Exercice 0.1

(a) **(2 points)** 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 :

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

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 $A \subset B$,

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

(b) **(2 points)** For $y \in (0,1]$, $F_Y(y) = \Pr(Y \le y) = \Pr(1/X \le y) = \Pr(X \ge 1/y) = y^2$. So,

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

(c) **(2 points)** Let $y_{0.5}$ be the median. Then as $X \sim U(a,b)$,

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

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

(d) **(2 points)** Note first that

$$\Pr(X > 6) = 1 - \Pr\left(\tfrac{X-2}{\sigma} < \tfrac{6-2}{\sigma}\right) = 1 - \Phi(4/\sigma).$$

Using the table for the standard normal distribution, we can easily see that

$$\Pr(X > 6) = 0.1 \Rightarrow \Phi^{-1}(0.9) = 4/\sigma \Rightarrow 1.282 = 4/\sigma \Rightarrow \sigma \approx 3.12.$$

Hence

$$\Pr(X < 0) = \Pr\left(\tfrac{X-2}{3.12} < \tfrac{0-2}{3.12}\right) = \Phi(-2/3.12) \approx \Phi(-0.64) \approx 0.26.$$

(e) **(3 points)** 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 $\Pr(X = 1 \mid Y = 2) \neq \Pr(X = 1)$. We accept other similar arguments.

(f) **(2 points)** The density of X is $f(x) = \lambda \exp(-\lambda x)$, for $x > 0$ and $\lambda > 0$, so

$$M_X(t) = \mathrm{E}\left\{\exp(tX)\right\} = \int_0^\infty \exp(tx)\lambda\exp(-\lambda x)\mathrm{d}x = \frac{\lambda}{\lambda - t}, \quad t < \lambda.$$

Hence independence of $X_1, \ldots, X_n$ gives

$$M_S(t) = \mathrm{E}\left\{\exp(tS)\right\} = \mathrm{E}\left\{\exp(tX_1 + \cdots + tX_n)\right\} = \prod_{j=1}^n \mathrm{E}\left\{\exp(tX_j)\right\} = \frac{\lambda^n}{(\lambda - t)^n}, \quad t < \lambda.$$

(g) **(3 points)** The null hypothesis $\mathcal{H}_0$ represents the theory/model we want to test. The test statistic T is chosen such that large values of T provide evidence against $\mathcal{H}_0$. The p-value $= p_{obs} = \Pr_0(T \geq t_{obs})$, where t_{obs} is the observed t-statistic and $\Pr_0$ is the probability under the null. Thus the p-value is the probability that we observe a value of T bigger than or equal to t_{obs}, when the null hypothesis is true. For the given situation, we reject H_0 at level α for all $\alpha > 0.001$, which suggests that H_0 is rejected at the conventional levels 0.05, 0.01.

(h) **(2 points)** The density function is $f(x;\alpha) = \alpha 4^\alpha / x^{\alpha+1}$, for $x > 4$ and $\alpha > 0$. Hence the likelihood is

$$L(\alpha) = \prod_{j=1}^n \frac{\alpha 4^\alpha}{x_j^{\alpha+1}}, \quad \alpha > 0,$$

and the log likelihood is given by

$$\ell(\alpha) = n\left\{\log\alpha + \alpha\log 4 - (\alpha+1)s\right\}, \quad \alpha > 0, \quad s = n^{-1}\sum_j \log x_j,$$

and has derivatives

$$\ell'(\alpha) = n\left\{1/\alpha + \log 4 - s\right\}, \quad \ell''(\alpha) = -n/\alpha^2.$$

The second derivative is always negative, so ℓ is concave, and therefore the solution to $\ell'(\alpha) = 0$ gives the unique and global maximum. This is $\widehat{\alpha} = 1/(s - \log 4)$.

