

Exercise 0.1 (a) **(4 points)** We can write

$$\Omega = \{(d_1, d_2, d_3) : d_1, d_2, d_3 \in \{1, \dots, 6\}\} = \{\omega_1, \dots, \omega_{216}\},$$

say, with $\omega_1 = (1, 1, 1), \omega_2 = (1, 1, 2), \dots, \omega_{216} = (6, 6, 6)$.

The list of all possibilities for getting 9 is every permutation of the three elements of

$$(1, 2, 6), \quad (1, 3, 5), \quad (1, 4, 4), \quad (2, 2, 5), \quad (2, 3, 4), \quad (3, 3, 3),$$

and with x, y, z all distinct the number of distinct permutations of the ordered triple (x, x, x) is 1, that of (x, y, y) is 3, and that of (x, y, z) is 6.

Hence

$$\Pr(X = 9) = (6 + 6 + 3 + 3 + 6 + 1)/216 = 25/216.$$

Similarly, the list of possibilities for getting 10 is

$$(1, 3, 6), \quad (1, 4, 5), \quad (2, 2, 6), \quad (2, 3, 5), \quad (2, 4, 4), \quad (3, 3, 4),$$

and the associated probability is

$$\Pr(X = 10) = (6 + 6 + 3 + 6 + 3 + 3)/216 = 27/216.$$

Hence it is more probable that $X = 10$ than that $X = 9$.

(b) **(2 points)**

$$\begin{aligned} \Pr(W \leq w) &= \Pr(\min(X_1^3, \dots, X_n^3) \leq w), \\ &= 1 - \Pr(\min(X_1^3, \dots, X_n^3) > w), \\ &= 1 - \Pr(X_i^3 > w)^n, \\ &= 1 - \{1 - \Pr(X_i \leq w^{1/3})\}^n, \\ &= 1 - \exp^{-n\lambda w^{1/3}}, \quad w \geq 0 \end{aligned}$$

The density is found by differentiation :

$$f_W(w) = \begin{cases} \frac{n\lambda}{3w^{2/3}} \exp^{-n\lambda w^{1/3}}, & w \geq 0, \\ 0, & w < 0. \end{cases}$$

(c) **(2 points)** Note that $\min(u, 1 - u) = u$ if $0 < u < 1/2$ and otherwise equals $1 - u$, and that the restriction of a uniform density to an interval remains uniform on that interval. Now

$$\begin{aligned} \Pr(P \leq x) &= \Pr\{2 \min(U, 1 - U) \leq x\} \\ &= \Pr\{\min(U, 1 - U) \leq x/2\} \\ &= \Pr(0 \leq U \leq x/2 \mid U \leq 1/2) \Pr(U \leq 1/2) \\ &\quad + \Pr(1 - x/2 \leq U \leq 1 \mid 1/2 \leq U \leq 1) \Pr(1/2 \leq U \leq 1), \\ &= \frac{x/2}{1/2} \times \frac{1}{2} + \frac{1 - (1 - x/2)}{1/2} \times \frac{1}{2} \\ &= x, \end{aligned}$$

so $P = \min(U, 1 - U) \sim U(0, 1)$.

- (d) **(2 points)** For any continuous distribution,

$$\Pr(x_{0.1} < X \leq x_{0.95}) = \Pr(X \leq x_{0.95}) - \Pr(X \leq x_{0.1}) = 0.95 - 0.1 = 0.85.$$

- (e) **(2 points)** We need $\text{cov}(X, Y)$, $\text{var}(X)$ and $\text{var}(Y)$, and as the additive constant can be ignored and the Z s are independent and thus have zero covariances, we obtain

$$\begin{aligned}\text{cov}(X, Y) &= \text{cov}(b_1 Z_1 + c_1 Z_3, b_2 Z_2 + c_2 Z_3) = c_1 c_2 \text{var}(Z_3) = c_1 c_2, \\ \text{var}(X) &= \text{var}(b_1 Z_1 + c_1 Z_3) = b_1^2 + c_1^2, \\ \text{var}(Y) &= \text{var}(b_2 Z_2 + c_2 Z_3) = b_2^2 + c_2^2,\end{aligned}$$

which gives

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\{\text{var}(X) \text{var}(Y)\}^{1/2}} = \frac{c_1 c_2}{\{(b_1^2 + c_1^2)(b_2^2 + c_2^2)\}^{1/2}}.$$

- (f) **(2 points)** Soit X le nombre d'essais jusqu'au premier succès dans une suite d'essais de Bernoulli indépendants avec probabilité p de succès. Alors si l'on conditionne sur le résultat du premier essai et utilisant la manque de mémoire on a

$$\begin{aligned}\mathbb{E}(X) &= \Pr(X = 1) \times 1 + \Pr(X > 1) \mathbb{E}(X \mid X > 1), \\ &= p + (1 - p) \mathbb{E}(X + 1), \\ &= p + (1 - p) \{1 + \mathbb{E}(X)\},\end{aligned}$$

et ainsi $\mathbb{E}(X) = 1/p$.

