

CC1 2025/2026 – Durée 1h30

Les documents et appareils électroniques (calculatrice, téléphone, ordinateur, ...) sont interdits. Toutes les réponses doivent être justifiées.

Exercice 1 - Plat végétarien au RU

Par conviction écologique, vous décidez d'adopter un régime végétarien, mais par nécessité vous continuez à fréquenter le RU, qui propose un repas végétarien, dont le stock, noté $K \in \mathbb{N}^*$ est supposé fixe, et malheureusement limité.

I- Calibration du nombre de tupperwares hebdomadaire

Vous fréquentez le RU n fois, et notez Z_i la variable qui vaut 1 s'il ne reste plus de plat végétarien lorsque vous arrivez au stand. On note N l'espérance du nombre de fois où cela arrive sur une semaine (5 jours). À partir de $Z_1, \dots, Z_n$, on veut estimer N (dans cette partie).

1. Proposer un modèle pour cette expérience.
2. Exprimer N en fonction de l'espérance de Z_1 . En déduire un estimateur pour N , noté $\hat{N}$.
3. Montrer que $\hat{N}$ est consistant.
4. Déterminer le comportement asymptotique de $\hat{N}$.
5. En déduire un intervalle de niveau de confiance asymptotique 95% sur N , de type $] - \infty, \hat{N}_+]$.
6. **Application** : Après $n = 25$ tentatives au RU, vous n'avez pas pu manger végétarien 5 fois. Quel est le nombre de tupperwares (végétariens) que vous devez préparer pour une semaine afin d'être asymptotiquement sûr à 95% de manger à votre faim ? Ci-dessous quelques quantiles d'une loi normale standard.

α	0,9	0,95	0,975
q_α	1,28	1,65	1,96

II- Estimation de la demande en fonction du temps d'arrivée

On suppose que vous arrivez au RU au temps T_0 (fixe et connu) après son ouverture. Pour un passage i (toujours entre 1 et n), on note $Y_{i,1}$ le nombre d'usagers vous précédant dans la file qui demandent le plat végétarien, et on suppose $Y_{i,1} \sim \mathcal{P}(\lambda T_1)$, où λ **est inconnu** et T_1 est le temps d'attente au stand (supposé fixe et connu). On notera aussi $Y_{0,i}$ le nombre de plats végétariens demandés avant votre arrivée (non observé donc), supposé suivre un loi $\mathcal{P}(\lambda T_0)$, et on supposera de plus $Y_{i,0} \perp\!\!\!\perp Y_{i,1}$. Vous observez $(Y_{i,1})_{i=1,\dots,n}$, et pas $(Y_{i,0})_{i=1,\dots,n}$.

7. Donner un modèle pour cette expérience, et proposer un estimateur $\hat{\lambda}$ pour le paramètre λ .
8. Trouver un intervalle de niveau de confiance **non asymptotique** 95% sur λ , de type $[0, \hat{\lambda}_+]$.
9. On suppose le nombre total de repas végétariens K connu. Exprimer la probabilité d'obtenir le repas végétarien désiré (probabilité de non-rupture de stock donc), en fonction de $F_{\lambda'}$, où $F_{\lambda'}$ désigne la fonction de répartition d'une loi de Poisson de paramètre λ' (λ' à déterminer).
10. On note $\psi : \lambda' \mapsto q_{95\%}(\mathcal{P}(\lambda'))$ (quantile d'ordre 95% d'une loi de Poisson de paramètre λ'). Montrer que ψ est croissante.
11. (**) À partir de $\hat{\lambda}$, construire $\hat{T}_0$, temps d'arrivée (maximal) au RU vous garantissant une probabilité globale supérieure à 90% d'avoir votre repas végétarien le $n + 1$ -ème jour (celui après la collecte des n premières données). On supposera ψ connue, et on admettra qu'elle est continue à droite.

III- Test sur le nombre de repas disponibles

On se place enfin dans la situation plus réelle où ni λ ni K ne sont connus. Le CROUS allègue mettre à disposition plus de 100 repas végétariens, et vous voulez vous en assurer, sur la base des observations $(Z_i, N_{i,1})$ collectées comme précédemment.

12. Vous voulez prouver que le CROUS ne respecte pas ses engagements, c'est à dire $K < 100$. Quelles sont les hypothèses du test associé ?
13. Dans le cas où λ est connu, trouver une statistique de test pertinente, et en déduire un test de niveau 95%. *Indication : si ce n'est déjà fait, on pourra prouver que la probabilité de non-rupture de stock de la question 9 est croissante en K et décroissante en λ .*
14. (*) Dans la situation où λ est inconnu, adapter le test précédent pour en déduire un test de niveau 90%.

