Durée de préparation : 1 heure 30

Question

Le tableau ci-dessous résume les surfaces, en millier de m^2, utilisées pour l'installation de panneaux solaires destinée à la production électrique en France, entre 2008 et 2017.

Année	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Surface	527	583	744	917	1139	1302	1447	1595	1762	1916

1. Calculer le taux d'évolution de la surface dédiée à l'énergie solaire entre 2008 et 2017.

   Corrigé: Le taux d'évolution correspond à la valeur T tel que
   $$S' = S \times (1+T),$$
   qui peut s'exprimer aussi par
   $$T = \frac{S'-S}{S}.$$
   Dans notre cas,
   $$T = \frac{1916-527}{527} = 2,636$$
   qui correspond à un taux d'évolution en pourcentage de 263,6%.

2. Déterminer le taux d'évolution annuel moyen entre 2008 et 2017.

   Corrigé: On cherche un taux d'évolution annuel t tel que
   $$S' = S \times (1+t)^n,$$
   où n est la quantité d'années. Le taux t s'obtient avec
   $$t = \left(\frac{S'}{S}\right)^{\frac{1}{n}} - 1.$$
   Dans notre cas,
   $$t = \left(\frac{1916}{527}\right)^{\frac{1}{9}} - 1 = 0,154$$
   qui correspond à un taux d'évolution annuel moyen en pourcentage de 15,4%.

3. En supposant que le taux d'évolution annuel moyen se stabilise à 15% après 2017, estimer à partir de quelle année la surface dédiée à l'énergie solaire dépassera 7 millions de m^2.

   Corrigé: On peut estimer la surface pour l'année $2017+n$ par
   $$S_n = 1916 \times (1+0,15)^n$$
   ce qui donne $S_9 \approx 6740$ et $S_{10} \approx 7751$. L'estimation indique que la surface dédiée à l'énergie solaire dépassera 7 millions de m^2 à partir de l'année 2027.

Exercice 1

La direction d'un hypermarché a effectué une étude sur les modes de paiement en caisse de ses clients. Pour cela, elle a effectué une enquête pendant une heure de pointe sur 10 de ses caisses dont les résultats sont donnés dans le tableau ci-dessous :

Caisse n°	1	2	3	4	5	6	7	8	9	10
Nombre total de clients	36	45	34	41	38	46	30	47	37	41
Nombre de paiements par CB	9	13	13	12	14	16	11	12	16	15

Partie 1 : estimation des paramètres du modèle

On suppose que le nombre de personnes s'étant présentées à chacune des caisses est, pour chaque caisse, une variable aléatoire qui suit une loi de Poisson de même paramètre λ, et que ces variables aléatoires sont indépendantes.

1. A partir des données du tableau, expliquer pourquoi une estimation du paramètre λ peut être donnée par : $\lambda \approx 39,5$.

   Corrigé: L'espérance d'une loi de Poisson est égale à son paramètre. Par conséquent, une bonne estimation de λ est donnée par la moyenne empirique du nombre de clients à chaque caisse, qui donne $39,5$.

2. A l'aide de cette estimation, déterminer la probabilité qu'au moins 5 clients se présentent à une caisse donnée.

   Corrigé: $\mathbf{P}(N \geqslant 5) = 1-\mathbf{P}(N<5) = 1-\mathbf{P}(N=0)-\mathbf{P}(N=1)-\mathbf{P}(N=2)-\mathbf{P}(N=3)-\mathbf{P}(N=4)$. On sait que $\mathbb{P}(N=k) = e^{-\lambda}\lambda^k/k!$, il suffit donc de faire le calcul en remplaçant λ par $39,5$.

On suppose qu'à chaque caisse, la probabilité qu'un client paye par carte bancaire (CB) est p, et que les modes de paiements des clients successifs sont indépendants.

