Solution 1

(a) The sample space is $\Omega = \{\{R,R\},\{R,W\},\{W,W\}\}$, since there is no mention of order in the sampling.

The elements of Ω are not equiprobable:

$$\mathrm{P}\{\{R,R\}\} = \frac{\binom{3}{0}\binom{2}{2}}{\binom{5}{2}} = \frac{1}{10}, \quad \mathrm{P}\{\{R,W\}\} = \frac{\binom{3}{1}\binom{2}{1}}{\binom{5}{2}} = \frac{6}{10}, \quad \mathrm{P}\{\{R,R\}\} = \frac{\binom{3}{2}\binom{2}{0}}{\binom{5}{2}} = \frac{3}{10}.$$

Let A and B denote the events 'both balls red' and 'at least one red' respectively. So, $A = \{\{R,R)\}\}$ and $B = \{\{R,W\},\{R,R\}\}$, and since $A \subset B$,

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(A)}{\mathrm{P}(B)} = \frac{1/10}{1/10 + 6/10} = \frac{1}{7}.$$

(b) For $y \in (0,1]$, $F_Y(y) = \mathrm{P}(Y \le y) = \mathrm{P}(1/X \le y) = \mathrm{P}(X \ge 1/y) = y^2$. So,

$$f_Y(y) = \frac{\partial F_Y(y)}{\partial y} = 2y, \quad 0 < y \le 1.$$

(c) Let $y_{0.5}$ be the median. Then as $X \sim U(a,b)$,

$$\tfrac{1}{2} = \mathrm{P}(Y \le y_{0.5}) = \mathrm{P}(X \le \log y_{0.5}) = \mathrm{P}\{X \le (a+b)/2\},$$

so $y_{0.5} = \exp\{(a+b)/2\}$.

(d) $Y = \cos X$ takes values only in the range $-1 \le y \le 1$, so

$$F_Y(y) = \begin{cases} 0, & y < -1 \\ 1, & y \ge 1. \end{cases}$$

A sketch of $\cos x$ for $x \ge 0$ shows that in the range $0 < x < 2\pi$, and for $-1 < y < 1$, we have $\cos X \le y \Leftrightarrow \cos^{-1} y \le X \le 2\pi - \cos^{-1} y$. Since the cosine function is periodic, we therefore have

$$\cos X \le y \quad \Leftrightarrow \quad X \in \bigcup_{j=0}^{\infty} \{x : 2\pi j + \cos^{-1} y \le x \le 2\pi(j+1) - \cos^{-1} y\},$$

and thus

$$\begin{aligned} \mathrm{P}(Y \le y) &= \sum_{j=0}^{\infty} \mathrm{P}\left\{2\pi j + \cos^{-1} y \le X \le 2\pi(j+1) - \cos^{-1} y\right\} \\ &= \sum_{j=0}^{\infty} \left(\exp[-\lambda\{2\pi j + \cos^{-1} y\}] - \exp[-\lambda\{2\pi(j+1) - \cos^{-1} y\}]\right) \\ &= \frac{\exp(-\lambda \cos^{-1} y) - \exp(\lambda \cos^{-1} y - 2\pi\lambda)}{1 - \exp(-2\pi\lambda)}, \end{aligned}$$

where we noticed that the summation is proportional to a geometric series.

If $y = 1$, then $\cos^{-1} y = 0$, and so $\mathrm{P}(Y \le 1) = 1$, and if $y = -1$, then $\cos^{-1} y = \pi$, and then $\mathrm{P}(Y \le -1) = 0$, as required. Here we used values of $\cos^{-1} y$ in the range $[0, \pi]$.

