Devoir 1 - IFT3655

Mathias La Rochelle

Le lundi 16 septembre 2024

Question 1 : Dans une file d'attente à un seul serveur, supposons que les clients arrivent selon un processus de Poisson (i.e., au hasard et indépendamment les uns des autres) à un taux de $\lambda = 20$ clients par heure, que les durées de service sont indépendantes et suivent une loi exponentielle de moyenne $\frac{1}{\mu}$ minutes, et que les clients sont servis par ordre d'arrivée. Quel sera le temps d'attente moyen par client pour $\frac{1}{\mu} = 1,\ 2,\ 2.9,\ 2.99,\ \text{et } 3.1$?

Nous avions vu en classe qu'il est possible de déterminer le temps d'attente moyen par client lorsque les clients arrivent complètement au hasard et indépendamment les uns des autres grâce à la formule suivante :

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

Désormais, il est facile de trouver notre temps d'attente moyen par client w pour chaque durée moyenne de service donnée.

1. Nombre de clients servis en moyenne par heure : $\mu = 60,\ 30,\ 21,\ 21$ et 20
2. Temps d'attente moyen par client en heure :
   (a) $\mu = 60 \to w = \frac{20}{60(60-20)} \approx \frac{1}{120}$h ou 30 sec
   (b) $\mu = 30 \to w = \frac{20}{30(30-20)} \approx \frac{1}{15}$h ou 4 min
   (c) $\mu = 21 \to w = \frac{20}{21(21-20)} \approx \frac{20}{21}$h ou 57 min
   (d) $\mu = 21$ donc même chose qu'au précédent, $w \approx 57$ min
   (e) $\mu = 20$, irréaliste car $\frac{20}{\mu} \geq s$ et donc qu'il n'y aurait pas d'attente pour aucun client

Question 2 : Soient b et c deux constantes réelles positives et soit X une variable aléatoire qui suit la loi uniforme (continue) sur l'intervalle $[0,\ b)$.

(a) Quelle est la fonction de répartition de X ?

On sait que la fonction de répartition d'une variable aléatoire X peut s'écrire à partir de la fonction de densité :

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{x} \frac{1}{b}\,dt$$

Pour $x < 0$ on a,

$$\int_{-\infty}^{x} 0\,dt = 0, \text{ parce que la densité est nulle pour } t < 0$$

Pour $0 \leq x < b$ on a,

$$\int_0^x \frac{1}{b}\,dt = \left[\frac{t}{b}\right]_0^x = \frac{x}{b}$$

Pour $x \geq b$ on a,

$$\int_0^b \frac{1}{b}\,dt = \left[\frac{t}{b}\right]_0^b = 1$$

Donc la fonction de répartition de X est

$$F_X(x) = \begin{cases} 0, & \text{si } x < 0 \\ \frac{x}{b}, & \text{si } 0 \leq x < b \\ 1, & \text{si } x \geq b \end{cases}$$

(b) Si $Y = (X + c)\ mod\ b$, quelle est la fonction de répartition de Y ? De quelle loi sagit-t-il ? Vous devez bien sûr démontrer votre réponse. Note : le "mod b" ici veut dire que l'on soustrait ou ajoute b autant de fois qu'il faut pour obtenir $0 \leq Y \leq b$.

La fonction de répartition de Y s'écrit comme :

$$\begin{aligned} F_Y(y) = \mathbb{P}(Y \leq y) &= \mathbb{P}\big((X+c)\ mod\ b \leq y\big) \\ &= \sum_t \mathbb{P}\big((X+c) - bt \leq y\big) \\ &= \sum_t \mathbb{P}(bt - c \leq X \leq y + bt - c) \end{aligned}$$

On peut désormais déterminer la fonction de répartition de Y à partir de X. Aussi, la variable Y est discrète à cause de l'opérateur modulo :

$$\begin{aligned} \sum_t \big(F_X(y + bt - c) - F_X(bt - c)\big) &= \sum_t \int_{bt-c}^{y+bt-c} f_X(x)\,dx \\ &= \sum_t \left[\frac{x}{b}\right]_{bt-c}^{y+bt-c} \\ &= \sum_t \left(\frac{y + bt - c}{b} - \frac{bt - c}{b}\right) \\ &= \frac{yt}{b} \end{aligned}$$

Les conditions que $0 \leq y \leq c$ et $c \leq y < b$ sont assez trivials. Donc la fonction de répartition de Y est

$$F_Y(y) = \begin{cases} 0, & \text{si } y < 0 \\ \frac{yt}{b}, & \text{si } 0 \leq y < b,\ \forall t \in \mathbb{N} \\ 1, & \text{si } y \geq b \end{cases}$$

Question 3 : On veut générer une variable aléatoire qui prend la valeur 0 ou 1, chacune avec probabilité $\frac{1}{2}$, en tirant à pile ou face. Pour cela, on dispose uniquement d'une pièce biaisée, qui tombe sur F avec une probabilité p inconnue, et sur P avec probabilité $(1-p)$. On considère les deux procédures suivantes.