3. A partir des résultats de l'enquête, montrer qu'une estimation de p peut être donnée par $p \approx 0,33$.

   Corrigé: Là encore, on estime p par la moyenne empirique, c-est-à-dire la proportion $\hat{p} = N_{CB}/N$ de clients qui a payé par carte:

$$\hat{p} = \frac{9+13+13+12+14+16+11+12+16+15}{36+45+34+41+38+46+30+47+37+41} \approx 0.3316.$$

4. Donner un intervalle de confiance pour p au niveau de confiance 95%.

   Corrigé: Il s'agit (presque) du problème du sondage, cf la section 4.2 page 18 dans les notes du cours de statistiques.

   Le nombre total de clients est $N = 395$, ce qui est très élevé : on peut donner un intervalle de confiance asymptotique en utilisant le TCL. Il devrait être possible de donner directement la forme de l'intervalle de confiance, mais il est très instructif de le construire soi-même, ce que l'on fait ici. On rappelle que si la probabilité qu'un client paie par carte bancaire est p, alors

$$\frac{\sqrt{N}}{\sqrt{p(1-p)}}\left(\frac{N_{CB}}{N} - p\right) \xrightarrow{\text{loi}} \mathscr{N}(0,1).$$

   En particulier, si t est tel que $\mathbf{P}(|\mathscr{N}[0,1)| > t) = 0.05$ (on regarde une table et on trouve $t = 1.96$), alors

$$\lim_{n\to\infty} \mathbf{P}\left(-t < \frac{\sqrt{N}}{\sqrt{p(1-p)}}\left(\frac{N_{CB}}{N} - p\right) < t\right) = 0.95.$$

On pivote et on obtient que $\mathbf{P}(p \in J_N) \to 0.95$, où

$$J_N = \left]\hat{p} - \frac{t\sqrt{p(1-p)}}{\sqrt{N}}, \hat{p} + \frac{t\sqrt{p(1-p)}}{\sqrt{N}}\right[.$$

Comme d'habitude, le problème vient du fait que p n'est pas connu, donc on ne peut pas évaluer cet intervalle; heureusement, on utilise l'astuce classique $p(1-p) \leqslant 1/4$ (car $p \in [0,1]$). L'intervalle $I_N =]\hat{p} \pm t/s\sqrt{N}[$ contient J_n, donc $\mathbf{P}(\hat{p} \in J_n) \geqslant 0.95$. Avec les données de l'énoncé, on obtient que $p \in]0.33 \pm 1.96/2\sqrt{295}[$ avec probabilité supérieure à 0.95, c'est-à-dire

$$0.28 \leqslant p \leqslant 0.38.$$

Caveat. Il y a une petite difficulté ici, c'est que le nombre total de personnes qui se sont présentées à la caisse, N, est une variable aléatoire, donc le TCL tel que vu en cours ne s'applique pas *stricto sensu*. Cependant, l'énoncé suppose que le comportement des clients ne dépend pas de N en disant que la probabilité qu'un client paie par carte bleue est fixe et égale à p. Si c'est vraiment le cas, alors on peut raisonner conditionnellement à N et considérer ce dernier comme fixé, comme dans la correction. Mais dans la réalité, il y a de fortes chances pour que ce ne soit pas le cas : le mode de paiement peut très bien être influencé par la taille de la file d'attente, qui dépend de N.

5. Sachant que 30 clients s'y sont présentés, quelle est la probabilité qu'au plus 4 clients aient payé par carte bancaire à une caisse donnée ?

   Corrigé: Si $n = 30$ clients se sont présentés à la caisse et que le mode de paiement de chaque client est une Bernoulli de paramètre p indépendante des autres, alors le nombre X de clients ayant payé par carte bancaire est une binomiale $\mathrm{Bin}(n,p) = \mathrm{Bin}(30, 0.33)$, et donc

$$\mathbf{P}(X \leqslant 4) = \sum_{i=0}^{4} \binom{30}{i} p^i (1-p)^{30-i}.$$

Partie 2 : étude d'indépendance

Nous étudions dans cette partie l'indépendance entre le nombre de clients d'une caisse donnée payant par carte bancaire, et le nombre de clients utilisant un autre mode de paiement à la même caisse pendant une heure de pointe.

