Série 6

Solution 24. a) Soit A l'événement que “le premier essai est un succès”. En conditionnant sur A, on a deux cas :

i) soit un succès avec une probabilité $\Pr(A) = p$.

ii) soit un échec avec une probabilité $\Pr(A^c) = 1 - p$. Dans ce cas le nombre prévu d'échecs supplémentaires avant le premier succès est $\mathrm{E}(X)$. L'espérance conditionnelle de X, étant donné qu'il y a eu un échec au premier essai, est donc $\mathrm{E}(X \mid A^c) = 1 + \mathrm{E}(X)$.

En fait, pour toute fonction g raisonnable (c'est à dire mesurable) de X, ce raisonnement donne

$$\mathrm{E}\{g(1) \mid A\} = \mathrm{E}\{g(1)\} = g(1), \quad \mathrm{E}\{g(X) \mid A^c\} = \mathrm{E}\{g(1 + X)\}.$$

Ainsi

$$\begin{aligned}
\mathrm{E}(X) &= \mathrm{E}(X \mid A) \times \Pr(A) + \mathrm{E}(X \mid A^c) \times \Pr(A^c) \\
&= 1 \times p + \{1 + \mathrm{E}(X)\} \times (1 - p) \\
&= 1 + (1 - p)\mathrm{E}(X),
\end{aligned}$$

et en résolvant cette équation pour $\mathrm{E}(X)$, on obtient $\mathrm{E}(X) = 1/p$. De même façon,

$$\begin{aligned}
\mathrm{E}(X^2) &= \mathrm{E}(X^2 \mid A) \times \Pr(A) + \mathrm{E}(X^2 \mid A^c) \times \Pr(A^c) \\
&= 1 \times p + \mathrm{E}\{(1 + X)^2\} \times (1 - p) \\
&= p + (1 - p)\mathrm{E}(1 + 2X + X^2) \\
&= \frac{2 - p}{p} + (1 - p)\mathrm{E}(X^2),
\end{aligned}$$

nous donnant $\mathrm{E}(X^2) = (2 - p)/p^2$. La variance est ainsi

$$\mathrm{var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = (2 - p)/p^2 - 1/p^2 = (1 - p)/p^2.$$

b) On a

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{k=1}^{\infty} k\Pr(X = k) = \sum_{k=1}^{\infty} kp(1 - p)^{k-1} \\
&= p\sum_{k=1}^{\infty} k(1 - p)^{k-1} = p\left\{-\frac{d\left(\frac{1}{p}\right)}{dp}\right\} = \frac{1}{p}, \\
\mathrm{E}\{X(X - 1)\} &= \sum_{k=1}^{\infty} k(k - 1)p(1 - p)^{k-1} = p(1 - p)\sum_{k=1}^{\infty} k(k - 1)(1 - p)^{k-2} \\
&= p(1 - p)\left\{\frac{d^2\left(\frac{1}{p}\right)}{dp^2}\right\} = 2\frac{1 - p}{p^2},
\end{aligned}$$

donc $\mathrm{var}(X) = \mathrm{E}\{X(X - 1)\} + \mathrm{E}(X) - \mathrm{E}(X)^2 = (1 - p)/p^2$.

Solution 25. a) Soit I_A une variable indicatrice. Nous avons prouvé dans la série 5 (ex 4.1-a) que $\mathrm{E}(I_A) = \Pr(A)$. De la même manière, on montre que $E(I_A^2) = \Pr(A)$. Ainsi, $\mathrm{var}(I_A) = \mathrm{E}(I_A^2) - \mathrm{E}(I_A)^2 = \Pr(A)\big(1 - \Pr(A)\big)$.

b) Soit $X \sim B(n,p)$, alors

$$
\begin{aligned}
\mathrm{E}\{X(X-1)\} &= \sum_{x=0}^{n} x(x-1)\Pr(X=x)\\
&= \sum_{x=0}^{n} x(x-1)\binom{n}{x} p^x(1-p)^{n-x}\\
&= \sum_{x=0}^{n} x(x-1)\frac{n!}{x!(n-x)!} p^x(1-p)^{n-x}\\
&= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2}(1-p)^{n-x}\\
&= n(n-1)p^2(p+q)^{n-2}\\
&= n(n-1)p^2,
\end{aligned}
$$

et donc

$$
\mathrm{var}(X) = \mathrm{E}\{X(X-1)\} + \mathrm{E}(X) - \mathrm{E}(X)^2 = n(n-1)p^2 + np - (np)^2 = np(1-p).
$$

c) On a montré dans la série 5 (ex 4.1-d) que la loi binomiale $B(n,p)$ peut être écrite comme une somme de n variables de Bernoulli indépendantes, identiquement distribuées avec une probabilité p, $X = I_1 + \cdots + I_n$. La variance est alors donnée par :

$$
\begin{aligned}
\mathrm{var}(X) &= \mathrm{var}(I_1 + \cdots + I_n)\\
&= \sum_i \mathrm{var}(I_i) \text{ (par independence des variables)}\\
&= np(1-p) \text{ (en utilisant (a)}, \mathrm{var}(I_i) = p(1-p))
\end{aligned}
$$

