F_X(z)\}^n = 1 - (1 - (1 - e^{-\lambda y}))^n = 1 - e^{-n\lambda y} \qquad 0 \le z \le 1$$

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

$$F_{Z_n}(z) \to F_Z(z) = \begin{cases} 0 & z < 0 \\ 1 & z \ge 0 \end{cases}$$

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

$$\mathrm{P}[Z = 0] = 1$$

that is, the limiting distribution is *degenerate* at 0.

Note again that these results are intuitively reasonable as, as the sample size gets increasingly large, we will obtain a random variable arbitrarily close to the lower end of the range, but will always have a positive probability of observing a larger maximum value than had previously been observed and hence no limiting distribution.

6. $Z_n = \min\{X_1, ..., X_n\}$ 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

$$F_{Z_n}(z) \to F_Z(z) = \begin{cases} 0 & z \leq 1 \\ 1 & z > 1 \end{cases}$$

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

$$\mathrm{P}[Z = 1] = 1$$

that is, the limiting distribution is *degenerate* at 1.

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

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

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

$$F_U(u) = 1 - \frac{1}{u} \qquad u > 1$$

7. $Y_n = \max\{X_1, ..., X_n\}$ 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

$$F_{Y_n}(y) \to 0 \qquad \text{for all } y$$

Hence there is *no limiting distribution.*

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

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

so that

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

which is a valid cdf. Hence the limiting distribution is

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

8. $Y_n = \max\{X_1, ..., X_n\}$ 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

$$F_{Y_n}(y) \to 0 \qquad \text{for all } y$$

Hence there is *no limiting distribution.*

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

$$F_{Z_n}(z) = 1 - \{1 - F_X(z)\}^n = 1 - \left(1 - \left(1 - \frac{1}{1+\lambda z}\right)\right)^n = 1 - \frac{1}{(1+\lambda z)^n} \qquad z > 0$$

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

$$F_{Z_n}(z) \to F_Z(z) = \begin{cases} 0 & z \le 0 \\ 1 & z > 0 \end{cases}$$

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

$$\mathrm{P}[Z = 0] = 1$$

that is, the limiting distribution is *degenerate* 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

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

which is a valid cdf. Hence the limiting distribution is

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

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

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

which is a valid cdf. Hence the limiting distribution is

$$F_V(v) = 1 - \exp\{-\lambda v\} \qquad v > 0$$

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

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

$$F_{Y_n}(y) = \{F_X(y)\}^n = (1 - e^{-\lambda y})^n \qquad y > 0$$

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

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

so that

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

which is a valid cdf. Hence the limiting distribution is

$$F_U(u) = \exp\{-e^{-u}\} \qquad u \in \mathbb{R}$$

10. $X_i \sim Bernoulli(\theta)$ so $M_{X_i}(t) = 1 - \theta + \theta e^t$ and hence

$$M_{S_n}(t) = \{M_{X_i}(t)\}^n = \left(1 - \theta + \theta e^t\right)^n$$

so that if $M_n = S_n/n$

$$\begin{aligned} M_{M_n}(t) &= M_{S_n}(t/n) = \left(1 - \theta + \theta e^{t/n}\right)^n \\ &= \left(1 + (e^{t/n} - 1)\theta\right)^n \\ &= \left(1 + \left(\frac{t}{n} + \frac{t^2}{2n^2} + \frac{t^3}{6n^3} + \ldots\right)\theta\right)^n \\ &\to e^{\theta t} \end{aligned}$$

as $n \to \infty$. This is the mgf of a discrete variable with all probability mass at θ. Hence the limiting distribution of variable $M_n = S_n/n = \dfrac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$ (that is, the sample mean random variable) is *degenerate* at θ.

This result indicates that the sample mean is a good estimator of the parameter θ in a Bernoulli sampling experiment.

11. $X_i \sim Poisson(\lambda)$ so $\sum_{i=1}^{n} X_i \sim Poisson(n\lambda)$ by mgfs and hence (by Q2 result) using the Central Limit Theorem,

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

and hence

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

and 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

$$\overline{X} \to \lambda$$

Note that this result can also be proved using mgfs, as in Q10.

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 $e^{-\lambda}$