Pour cela, on considère les variables aléatoires suivantes :

$Z =$ « nombre total de clients payant à cette caisse pendant une heure de pointe »,

$X =$ « nombre de clients payant à cette caisse par carte bancaire pendant cette même heure de pointe »,

$Y =$ « nombre de clients utilisant à cette caisse un autre mode de paiement pendant la même heure de pointe ».

On a donc : $Z = X + Y$.

On suppose que :

- Z suit une loi de Poisson de paramètre λ.
- la probabilité qu'un client paye par carte bancaire à une heure de pointe est constante et égale à p.

1. Pour tout entier naturel n et tout entier k positif inférieur ou égal à n, déterminer la probabilité conditionnelle $P_{\{Z=n\}}(X = k)$, que k clients payent par carte bancaire à une caisse donnée sachant que n clients sont passés à cette caisse.

   Corrigé: Conditionnellement au fait que n clients se soient présentés à la caisse, X suit une loi binomiale de paramètres n, p, donc $\mathbf{P}(X = k|Z = n) = 0$ si $k > n$, et si $k \in \{0, \dots, n\}$ c'est $\binom{n}{k}p^k(1-p)^{n-k}$.

2. (a) Déterminer a tel que $\sum_{n=0}^{+\infty} \frac{(1-p)^n}{n!}\lambda^n = e^a$.

   Corrigé: On reconnaît la série qui définit l'exponentielle:

$$\sum_{n=0}^{\infty} \frac{(1-p)^n}{n!}\lambda^n = \sum_{n=0}^{\infty} \frac{[(1-p)\lambda]^n}{n!} = e^{(1-p)\lambda}.$$

   (b) Pour k entier naturel, exprimer $P(X = k)$ en fonction des probabilités $P_{\{Z=n\}}(X = k)$ et $P(Z = n)$ où $n \in \mathbb{N}$.

   Corrigé: La formule des probabilités totales et la réponse de la question 1 de la partie 2 donnent :

$$\begin{aligned}
\mathbf{P}(X = k) &= \sum_{n=0}^{\infty} \mathbf{P}(X = k|Z = n)\mathbf{P}(Z = n) \\
&= \sum_{n=k}^{\infty} \binom{n}{k} p^k(1-p)^{n-k} \times e^{-\lambda}\frac{\lambda^n}{n!} \\
&= \sum_{n=k}^{\infty} p^k \frac{n!}{k!(n-k)!}(1-p)^{n-k} \times e^{-\lambda}\frac{\lambda^{n-k}\lambda^k}{n!} \\
&= e^{-\lambda}\lambda^k p^k \frac{1}{k!}\sum_{n=k}^{\infty} \frac{(1-p)^{n-k}\lambda^{n-k}}{(n-k)!}.
\end{aligned}$$

   En effectuant le changement de variables $i = n - k$, on obtient

$$\begin{aligned}
\mathbf{P}(X = k) &= e^{-\lambda}\lambda^k p^k \frac{1}{k!}\sum_{i=0}^{\infty} \frac{(1-p)^i\lambda^i}{i!} \\
&= e^{-\lambda}\frac{(\lambda p)^k}{k!}e^{(1-p)\lambda} \\
&= e^{-p\lambda}\frac{(p\lambda)^k}{k!}.
\end{aligned}$$

   (c) En déduire que la variable aléatoire X suit une loi de Poisson dont on précisera le paramètre.

   Corrigé: X suit donc une loi de Poisson de paramètre λp, et le même raisonnement montre que Y suit une loi de Poisson de paramètre $\lambda(1-p)$.

3. On admet que Y suit une loi de Poisson de paramètre $\lambda(1-p)$.

   Montrer que le nombre de personnes payant par carte bancaire est indépendant du nombre de personnes utilisant un autre mode de paiement.