Solution 26. Une variable aléatoire $Z \sim \mathcal{N}(0,1)$ a pour fonction de densité $\phi(z) = (2\pi)^{-1/2}\exp(-z^2/2)$ et pour fonction de répartition $\Phi(z) = \int_{-\infty}^{z} \phi(x)\,\mathrm{d}x$ $(z \in \mathbb{R})$.

a) $\phi'(z) = -z\phi(z)$ et $\phi''(z) = (z^2-1)\phi(z)$. Ainsi,

$$
\mathrm{E}(Z) = \int_{-\infty}^{\infty} z\phi(z) = -\int_{-\infty}^{\infty} \phi'(z)dz = [\phi(z)]_{-\infty}^{\infty} = 0,
$$

et

$$
\mathrm{var}(Z) = \mathrm{E}(Z^2) = \mathrm{E}(1+Z^2-1) = 1+\int_{-\infty}^{-\infty}(z^2-1)\phi(z)dz = \int_{-\infty}^{-\infty}\phi'(z)dz = 1+[\phi'(z)]_{-\infty}^{\infty} = 1.
$$

b) La fonction de répartition de X est :

$$
\Pr(X \le x) = \Pr(\mu + \sigma Z \le x) = \Pr\left(Z \le \tfrac{x-\mu}{\sigma}\right) = \Phi\left(\tfrac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R}.
$$

En dérivant la fonction de répartition, on obtient la fonction de densité de X :

$$
\phi_X(x) = \tfrac{1}{\sigma}\phi\left(\tfrac{x-\mu}{\sigma}\right) = (2\pi\sigma^2)^{-1/2}\exp\left\{-\tfrac{1}{2}\left(\tfrac{x-\mu}{\sigma}\right)^2\right\}, \quad x \in \mathbb{R}.
$$

L'espérance et la variance sont respectivement : $\mathrm{E}(X) = \mathrm{E}(\mu+\sigma Z) = \mu + \sigma E(Z) = \mu$ et $\mathrm{var}(X) = \mathrm{var}(\mu+\sigma Z) = \sigma^2 \mathrm{var}(Z) = \sigma^2$.

c) Pour $m \in \{1, 2, \ldots\}$, on a

$$\begin{aligned}
\mathrm{E}(Z^{2m+1}) &= \int_{-\infty}^{+\infty} z^{2m+1}\phi(z)dz \\
&= (2\pi)^{-1/2}\int_{-\infty}^{+\infty} z^{2m} z \exp(-\tfrac{z^2}{2})dz \\
&= (2\pi)^{-1/2}\left\{[-z^{2m}\phi(z)]_{-\infty}^{\infty} + 2m\int_{-\infty}^{+\infty} z^{2m-1}\exp(-\tfrac{z^2}{2})dz\right\} \\
&= 2m\mathrm{E}(Z^{2m-1}) \\
&= 2m\cdot(2m-2)\cdots\mathrm{E}(Z) \\
&= 0.
\end{aligned}$$

De la même manière, on obtient

$$\begin{aligned}
\mathrm{E}(Z^{2m}) &= \int_{-\infty}^{+\infty} z^{2m}\phi(z)dz \\
&= (2\pi)^{-1/2}\int_{-\infty}^{+\infty} z^{2m-1} z \exp(-\tfrac{z^2}{2})dz \\
&= (2\pi)^{-1/2}\{[-z^{2m-1}\phi(z)]_{-\infty}^{\infty} + (2m-1)\int_{-\infty}^{+\infty} z^{2m-2}\exp(-\tfrac{z^2}{2})dz\} \\
&= (2m-1)\mathrm{E}(Z^{2m-2}) \\
&= (2m-1)\cdot(2m-3)\cdots 3\cdot\mathrm{E}(Z^2) \\
&= (2m-1)\cdot(2m-3)\cdots 3\cdot 1.
\end{aligned}$$

d)
- $\phi(-z) = (2\pi)^{-1/2}\exp(-(-z)^2/2) = (2\pi)^{-1/2}\exp(-z^2/2) = \phi(z)$
- $\Phi(z) = \int_{-\infty}^{z}\phi(x)\,\mathrm{d}x = \int_{-\infty}^{\infty}\phi(x)\,\mathrm{d}x - \int_{z}^{\infty}\phi(x)\,\mathrm{d}x = \int_{-\infty}^{\infty}\phi(x)\,\mathrm{d}x - \int_{-\infty}^{-z}\phi(x)\,\mathrm{d}x = 1-\Phi(-z)$
- Par symétrie de la densité, les quantiles de Z satisfont : $z_p = -z_{1-p}$ $(0 < p < 1)$, sinon on a $p = \Phi(z_p) = 1 - \Phi(-z_p)$, et donc