Dans chacun des deux cas, est-ce que le résultat est équitable ou pas ? Dans chaque cas, vous devez le démontrer.

(a) On lance la pièce deux fois. Si on obtient PF, on retourne 0; si on obtient FP, on retourne 1; si on obtient FF ou PP, on recommence.

La question est $\mathbb{P}(0) \stackrel{?}{=} \mathbb{P}(1)$. Sachant que $\mathbb{P}(F) = p$ et $\mathbb{P}(P) = 1-p$, on peut calculer les probabilités des événements dans $\Omega = \{FF,\ FP,\ PF,\ PP\}$:

$$\begin{aligned}
\mathbb{P}(FF) &= p^2 \\
\mathbb{P}(FP) &= \mathbb{P}(PF) = p - p^2 \\
\mathbb{P}(PP) &= 1 - 2p + p^2
\end{aligned}$$

Étant donnée que $\sum_i^{|\Omega|} \mathbb{P}(E_i) = 1$, alors

$$\mathbb{P}(\text{Ne pas recommencer}) = 1 - \big(\mathbb{P}(FF) + \mathbb{P}(PP)\big) = 1 - (p^2 + 1 - 2p + p^2) = 2p(1-p)$$

Puis,

$$\begin{aligned}
\mathbb{P}(0) &= \frac{\mathbb{P}(PF)}{\mathbb{P}(\text{Ne pas recommencer})} = \frac{p-p^2}{2p(1-p)} = \frac{1}{2} \\
\mathbb{P}(1) &= \frac{\mathbb{P}(FP)}{\mathbb{P}(\text{Ne pas recommencer})} = \frac{p-p^2}{2p(1-p)} = \frac{1}{2}
\end{aligned}$$

(b) On lance jusqu'à ce que le dernier tirage soit différent du précédent. Si on a P$\cdots$PF on retourne 0, si on a F$\cdots$FP on retourne 1.

Dans la première (P$\cdots$PF) et deuxième (F$\cdots$FP) séquence, on termine avec l'événement PF et FP respectivement. On a vu que $\mathbb{P}(PF) = p - p^2 = \mathbb{P}(FP)$ mais les événements sont équitables si et seulement si $\mathbb{P}(P) = \mathbb{P}(F)$. Si ces probabilités sont différentes alors la séquence d'événements qui vient avant FP ou PF feront en sorte que $\mathbb{P}(P...PF) \neq \mathbb{P}(F...FP)$.

Question 4 : Une petite ville est divisée en quatre quartiers: A, B, C, D, qui ont des populations respectives de 1500, 1800, 3200, et 500. Le nombre de personnes ayant déjà fréquentée l'université dans chacun de ces quartiers est de 150, 120, 60, et 100. Si vous rencontrez une personne au hasard venant de cette ville et qu'elle a fréquenté l'université, quelle est la probabilité qu'elle vienne du quartier X pour X = A, B, C, D (on veut la probabilité pour chacun des quartiers) ? Quelle est la somme de ces probabilités ?

$$\mathbb{P}(X \mid \text{Université}) = \frac{\mathbb{P}(\text{Université} \mid X) \cdot \mathbb{P}(X)}{\mathbb{P}(\text{Université})}$$

1. $X = A \to \mathbb{P}(A \,|\, \text{Université}) = \frac{\frac{150}{1500} \cdot \frac{1500}{7000}}{\frac{430}{7000}} = \frac{150}{430} \approx 34.88\%$

2. $X = B \to \mathbb{P}(B \,|\, \text{Université}) = \frac{\frac{120}{1800} \cdot \frac{1800}{7000}}{\frac{430}{7000}} = \frac{120}{430} \approx 27.9\%$

3. $X = C \to \mathbb{P}(C \,|\, \text{Université}) = \frac{\frac{60}{3200} \cdot \frac{3200}{7000}}{\frac{430}{7000}} = \frac{60}{430} \approx 13.95\%$

4. $X = D \to \mathbb{P}(D \,|\, \text{Université}) = \frac{\frac{100}{500} \cdot \frac{500}{7000}}{\frac{430}{7000}} = \frac{100}{430} \approx 23.26\%$

La somme de ces probabilités est de 1.

Question 5 : Un vol d'Air Canada a 220 sièges, mais la demande dépasse le nombre de sièges disponibles. Considérant que certains passagers qui achètent un billet ne vont pas se présenter pour leur vol, la compagnie décide de faire de la sur-réservation ("overbooking") et de vendre 226 billets. Soit $X \geq 0$ le nombre de passagers qui vont se présenter avec un billet et qui n'auront pas de siège. En supposant que chaque détenteur de billet a 3% des chances de ne pas se présenter et que ces décisions sont indépendantes, ce X est une variable aléatoire qui suit quelle loi de probabilité ? Attention: Notez que X ne peut pas prendre une valeur négative. Il faut d'abord bien identifier l'ensemble des valeurs que X peut prendre. Vous pouvez ensuite trouvez une expression pour $\mathbb{P}[X = x]$ pour les valeurs positives de x, puis une expression pour $\mathbb{P}[X = 0]$. On vous demande ensuite de calculer les valeurs numériques de ces probabilités. Si vous ne pouvez pas calculer les valeurs exactes, vous pouvez approximer $\mathbb{P}[X = x]$ pour les valeurs positives de x en utilisant une approximation appropriée de la loi exacte. Supposons maintenant que le gestionnaire décide plutôt de vendre le plus grand nombre de billets possible, mais sous la contrainte que $\mathbb{P}[X > 0] \leq 0.1$. Quel est le nombre maximal de billets qui satisfait cette contrainte?