The density function is found by differentiation: since $\cos(\cos^{-1} y) = y$, we have

$$\frac{d \cos^{-1} y}{dy} = -\frac{1}{\sin(\cos^{-1} y)},$$

and this gives

$$f_Y(y) = \frac{\lambda}{\sin(\cos^{-1} y)} \times \frac{\exp(-\lambda \cos^{-1} y) + \exp(\lambda \cos^{-1} y - 2\pi\lambda)}{1 - \exp(-2\pi\lambda)}, \quad y \in (-1, 1).$$

(e) We have

$$\mathrm{P}(X<3)=\mathrm{P}(X=0)+\mathrm{P}(0<X<3\mid X>0)\mathrm{P}(X>0)=(1-p)+p(1-e^{-3\lambda});$$

equivalently $\mathrm{P}(X<3)=1-\mathrm{P}(X\geq 3)=1-pe^{-3\lambda}$. Hence $\mathrm{P}(X=0\mid X<3)=(1-p)/(1-pe^{-3\lambda})$. The mean and variance of X, p/λ and $(2-p)p/\lambda^2$, can be computed by noting that

$$\mathrm{E}(X^r)=\mathrm{E}(X^r\mid X=0)\mathrm{P}(X=0)+\mathrm{E}(X^r\mid X>0)\mathrm{P}(X>0)=0^r(1-p)+\lambda^{-r}\Gamma(r+1)p,$$

or via the moment-generating function,

$$M_X(t)=\mathrm{E}(e^{tX})=\mathrm{E}(e^{tX}\mid X=0)\mathrm{P}(X=0)+\mathrm{E}(e^{tX}\mid X>0)\mathrm{P}(X>0)=1-p+p\frac{\lambda}{\lambda-t},\quad t<\lambda.$$

(f) Since all the entries of the table must sum to 1, we must have $c=1/12$. Hence

$$\mathrm{E}(X)=1\cdot\tfrac{4}{12}+3\cdot\tfrac{4}{12}+5\cdot\tfrac{4}{12}=\tfrac{36}{12}=3,$$

and the conditional expectation is

$$\mathrm{E}(X\mid Y=4)=\frac{1\times 3/12+3\times 2/12+5\times 1/12}{6/12}=\frac{14}{6}.$$

The random variables X and Y are clearly dependent, since $\mathrm{P}(X=1\mid Y=2)\neq\mathrm{P}(X=1)$.

(g) We can write

$$X_{n\times 1}=(X_1,\ldots,X_n)^{\mathrm{T}}=(Y+Z_1,\ldots,Y+Z_n)^{\mathrm{T}}=B_{n\times(n+1)}(Y,Z_1,\ldots,Z_n)^{\mathrm{T}},$$

where $(Y,Z_1,\ldots,Z_n)^{\mathrm{T}}\sim N_{n+1}\{(\mu,a_1,\ldots,a_n)^{\mathrm{T}},\mathrm{diag}(\sigma^2,1,\ldots,1)\}$. Hence X has a joint normal distribution (see slide 18), and it is easy to check that $\mathrm{E}(X)=(\mu+a_1,\ldots,\mu+a_n)^{\mathrm{T}}$ and $\mathrm{var}(X)=I_n+\sigma^2 1_n 1_n^{\mathrm{T}}$. This distribution is finitely exchangeable if we have constant $\mathrm{var}(X_j)$, constant $\mathrm{cov}(X_j,X_k)$ for $j\neq k$, which are both true, and equal means; the latter occurs if and only if all the a_j are equal.

(h) Since $X_1,\ldots,X_N\overset{\text{iid}}{\sim}\mathrm{Poiss}(\lambda)$, each indicator variable $I(X_j=0)$ is Bernoulli with success probability $q=e^{-\lambda}$. So, conditional on $N=n$, $T\sim\mathrm{Bin}(n,q)$, which gives $\mathrm{E}(T\mid N=n)=nq$ and $\mathrm{var}(T\mid N=n)=nq(1-q)$. As N has a geometric distribution with success probability p, we have (by direct calculation or looking online ...) $\mathrm{E}(N)=1/p$ and $\mathrm{var}(N)=(1-p)/p^2$, and thus

$$\begin{aligned}
\mathrm{E}(T) &= \mathrm{E}_N\mathrm{E}(T\mid N)=\mathrm{E}_N(Nq)=q/p,\\
\mathrm{var}(T) &= \mathrm{E}_N\mathrm{var}(T\mid N)+\mathrm{var}_N\mathrm{E}(T\mid N)\\
&= \mathrm{E}_N\{Nq(1-q)\}+\mathrm{var}_N(Nq)=q(1-q)/p+q^2(1-p)/p^2.
\end{aligned}$$

If the X_j's are dependent, $\mathrm{E}(T)$ is unchanged because expectation is a linear operator.