Corrigé: Pour montrer que X, Y sont indépendantes, il faut montrer que $\mathbf{P}(X = i, Y = j) = \mathbf{P}(X = i)\mathbf{P}(Y = j)$. Par la formule des probabilités totales,

$$\begin{aligned}\mathbf{P}(X = i, Y = j) &= \sum_{n=0}^{\infty} \mathbf{P}(X = i, Y = j | Z = n) e^{-\lambda} \frac{\lambda^n}{n!} \\ &= \sum_{n=0}^{\infty} \mathbf{1}_{n \geqslant i, n \geqslant j} \binom{n}{i} p^i (1-p)^j \mathbf{1}_{i+j=n} e^{-\lambda} \frac{\lambda^n}{n!}\end{aligned}$$

en effet, si $Z = n$ et $X = i$ on a nécessairement $Y = n - j$. Finalement, la somme ci-dessus n'a qu'un seul terme, qui correspond à $n = i + j$, et donc

$$\begin{aligned}\mathbf{P}(X = i, Y = j) &= \binom{n}{i} p^i (1-p)^j e^{-\lambda} \frac{\lambda^n}{n!} \\ &= \frac{p^i}{i!} \frac{(1-p)^{n-i}}{(n-i)!} e^{-\lambda} \lambda^n \\ &= \frac{p^i}{i!} \frac{(1-p)^j}{j!} e^{-\lambda p - \lambda(1-p)} \lambda^i \lambda^j \\ &= e^{-\lambda p} \frac{(\lambda p)^i}{i!} e^{-\lambda(1-p)} \frac{(\lambda(1-p))^j}{j!} \\ &= \mathbf{P}(X = i)\mathbf{P}(Y = j)\end{aligned}$$

d'où l'indépendance de X et Y.

Exercice 2

En 1928, C. Cobb et P. Douglas ont modélisé la croissance américaine entre 1899 et 1922 par la relation :

$$Q(L, K) = 1.01\, L^{0.25} K^{0.75}$$

où Q, K, L désignent respectivement la production totale, le capital investi et la quantité de travail.

1. Calculer $Q(147, 208)$ et $Q(147, 210)$ et évaluer en pourcentage, les variations relatives $\frac{\Delta Q}{Q}$ et $\frac{\Delta K}{K}$ quand K passe de la valeur 208 à 210, L restant égal à 147.

   Corrigé:

   $$Q(147, 208) \approx 192,6 \qquad Q(147, 210) \approx 194$$

   $$\frac{\Delta Q}{Q} = \frac{194 - 192,6}{192,6} \approx 0,0072,$$

   qui correspond à une variation relative de 0,72%.

   $$\frac{\Delta K}{K} = \frac{210 - 208}{208} \approx 0,0096,$$

   qui correspond à une variation relative de 0,96%.

2. Calculer $Q(aL, aK)$, a réel positif, en fonction de $Q(L, K)$ et de a. Interpréter votre résultat.

Corrigé:

$$\begin{aligned} Q(aL, aK) &= 1,01\,(aL)^{0,25}\,(aK)^{0,75} \\ &= 1,01\,a^{0,25}\,L^{0,25}\,a^{0,75}\,K^{0,75} \\ &= 1,01\,a^{0,25+0,75}\,L^{0,25}\,K^{0,75} \\ &= a\,1,01\,L^{0,25}\,K^{0,75} \\ &= a\,Q(L, K). \end{aligned}$$

Q est une fonction homogène de degré 1.

3. Elasticité

   (a) Calculer l'élasticité partielle de Q par rapport à K au point $(147, 208)$. Interpréter le résultat.

   Corrigé: L'élasticité point de Q par rapport à K se définie comme

   $$e_{Q/K} = \frac{\partial Q}{\partial K}\frac{K}{Q}.$$

   $$\frac{\partial Q}{\partial K} = 1,01\,L^{0,25}\,0,75\,K^{0,75-1} = 0,75\cdot 1,01\frac{L^{0,25}}{K^{0,25}}.$$

   D'où,

   $$e_{Q/K} = 0,75\cdot 1,01\frac{L^{0,25}}{K^{0,25}}\frac{K}{1,01\,L^{0,25}K^{0,75}} = 0,75.$$

   L'élasticité point de Q par rapport à K vaut toujours 0,75 et ne dépend pas du point (L, K).

   (b) Vérifier la cohérence avec le résultat de la question 1).

   Corrigé: L'élasticité arc de Q par rapport à K se définie comme

   $$\frac{\frac{\Delta Q}{Q}}{\frac{\Delta K}{K}} = \frac{0,0072}{0,0096} \approx 0,75$$

   Dans ce cas on obtient le même résultat.