(i) **(2 points)** Since $X_1, \ldots, X_n \overset{\text{iid}}{\sim}$ Poiss(λ), each indicator variable $I(X_j = 0)$ is Bernoulli with success probability $p = e^{-\lambda}$. So, $R = nT \sim$ Bin($n, e^{-\lambda}$), as it is just a sum of i.i.d Bernoulli random variables. Hence,

$$\mathrm{E}(T) = \tfrac{1}{n}\mathrm{E}\left\{\sum_{j=1}^n I(X_j = 0)\right\} = \tfrac{1}{n}np = e^{-\lambda}$$

and (quickly) $\mathrm{var}(T) = \mathrm{var}(R/n) = n^{-2}\mathrm{var}(R) = np(1-p)/n^2 = e^{-\lambda}(1 - e^{-\lambda})/n$, or (in more detail)

$$\mathrm{var}(T) = \tfrac{1}{n^2}\mathrm{var}\left\{\sum_{j=1}^n I(X_j = 0)\right\} = \tfrac{1}{n}\mathrm{var}\Big(I(X_1 = 0)\Big) = \frac{e^{-\lambda}(1 - e^{-\lambda})}{n},$$

where the second equality holds because of the independence assumption. If the X_j's become dependent, the expectation of T remains the same because the expectation is a linear operator.

Exercice 0.2 (a) **(3 points)** We use Bayes' theorem. Let A_r be the event that there are r 9s in a row, and let M be the event that the generator is bad. Then we seek r such that

$$q \le \Pr(M \mid A_r) = \frac{\Pr(A_r \mid M)\Pr(M)}{\Pr(A_r \mid M)\Pr(M) + \Pr(A_r \mid M^c)\Pr(M^c)} = \frac{1 \times p}{1 \times p + (1-p)/10^r},$$

or equivalently

$$q^{-1} \ge 1 + (1-p)10^{-r}/p,$$

i.e.,

$$r \ge \frac{\log\left[(1-p)q/\{(1-q)p\}\right]}{\log 10}.$$

(b) **(5 points)** Let $f(x) = \binom{9}{x} a^x (1-a)^{9-x}$. Then by independence of X and U,

$$\Pr(X = x, U \le u) = \Pr(X = x)\Pr(U \le u) = \frac{1}{10} \times u, \quad x \in \{0, \ldots, 9\}, 0 < u < 1,$$

so because $0 \le f(x) \le 1$ for each $x \in \{0, \ldots, 9\}$, we have

$$\begin{aligned} \Pr\{U \le f(X), X = x\} &= \Pr\{U \le f(X) \mid X = x\} \times \Pr(X = x) \\ &= \Pr\{U \le f(x)\} \times \Pr(X = x) \\ &= \binom{9}{x} a^x (1-a)^{9-x} \times \frac{1}{10}, \end{aligned}$$

so the acceptance probability is

$$\Pr\{U \le f(X)\} = \sum_{x=0}^{9} \Pr\{U \le f(X), X = x\} = \sum_{x=0}^{9} \binom{9}{x} a^x (1-a)^{9-x} \times \frac{1}{10} = \frac{1}{10}.$$

Hence

$$\Pr(R = x) = \Pr\{X = x \mid U \le f(X)\} = \binom{9}{x} a^x (1-a)^{9-x}, \quad x \in \{0, \ldots, 9\},$$

and this implies that $R \sim B(9, a)$, as required.

We could accept alternative arguments here, as they've never seen anything like this before.

(c) **(2 points)** The acceptance probability is $1/10$ and iterations are independent, so N is geometric with success probability $1/10$ and hence mean 10.

Exercice 0.3 **(a) (2 points)** The random variable $T = X_1 + X_2$ follows a normal distribution since both X_1 and X_2 are normally distributed and the sum of normally distributed random variables is normally distributed. We have

$$\mathrm{E}(T) = \mathrm{E}(X_1 + X_2) = \mathrm{E}(X_1) + \mathrm{E}(X_2) = 8 + 16 = 24,$$

and by the independence of X_1 and X_2,

$$\mathrm{var}(T) = \mathrm{var}(X_1 + X_2) = \mathrm{var}(X_1) + \mathrm{var}(X_2) = 9 + 16 = 25.$$

So, $T \sim \mathcal{N}(24, 5^2)$.