Solution 1 -

1. On peut supposer les Z_i i.i.d. de loi $\mathcal{B}(p)$, p représentant la probabilité qu'il ne reste plus de plats végétariens quand vous arrivez au stand. Le modèle s'écrit alors

$$(\{0, 1\}^n, \mathcal{P}(\{0, 1\}^n), (\mathcal{B}(p)^{\otimes n})_{p \in [0, 1]}) .$$

2. Par définition, $N = E_p \left(\sum_{i=1}^5 Z_i \right) = 5E_p(Z_1) = 5p$. Un estimateur naturel pour N est alors

$$\hat{N} = 5\bar{Z}_n.$$

3. Comme $E_p|Z_1| = p < +\infty$, la loi des grands nombres donne $\bar{Z}_n \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} p$, et donc

$$\hat{N} \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} N.$$

4. Comme $E_p Z_1^2 = p < +\infty$, le Théorème central limite donne

$$\sqrt{n}(\bar{Z}_n - p) \rightsquigarrow \mathcal{N}(0, p(1-p)).$$

On en déduit

$$\sqrt{n}(\hat{N} - N) \rightsquigarrow \mathcal{N}(0, 25p(1-p)).$$

5. D'après la question précédente, si q_{95} est le quantile d'ordre 95% d'une loi $\mathcal{N}(0, 1)$, on a (application du Lemme de Portmanteau)

$$\mathbb{P}_p \left(\sqrt{\frac{n}{25p(1-p)}}(N - \hat{N}) \leq q_{95} \right) \xrightarrow[n \rightarrow +\infty]{} 95\%.$$

Or

$$\begin{aligned} \sqrt{\frac{n}{25p(1-p)}}(N - \hat{N}) \leq q_{95} &\Leftrightarrow N \leq \hat{N} + \sqrt{\frac{25p(1-p)}{n}}q_{95} \\ &\Rightarrow N \leq \hat{N} + \sqrt{\frac{25}{4n}}q_{95}, \end{aligned}$$

en utilisant $p(1-p) \leq 1/4$. On en déduit alors que $] -\infty; \hat{N} + \sqrt{\frac{25p}{4n}}q_{95}]$ est un intervalle de niveau de confiance asymptotique 95% sur N .

6. On a $\hat{N} = 5 \times (5/25) = 1$, et $\sqrt{25/(4n)}q_{95} = 1.65/2 \in]0, 5; 1[$. Pour être sûr de manger à votre faim sur une semaine il vous faudra préparer 2 tupperwares.

7. En supposant toujours les passages au RU i.i.d., on a le modèle

$$(\mathbb{N}^n, \mathcal{P}(N^n), (\mathcal{P}(\lambda T_1)^{\otimes n})_{\lambda > 0})$$

Un estimateur pour λ est donné par $\hat{\lambda} = \frac{\bar{Y}_{n,1}}{T_1}$.

8. L'inégalité de Bienaymé-Cebicev donne dans ce cas, pour $t > 0$,

$$\mathbb{P}_\lambda \left(|\hat{\lambda} - \lambda| \geq t \right) \leq \frac{\lambda T_1}{n T_1^2 t^2} = \frac{\lambda}{n T_1 t^2}.$$

En notant $\alpha = 5\%$, on en déduit que

$$\mathbb{P}_\lambda \left(\lambda \geq \hat{\lambda} + \sqrt{\frac{\lambda}{n T_1 \alpha}} \right) \leq \alpha.$$

Or, pour tout $\lambda > 0$,

$$\lambda - \sqrt{\frac{\lambda}{nT_1\alpha}} - \hat{\lambda} < 0 \quad \Rightarrow \quad \sqrt{\lambda} < \frac{1}{2\sqrt{nT_1\alpha}} + \frac{1}{2}\sqrt{4\hat{\lambda} + (n\alpha T_1)^{-1}} := \sqrt{\hat{\lambda}_+},$$

en résolvant l'inéquation du second degré en $\sqrt{\lambda}$. On en déduit que $[0; \hat{\lambda}_+[$ est un intervalle de niveau de confiance $1 - \alpha$ sur λ .

9. En notant Y le nombre total de demandes entre l'ouverture du RU et le moment où vous arrivez au stand, on a $Y = Y_0 + Y_1$, avec $Y_0 \sim \mathcal{P}(\lambda T_0)$, $Y_1 \sim \mathcal{P}(\lambda T_1)$, $Y_0 \perp\!\!\!\perp Y_1$. On a alors $Y \sim \mathcal{P}(\lambda(T_1 + T_0))$. La probabilité d'obtenir votre repas végétarien est alors

$$\mathbb{P}(\mathcal{P}(\lambda(T_1 + T_0)) \leq K - 1) = F_{\lambda(T_1+T_0)}(K - 1).$$

10. Soit $\lambda_1 \leq \lambda_2$, et $\delta = \lambda_2 - \lambda_1$. On se donne $N_1 \sim \mathcal{P}(\lambda_1)$ et $N_\delta \sim \mathcal{P}(\delta)$ indépendantes, et on a immédiatement

$$\mathcal{P}(\lambda_1) \sim N_1 \leq N_1 + N_\delta \sim \mathcal{P}(\lambda_2).$$

On en déduit que $\mathcal{P}(\lambda_1) \preceq \mathcal{P}(\lambda_2)$ (domination stochastique), et donc que ψ est croissante.