Il n'y a pas de loi de probabilité exacte à laquelle X s'associe. C'est plutôt par l'intermédiaire d'une autre variable aléatoire Y décrivant le nombre de détenteurs d'un billet que X peut être expliqué. Il y a 226 billets vendus et la probabilité qu'un détenteur se présente est de $1 - 0.03 = 97\%$.

$$\text{Y } : Y \sim \text{Bin}(n = 226,\, p = 0.97)$$
$$\text{X } : X = Y - N, \text{ où N} = 220$$

Probabilité que les détenteurs de billets aient tous une place ($X = 0$) :

$$\mathbb{P}(Y \leq 220) = \sum_{k=0}^{N} \binom{226}{k} \cdot 0.97^k \cdot 0.03^{226-k} = 67.4\%$$

Probabilité que x détenteurs de billets n'aient pas de place $X = 1, 2, 3, 4, 5, 6$:

$$\mathbb{P}(Y = 221) = \binom{226}{221} \cdot 0.97^{221} \cdot 0.03^{5} = 13.62\%$$
$$\mathbb{P}(Y = 222) = \binom{226}{222} \cdot 0.97^{222} \cdot 0.03^{4} = 9.92\%$$
$$\mathbb{P}(Y = 223) = \binom{226}{223} \cdot 0.97^{223} \cdot 0.03^{3} = 5.75\%$$
$$\mathbb{P}(Y = 224) = \binom{226}{224} \cdot 0.97^{224} \cdot 0.03^{2} = 2.49\%$$
$$\mathbb{P}(Y = 225) = \binom{226}{225} \cdot 0.97^{225} \cdot 0.03^{1} = 0.72\%$$
$$\mathbb{P}(Y = 226) = \binom{226}{226} \cdot 0.97^{226} \cdot 0.03^{0} = 0.1\%$$

Ces valeurs ont été calculées numériquement grâce au langage de programmation R et l'IDE RStudio.

```r
n <- 226
p <- 0.97
k <- 220

for (i in 1:6) {
  probX <- dbinom(k + i, size = n, prob = p)
  message <- paste("Probabilite que", i, "clients arrivent et n'ont pas
      de siege :", round(probX * 100, digits=2), "%")
  print(message)
}

probX0 <- pbinom(k, size=n, prob=p)
message <- paste("Probabilite qu'aucun client arrive et qu'il n'y a plus
de siege : ", round(probX0 * 100, digits=2), "%")
print(message)
```

output.txt

```
[1] "Probabilité que 1 clients arrivent et n'ont pas de siège : 13.62 %"
[1] "Probabilité que 2 clients arrivent et n'ont pas de siège : 9.92 %"
[1] "Probabilité que 3 clients arrivent et n'ont pas de siège : 5.75 %"
[1] "Probabilité que 4 clients arrivent et n'ont pas de siège : 2.49 %"
[1] "Probabilité que 5 clients arrivent et n'ont pas de siège : 0.72 %"
[1] "Probabilité que 6 clients arrivent et n'ont pas de siège : 0.1 %"
[1] "Probabilité qu'aucun client arrive et qu'il n'y a plus de siège :  67.4 %"
```

Pour trouver le nombre maximum de billets pouvant être vendus selon la contrainte $\mathbb{P}(X > 0) \leq 0.1$, on peut procéder de la manière suivante :

$$\begin{aligned}\mathbb{P}(Y > 220) \leq 0.1 &\equiv \mathbb{P}(Y \leq 220) \geq 0.9 \\ &\equiv \sum_{k=0}^{220} \binom{x}{k} \cdot 0.97^{k} \cdot 0.03^{x-k} \geq 0.9\end{aligned}$$

Pour résoudre cette équation, nous allons encore utiliser RStudio. Le seul moyen que j'ai trouvé pour chercher le nombre maximum de billets c'est d'itérer à travers un certain intervalle et d'imprimer la probabilité si et seulement si elle respecte la condition ≥ 0.9.

```r
n <- 226
p <- 0.97
k <- 220

compute_cdf <- function(n, p, k) {
  pbinom(k, size=n, prob = p)
}

for (n in seq(240, 200, by = -1)) { # l'intervalle 200-240 a été tester,
                                    # ce n'est pas au hasard
  prob <- compute_cdf(n, p, k)
  if (prob >= 0.9) {
    message <- paste("Le nombre maximal de billets à vendre est ", n, ".")
    print(message)
    break
  }
}
```

output.txt

```
[1] "Le nombre maximal de billets à vendre est 224."
```