(i) The central limit theorem implies that $\overline{X}\,\dot\sim\,\mathcal{N}(\mu,\mu/n)$, so applying the delta method with $g(u)=2\sqrt{u}$, giving $g'(u)=u^{-1/2}$, leads to

$$Y=g(\overline{X})\,\dot\sim\,\mathcal{N}\left\{g(\mu),g'(\mu)^2\times\mu/n\right\}=\mathcal{N}(2\sqrt{\mu},1/n),\quad n\to\infty.$$

Thus the variance of Y does not depend on μ, at least to this order of approximation: the square root transformation is variance-stabilizing for the Poisson distribution. The related *Anscombe transformation* $Y=(4\overline{X}+3/2)^{1/2}$ is widely used in certain imaging settings.

(j) As $n \to \infty$, the weak law of large numbers gives (in the usual notation) $\overline{X} \xrightarrow{P} \mu_X$ and $\overline{Y} \xrightarrow{P} \mu_Y$, so as these convergences are also in distribution, Slutsky's theorem gives $T \xrightarrow{D} \mu_Y/\mu_X = \theta$. As this is convergence in distribution to a constant, we also have $T \xrightarrow{P} \theta$. The (multivariate) central limit theorem gives

$$n^{1/2}\left\{\begin{pmatrix}\overline{X}\\ \overline{Y}\end{pmatrix} - \begin{pmatrix}\mu_X\\ \mu_Y\end{pmatrix}\right\} \xrightarrow{D} \mathcal{N}\left\{\begin{pmatrix}0\\ 0\end{pmatrix}, \begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}\right\}, \quad n \to \infty,$$

and as $\theta = g(\mu_X, \mu_Y) = \mu_Y/\mu_X$ has derivatives $g_1(\mu_X, \mu_Y) = -\mu_Y/\mu_X^2 = -\theta/\mu_X$ and $g_2(\mu_X, \mu_Y) = 1/\mu_X$, simplifying the variance expression

$$\sigma_\theta^2 = \mu_X^{-2}(-\theta, 1)\begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}\begin{pmatrix}-\theta\\ 1\end{pmatrix}$$

and an application of the delta method gives the result.

σ_θ^2 could be estimated by replacing μ_Y by $\overline{Y}$, σ_Y^2 by $n^{-1}\sum_j (Y_j - \overline{Y})^2$, $\rho\sigma_X\sigma_Y$ by $n^{-1}\sum_j (X_j - \overline{X})(Y_j - \overline{Y})$, etc.; note that as the variances are finite, all these expressions will converge in probability to the corresponding theoretical quantities.

(k) (i) The random variable $T = X_1 + X_2$ follows a normal distribution since a linear combination of normal variables is normal, and $\mathrm{E}(T) = \mathrm{E}(X_1 + X_2) = \mathrm{E}(X_1) + \mathrm{E}(X_2) = 8 + 16 = 24$ and $\mathrm{var}(T) = \mathrm{var}(X_1) + \mathrm{var}(X_2) = 9 + 16 = 25$, so $T \sim \mathcal{N}(24, 5^2)$.

(ii) The random variable $Z = (T - 24)/5 \sim \mathcal{N}(0, 1)$, so

$$\mathrm{P}(T > 30) = \mathrm{P}\left(Z > \tfrac{30-24}{5}\right) = 1 - \mathrm{P}(Z \leq 1.2) = 1 - \Phi(1.2) = 1 - 0.88493 \approx 0.115,$$

where $\Phi(\cdot)$ denotes the standard normal CDF.