(b) (2 points) From (a), the random variable $Z = (T - 24)/5$ follows a standard normal distribution. Then, the probability that the total download time exceeds 30 minutes is

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

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

(c) (2 points) The probability that the total download time T exceeds 30 minutes given that $X_1 = 10$ is

$$\begin{aligned}\Pr(T > 30 \mid X_1 = 10) &= \Pr(X_1 + X_2 > 30 \mid X_1 = 10) \\ &= \Pr(X_2 > 20 \mid X_1 = 10) = \Pr(X_2 > 20)\end{aligned}$$

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

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

Alternatively : Similarly to the development in (d), we have $T \mid X_1 = 10 \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$ where $\tilde{\mu} = 24 + 9 \times (10 - 8)/9 = 26$ and $\tilde{\sigma}^2 = 25 - 9^2/9 = 4^2$. Thus, $Z_T = (T - \tilde{\mu})/\tilde{\sigma} \mid X_1 = 10 \sim \mathcal{N}(0, 1)$, and therefore,

$$\Pr(T > 30 \mid X_1 = 10) = \Pr\left(Z_T > \tfrac{30-26}{4}\right) = 1 - \Phi(1) = 1 - 0.84134 \approx 0.152.$$

(d) (4 points) The random vector $Y = (X_1, T)^{\mathrm{T}} = (X_1, X_1 + X_2)^{\mathrm{T}} = B(X_1, X_2)^{\mathrm{T}}$ 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 the hint in the question gives $\mu = 8 + 9 \times (30 - 24)/25 = 10.16$ 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

$$\begin{aligned}\Pr(X_1 < 7 \mid T = 30) &= \Pr\left(Z_1 < \tfrac{7-10.16}{2.4}\right) \approx \Phi(-1.32) = 1 - \Phi(1.32) \\ &= 1 - 0.90658 \approx 0.093.\end{aligned}$$

Exercice 0.4 (a) **(2 points)** There should be

(i) a description of the whiskers : the maximum is higher and the minimum is lower in group 1

(ii) a description of the inter-quartile range (IQR) : group 1 has a higher IQR

(iii) a description of the variability of both groups : group 1 has higher variability while group 2's marks are more concentrated

(iv) the median : group 1 has a slightly lower median.

(b) **(3 points)** 1 marks should be reserved for the description of group 1's Q-Q plot. The graph is not close to a straight line, which seems to suggest that the observations may not be well fitted by a normal model. The slope and the intercept of the best fitted straight line at $x = 0$ give estimates of σ and μ respectively, which should be close to 1.06 and 3.66. The Q-Q plot drawn for group 2 should have a best fitted straight line with a slope of approximately 0.75 and an intercept of 4 at $x = 0$.

(c) **(3 points)** Using the assumptions given in the question, we have

$$\sigma^2_{\text{diff}} = \text{var}(\overline{x}_1 - \overline{x}_2) = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} = \frac{1.06^2}{100} + \frac{0.75^2}{80} \approx 0.018267 \approx 0.135^2$$

Since the estimated mean difference in marks is $3.66 - 4.05 = -0.39$, an approximate 95% confidence interval for the mean difference in marks is

$$-0.39 \pm z_{0.975} \cdot \sigma_{\text{diff}} = [-0.39 - 1.96 \cdot 0.135, -0.39 + 1.96 \cdot 0.135] = [-0.645, -0.125],$$

where $z_{0.975}$ is the 97.5% quantile of the standard normal distribution. The 95% confidence interval for the mean difference in marks does not contain 0, so we reject the null hypothesis of equal group means at the 5% level.

(d) **(2 points)** The assumptions for the calculation in (b) are between-group independence (to argue that $\text{var}(\overline{x}_1 - \overline{x}_2) = s_1^2/n_1 + s_2^2/n_2$) and within-group independence, so that a normal approximation from the CLT applies to the group means. If every mark from the exam is independent, both of these assumptions will hold. This is not an unrealistic assumption for exam marks, which clearly have a finite mean and variance.

[A very critical student might question from what populations the two groups are drawn : is this test result applicable only to these particular students, and if so, what does it tell us more generally, if anything ? If it applies only to these students, then what CLT applies (there are no larger population parameters for the averages and sample variances to tend to ...) ? Good questions for which a bonus should be given.]