- (g) **(3 points)**

If $X \sim \mathcal{N}(0, \sigma^2)$, then using the fact that $\phi'(x) = -x\phi(x)$ we have

$$\mathbb{E}(|X|) = \int_{-\infty}^{\infty} |x| \phi(x) \, ds = 2 \int_0^{\infty} x \phi(x) \, dx = 2 [-\phi(x)]_0^{\infty} = 2\phi(0) = \sqrt{\frac{2}{\pi}} \sigma.$$

If $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$ we therefore see that

$$\tilde{\sigma} = (\pi/2)^{1/2} n^{-1} \sum_{j=1}^n |X_j|$$

is an unbiased estimator of σ .

- (h) **(2 points)** Un intervalle de confiance bilatéral, de la forme (L, U) , est une statistique sous forme d'intervalle qui contient R_e avec un probabilité spécifiée. Cette probabilité s'appelle le niveau de l'intervalle, pour cet exemple $\alpha = 0.95$.

Exercice 0.2 1. **(2 points)** We first obtain c :

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{S,T}(s,t) dt ds = 1 &\iff c \int_0^{\infty} \left[-(1+s)^{-3} e^{-t/(1+s)} \right]_{t=0}^{t=\infty} ds = 1 \\ &\iff c \int_0^{\infty} (1+s)^{-3} ds = c \left[-\frac{1}{2} (1+s)^{-2} \right]_0^{\infty} = 1, \end{aligned}$$

i.e., $c = 2$. The marginal density of S is therefore

$$f_S(s) = \int_{-\infty}^{\infty} f_{S,T}(s,t) dt = 2(1+s)^{-3}, \quad s > 0,$$

using the previous computations. Download time depends on file size, as

$$f_{T|S}(t | s) = \frac{f_{S,T}(s,t)}{f_S(s)} = \frac{1}{1+s} \cdot e^{-t/(1+s)}, \quad t > 0, s > 0,$$

depends on s and hence does not equal the marginal density $f_T(t)$.

(It helps below if we notice that this conditional density is exponential with mean $1 + s$ minutes.)

2. **(2 points)** The mean file size is

$$\begin{aligned} E(S) &= \int_{-\infty}^{\infty} s f_S(s) ds = \int_0^{\infty} 2s(1+s)^{-3} ds \\ &= \left[2s \cdot \frac{-1}{2} (1+s)^{-2} \right]_0^{\infty} - \int_0^{\infty} 2 \frac{-1}{2} (1+s)^{-2} ds \\ &= \left[-1 \cdot (1+s)^{-1} \right]_0^{\infty} = 1 \text{ GB.} \end{aligned}$$

3. **(2 points)** Using (a), we get

$$f_{T|S}(t | s = 1) = \frac{1}{2} \cdot e^{-t/2}, \quad t > 0,$$

which is exponential with mean 2 minutes.

4. **(2 points)** If we recognise the $\exp(1/3)$ density, we easily see that the conditional expectation is 3 minutes. Otherwise, we have to compute (integration by parts)

$$E(T | S = 2) = \int_{-\infty}^{\infty} t f_{T|S}(t | s = 2) dt = \int_0^{\infty} \frac{t}{3} \cdot e^{-t/3} dt = \left[\frac{t}{3} \cdot (-3) e^{-t/3} \right]_0^{\infty} - \int_0^{\infty} \frac{1}{3} \cdot (-3) e^{-t/3} dt = 3.$$

5. **(2 points)** Since (a) implies that $E(T | S = s) = 1 + s$, we get from (b) that

$$E(T) = E_S\{E(T | S)\} = E_S(1 + S) = 1 + E(S) = 2 \text{ minutes.}$$

Otherwise one can do the integral

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t f_{S,T}(s,t) dt ds = 2 \int_0^{\infty} \int_0^{\infty} t (1+s)^{-4} e^{-t/(1+s)} dt ds = \dots = 2,$$

after some struggle.

- Exercice 0.3**
1. **(2 points)** 0 n'est pas dans l'IC à 95% donc un départ à l'heure (retard nul) n'est pas dans les normes (au seuil 95%).
 2. **(3 points)** 0 est dans l'IC donc un départ à l'heure serait dans les normes, bien que l'IC contient des valeurs négatives! Mais il semble bizarre qu'en remplaçant 16 par 160 (un retard plus grand) on peut conclure qu'un retard moyen de zero seconds devient plus credible.
 3. **(4 points)** Description des éléments d'un QQplot.
L'hypothèse de normalité n'est pas bonne car un retard ne peut pas être négatif. Il faudrait utiliser une loi ayant pour support R^+ . On pourrait par exemple utiliser la loi exponentielle qui modélise habituellement les temps d'attente.