(iii) The probability that the total download time T exceeds 30 minutes given that $X_1 = 10$ is

$$\mathrm{P}(T > 30 \mid X_1 = 10) = \mathrm{P}(X_1 + X_2 > 30 \mid X_1 = 10) = \mathrm{P}(X_2 > 20 \mid X_1 = 10) = \mathrm{P}(X_2 > 20)$$

by the independence of X_1 and X_2. The random variable $Z_2 = (X_2 - 16)/4$ follows a standard normal distribution, so

$$\mathrm{P}(X_2 > 20) = 1 - \mathrm{P}\left(Z_2 \leq \tfrac{20-16}{4}\right) = 1 - \Phi(1) = 1 - 0.84134 \approx 0.152.$$

(iv) $Y = (X_1, T)^{\mathrm{T}} = (X_1, X_1 + X_2)^{\mathrm{T}}$ is a linear combination of normal variables, so it has a bivariate normal distribution, with mean and covariance matrix

$$\mu = \begin{pmatrix}\mathrm{E}(X_1)\\ \mathrm{E}(T)\end{pmatrix} = \begin{pmatrix}8\\ 24\end{pmatrix}, \quad \Omega = \begin{pmatrix}\mathrm{var}(X_1) & \mathrm{cov}(X_1, T)\\ \mathrm{cov}(X_1, T) & \mathrm{var}(T)\end{pmatrix} = \begin{pmatrix}9 & 9\\ 9 & 25\end{pmatrix},$$

since $\mathrm{cov}(X_1, T) = \mathrm{var}(X_1) = 9$ by the independence of X_1 and X_2. Thus

$$\begin{pmatrix}X_1\\ T\end{pmatrix} \sim \mathcal{N}_2\left\{\begin{pmatrix}8\\ 24\end{pmatrix}, \begin{pmatrix}9 & 9\\ 9 & 25\end{pmatrix}\right\}.$$

Now $X_1 \mid T = 30 \sim \mathcal{N}(\mu, \sigma^2)$ where $\mu = 8 + 9 \times (30 - 24)/25 = 10.16$ and $\sigma^2 = 9 - 9^2/25 = 2.4^2$. Thus, $Z_1 = (X_1 - \mu)/\sigma \mid T = 30 \sim \mathcal{N}(0, 1)$, so

$$\mathrm{P}(X_1 < 7 \mid T = 30) = \mathrm{P}\left(Z_1 < \tfrac{7-10.16}{2.4}\right) \approx \Phi(-1.32) = 1 - \Phi(1.32) = 1 - 0.90658 \approx 0.093.$$

(l) Let X_1, X_2, X_3 denote the arrival times after 19.00; as they arrive independently at random, we can suppose that $X_1, X_2, X_3 \stackrel{\text{iid}}{\sim} U(0,1)$. The arrival times of the first and the last are $U = \min(X_1, X_2, X_3)$ and $V = \max(X_1, X_2, X_3)$. The densities can be obtained directly or ...

As the X_j have distribution function $F(u) = u$ in $0 < u < 1$, we obtain

$$\begin{aligned} \mathrm{P}(U \leq u) &= 1 - \mathrm{P}(X_1 > u, X_2 > u, X_3 > u) \\ &= 1 - \mathrm{P}(X_1 > u)\mathrm{P}(X_2 > u)\mathrm{P}(X_3 > u) \\ &= 1 - (1-u)^3, \quad 0 < u < 1, \end{aligned}$$

and the density is $f_U(u) = 3(1-u)^2$, for $0 < u < 1$. Likewise

$$\mathrm{P}(V \leq v) = \mathrm{P}(X_1 \leq v)\mathrm{P}(X_2 \leq v)\mathrm{P}(X_3 \leq v) = v^3, \quad 0 < v < 1,$$

and the density is $f_V(v) = 3v^2$, for $0 < v < 1$.

For the final part, we seek $\mathrm{E}(V - U) = \mathrm{E}(V) - \mathrm{E}(U)$, and

$$\mathrm{E}(V) = \int_0^1 v f_V(v) \mathrm{d}v = \tfrac{3}{4}, \quad \mathrm{E}(U) = \int_0^1 u f_u(u) \mathrm{d}u = \int_0^1 3u(1-u)^2 \mathrm{d}u = \tfrac{1}{4},$$

using integration by parts. Thus $\mathrm{E}(V - U) = 2/4$, or 30 minutes.