4. On désigne respectivement par p_L et p_K les coûts unitaires du travail et du capital.

   (a) Expliciter l'expression de la fonction de coût C en fonction de L et de K.

   Corrigé:

   $$C(L, K) = p_L\,L + p_K\,K,$$

   définie pour $L > 0$ et $K > 0$.

   (b) On suppose que $p_L = 2$ et $p_K = 5$ et que la production est fixée à $Q = 300$. Le but des questions qui suivent est de minimiser le coût.

      i. Le niveau de production étant de 300, donner l'expression de K en fonction de L et construire dans un repère la courbe de niveau $Q = 300$, K étant portée en ordonnée et L en abscisse.

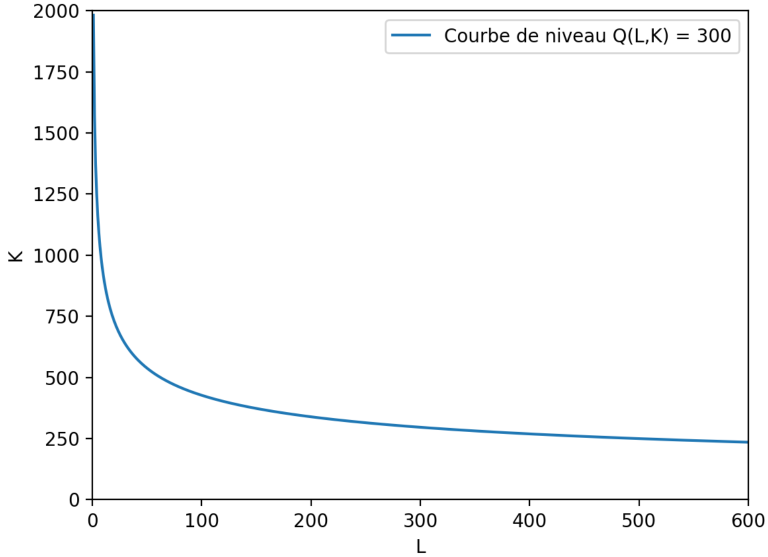
      Corrigé:

      $$Q(L, K) = 1,01\,L^{0,25}\,K^{0,75} = 300.$$

      Alors,

      $$K = \left(\frac{300}{1,01\,L^{0,25}}\right)^{\frac{1}{0,75}} = (297\,L^{-0,25})^{\frac{1}{0,75}} = 1981,8\,L^{-\frac{1}{3}}$$

      Finalement, $K(L) = 1981,8\,L^{-\frac{1}{3}}$, définie pour $L > 0$.



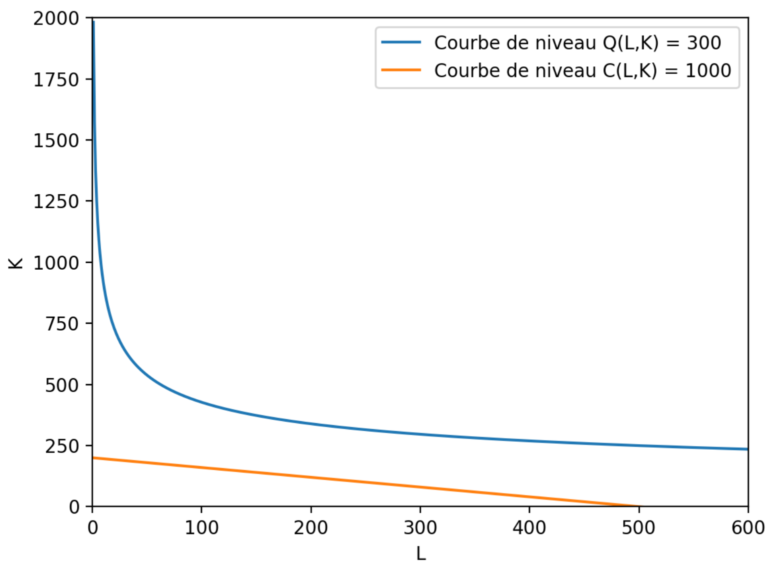
ii. Tracer dans le même repère la courbe de niveau $C(L, K) = 1000$.