  $$1 - p = \Phi(z_{1-p}) = \Phi(-z_p),$$

  nous donnant $z_{1-p} = -z_p$.
- $\Pr(X \le x_p) = \Pr(\frac{X-\mu}{\sigma} \le \frac{x_p-\mu}{\sigma}) = \Pr(Z \le z_p)$, Il s'en suit que $x_p = \mu + \sigma z_p$.

e) On a

$$\begin{aligned}
\mathrm{E}(X \mid X > a) &= \mathrm{E}(\mu + \sigma Z \mid \mu + \sigma Z > a) \\
&= \mu + \sigma\mathrm{E}\left(Z \mid Z > \frac{a-\mu}{\sigma}\right) \\
&= \mu + \sigma\frac{\int_b^\infty z\phi(z)}{\Pr(Z > b)} \\
&= \mu + \sigma\frac{-\int_b^\infty \phi'(z)}{1 - \Pr(Z \le b)} \\
&= \mu + \sigma\frac{\phi(b)}{1-\Phi(b)},
\end{aligned}$$

où $b = (a-\mu)/\sigma$.

f) On trouve $\Pr(Z \leq 0.53) = 0.70194$, $\Pr(Z \leq -1.86) = 1 - 0.96856 = 0.03144$, $\Pr(-1.86 < Z \leq 0.53) = 0.70194 - 0.03144 = 0.6705$, $z_{0.95} = 1.65$, $z_{0.025} = -1.96$, et $z_{0.5} = 0$.

Solution 27. Soit X la variable aléatoire qui représente les scores. X est normalement distribuée avec une moyenne de 50 et un écart-type de 20.

a) Pour $x = 67$, $z = \dfrac{(67-50)}{20} = 0,85$. La proportion P des élèves ayant obtenu une note inférieure à 67 est donc : $P = [\text{surface à gauche de } z = 0,85] = 0,8023 = 80,23\%$. Albert a donc obtenu un résultat meilleur que $80,23\%$ des étudiants qui ont passé le test et il sera admis à l'ETHZ.

b) Nous cherchons à calculer la valeur suivante $\mathrm{E}(X|X > a)$ où $a = \mu + \sigma z_{0.70} = 50 + 20 \cdot 0.53 = 60.6$

$$\begin{aligned}
\mathrm{E}(X|X > a) &= \mu + \sigma \frac{\phi(b)}{1 - \Phi(b)} \\
&= 50 + 20 \frac{\phi(0.53)}{1 - \Phi(0.53)} \\
&= 50 + 20 \frac{0.346}{1 - 0.7019} \\
&= 73.213.
\end{aligned}$$

La meilleure estimation de son score est 73.2.

Solution 28. Soit Y désignant le nombre d'allumettes restants dans l'autre boîte, notez que $Y \in \{0, \ldots, m\}$, et considérons en particulier les cas où $Y = m$ et $Y = 0$.

i) Pour $Y = m$, le capitaine Haddock a choisi la même boîte $m+1$ fois successivement, et ceci a une probabilité $1/2^{m+1}$. Mais puisqu'il y a deux façons de le faire (la boîte vide peut être dans n'importe quelle poche)

$$\Pr(Y = m) = \frac{1}{2^m}.$$

ii) L'événement $Y = 0$ correspond à toutes les séquences de longueur $2m$ avec exactement m D's et m G's, et ceci avec une probabilité

$$\Pr(Y = 0) = \binom{2m}{m} \frac{1}{2^{2m}},$$

parce que n'importe quelle poche qu'il choisira sera vide.

Considérons le cas général, en supposant qu'il mette sa main dans la poche droite et la trouve vide. Alors le nombre Y_G dans la poche gauche est égal à k s'il choisit la poche droite pour la $m+1$ fois sur le $(m-k)+(m+1)$ième choix, et ceci a une probabilité binomiale négative avec $x = 2m + 1 - k$ et $n = m + 1$, soit

$$\Pr(Y_G = k) = \binom{2m+1-k-1}{m+1-1} 2^{-(m+1)} 2^{-(m-k)} = \Pr(Y_D = k)$$

par symétrie. La probabilité requise est donc

$$\Pr(Y = k) = 2 \times \Pr(Y_G = k) = \binom{2m-k}{m} 2^{-(2m-k)}, \quad k = 0, \ldots, m.$$

Ceci correspond aux calculs pour $k = 0$ et $k = m$.