Exercice 0.4 1. **(4 points)** La vraisemblance est

$$L(\theta; x_1, \dots, x_k) = \prod_{i=1}^k \binom{n}{x_i} \theta^{x_i} (1-\theta)^{n-x_i}, \quad x_i = 0, 1, \dots, n, \quad 0 < \theta < 1$$

et la log-vraisemblance

$$\begin{aligned} \ell(\theta) &= \sum_{i=1}^k \log \binom{n}{x_i} + \sum_{i=1}^k \{x_i \log \theta + (n-x_i) \log(1-\theta)\}, \quad x_i = 0, 1, \dots, n, \quad 0 < \theta < 1 \\ &= \log \theta \sum_{i=1}^k x_i + \log(1-\theta) \sum_{i=1}^k (n-x_i) + \text{const.} \end{aligned}$$

La dérivée par rapport à θ est

$$\ell'(\theta) = \frac{1}{\theta} \sum_{i=1}^k x_i - \frac{1}{1-\theta} \sum_{i=1}^k (n-x_i),$$

qui s'annule lorsque

$$\begin{aligned} \frac{1-\theta}{\theta} &= \frac{1}{\theta} - 1 = \frac{\sum_{i=1}^k (n-x_i)}{\sum_{i=1}^k x_i} \\ \Leftrightarrow \frac{1}{\theta} &= \frac{\sum_{i=1}^k (n-x_i)}{\sum_{i=1}^k x_i} + 1 = \frac{nk}{\sum_{i=1}^k x_i} \\ \Leftrightarrow \theta &= \frac{\sum_{i=1}^k x_i}{nk} = \frac{\bar{X}_k}{n}. \end{aligned}$$

Notons que

$$\ell''(\theta) = -\frac{1}{\theta^2} \sum_{i=1}^k x_i - \frac{1}{(1-\theta)^2} \sum_{i=1}^k (n-x_i) < 0,$$

donc $\hat{\theta}$ donne bien un maximum de ℓ .

2. **(3 points)** Le nombre de points marqués par un étudiant suit une loi binomiale de paramètres $n = 6$ et θ . D'après (a), l'estimation du maximum de vraisemblance de θ est

$$\hat{\theta} = \frac{\sum_{i=1}^{15} x_i}{6 \times 15} = 4/9 = 0.4445.$$

En supposant $k = 15$ assez grand pour appliquer la théorie asymptotique,

$$\hat{\theta} \sim \mathcal{N}(\theta, J(\hat{\theta})^{-1}),$$

où $J(\hat{\theta}) = -\ell''(\hat{\theta})$. On a $\sum_i x_i = nk\hat{\theta}$, et donc

$$\begin{aligned} -\ell''(\hat{\theta}) &= \frac{1}{\hat{\theta}^2} \sum_{i=1}^k x_i + \frac{1}{(1-\hat{\theta})^2} (nk - \sum_{i=1}^k x_i) \\ &= nk \left\{ \frac{\hat{\theta}}{\hat{\theta}^2} + \frac{1-\hat{\theta}}{(1-\hat{\theta})^2} \right\} \\ &= nk \frac{1}{\hat{\theta}(1-\hat{\theta})} \end{aligned}$$

ce qui donne

$$J(\hat{\theta}) = -\ell(\hat{\theta}) = 364.49.$$

Donc

$$\hat{\theta} \sim \mathcal{N}(\theta, 0.0027).$$

Un IC à 95% pour θ est

$$(\hat{\theta} - z_{1-\alpha/2}\sqrt{J(\hat{\theta})^{-1}}, \hat{\theta} + z_{1-\alpha/2}\sqrt{J(\hat{\theta})^{-1}})$$

où $\alpha = 0.05$ et $z_{1-\alpha/2} = 1.96$. On obtient donc $(0.3426, 0.5463) \approx (0.34, 0.55)$.

3. **(2 points)** Un étudiant réussit l'examen avec la probabilité

$$P(Z \geq 3) = P(Z = 3) + P(Z = 4) + P(Z = 5) + P(Z = 6)$$

où Z est une variable aléatoire binomiale de paramètres $n = 6$ et $\hat{\theta} = 4/9$. Donc

$$P(Z \geq 3) = 0.5472.$$

Parmi les 85 étudiants restants, le nombre de ceux ayant réussi le test suit une loi binomiale de paramètres $n = 85$ et $p = 0.5472$. Donc sur les 85 étudiants restants, en moyenne $0.5472 \times 85 = 46.51$ ont réussi le test.