Corrigé:

$$C(L, K) = 2L + 5K = 1000$$

d'où

$$K = 200 - \frac{2}{5}L$$



La courbe de niveau $C(L, K) = 1000$ ne coupe pas la courbe de niveau $Q(L, K) = 300$. On peut conclure qu'avec un coût total de 1000 ce n'est pas possible d'obtenir une production de 300.

iii. Montrer que minimiser le coût se ramène à minimiser une fonction d'une variable que l'on explicitera.

Corrigé: Le coût à minimiser est

$$C(L,K) = 2L + 5K.$$

En remplaçant K par l'expression de $K(L)$ pour $Q = 300$ obtenue dans la partie 4(b),

$$C(L) = 2L + 5(1981,8\, L^{-\frac{1}{3}}) = 2L + 9909\, L^{-\frac{1}{3}},$$

définie pour $L > 0$. Cette dernière est bien une fonction d'une variable.

iv. Déterminer les valeurs de L et de K qui minimisent le coût et calculer le coût minimum sous la contrainte $Q = 300$.
Donner l'interprétation graphique.

Corrigé: La dérivée de $C(L)$ est

$$C'(L) = 2 - \frac{9909}{3}\, L^{-\frac{4}{3}}.$$

Le seul point critique qui donne $C'(L) = 0$ est

$$L = \left(\frac{3303}{2}\right)^{\frac{3}{4}} \approx 259.$$

La dérivée seconde de $C(L)$ est

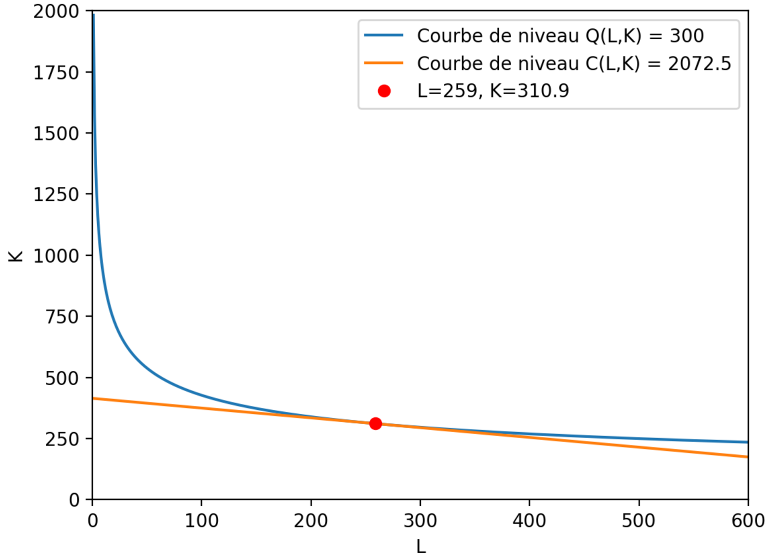
$$C''(L) = -3303 \cdot -\frac{4}{3}\, L^{-\frac{4}{3}-1} = 4404\, L^{-\frac{7}{3}}.$$

$C''(259) \approx 0,01 > 0$ et $L = 259$ est un minimum local de $C(L)$. En plus, $C''(L) > 0$ pour tout $L > 0$, c'est-à-dire pour tout le domaine de définition de $C(L)$. La fonction $C(L)$ est donc convexe, et on peut conclure que $L = 259$ est un minimum global.
La valeur de K est donnée par

$$K(259) = 1981,8 \times 259^{-\frac{1}{3}} \approx 310,9.$$

Le coût minimum sous la contrainte $Q = 300$ est

$$C(259) = 2 \times 259 + 9909 \times 259^{-\frac{1}{3}} \approx 2072,5.$$


Courbe de niveau Q(L,K) = 300
Courbe de niveau C(L,K) = 2072.5
L=259, K=310.9
K
L
0
250
500
750
1000
1250
1500
1750
2000
0
100
200
300
400
500
600