11. Si λ était connu, il faudrait résoudre l'inéquation

$$K > \psi(\lambda(T_1 + T_0))$$

pour garantir une proba plus grande que 95%. ψ étant croissante, on peut introduire l'inverse généralisé

$$\psi^{-1}(u) = \inf\{t \mid \psi(t) \geq u\},$$

et ψ étant càd, on a l'équivalence $\psi(t) \geq u \Leftrightarrow t \geq \psi^{-1}(u)$, comme pour les fonctions de répartitions en somme. L'équation se réécrit alors

$$\lambda(T_1 + T_0) < \psi^{-1}(K) \quad \Leftrightarrow \quad T_0 < \frac{\psi^{-1}(K)}{\lambda} - T_1,$$

ce qui donne une idée du plus grand T_0 possible. Posons alors

$$\hat{T}_0 = \frac{\psi^{-1}(K)}{\hat{\lambda}_+} - T_1.$$

On a $\mathbb{P}_\lambda(\lambda \geq \hat{\lambda}_+) \leq 5\%$ d'après la question précédente. Pour un nouveau passage au RU, si on arrive au temps $\hat{T}_0$, notons Y_{new} le nombre de commandes effectuées avant votre tour. On a $Y_{new} \mid \{Y_{1:n,1}\} \sim \mathcal{P}(\lambda(T_1 + \hat{T}_0))$, et donc

$$\begin{aligned} \mathbb{P}_\lambda(Y_{new} \leq (K - 1)) &= \mathbb{E}\left(\mathbb{P}_\lambda\left(Y_{new} \leq (K - 1) \mid \hat{\lambda}_+\right)\right) \\ &= \mathbb{E}\left(\mathcal{P}(\lambda\psi^{-1}(K)/\hat{\lambda}_+) \leq K - 1\right) \\ &\geq \mathbb{E}_\lambda\left(F_{\lambda\psi^{-1}(K)/\hat{\lambda}_+}(K - 1)\mathbb{1}_{\hat{\lambda}_+ > \lambda}\right). \end{aligned}$$

Or, en abusant de la relation $\psi(t) \geq u \Leftrightarrow t \geq \psi^{-1}(u)$, on peut écrire

$$\begin{aligned}\hat{\lambda}_+ > \lambda &\Leftrightarrow \psi^{-1}(K) > \frac{\lambda}{\hat{\lambda}_+} \psi^{-1}(K) \\ &\Leftrightarrow \psi\left(\frac{\lambda}{\hat{\lambda}_+} \psi^{-1}(K)\right) < K \\ &\Leftrightarrow \psi\left(\frac{\lambda}{\hat{\lambda}_+} \psi^{-1}(K)\right) \leq K - 1 \\ &\Leftrightarrow F_{\lambda \psi^{-1}(K)/\hat{\lambda}_+}(K - 1) \geq 95\%.\end{aligned}$$

On en déduit

$$\mathbb{P}_\lambda(Y_{new} \leq (K - 1)) \geq 95\% \mathbb{P}_\lambda(\hat{\lambda}_+ > \lambda) \geq 95\%(1 - \mathbb{P}_\lambda(\hat{\lambda}_+ \leq \lambda)) \geq 90\%,$$

on est donc sûrs à 90% d'avoir à manger le jour suivant la collecte de données en suivant cette stratégie.

12. On met en alternative ce dont on veut être sûr, les hypothèses sont alors $H_0 : K \geq 100$ et $H_1 : K < 100$.
13. Notons $S = \sum_{i=1}^n Z_i$ (nombre de ruptures de stock observées). Sous H_1 , on s'attend à ce que S soit plus grande que sous H_0 . La zone de rejet correspondante est alors $\llbracket s_\alpha, n \rrbracket$. On note $T = \mathbb{1}_{S \geq s_\alpha}$.

Notons $p(\lambda, K) = \mathbb{P}_{\lambda, K}(Z_1 = 1)$. D'après la question 9–, on a $p(\lambda, K) = 1 - F_{\lambda(T_1+T_0)}(K - 1)$, décroissante en K . Comme $p_1 \leq p_2$ implique $\mathcal{B}(n, p_1) \preceq \mathcal{B}(n, p_2)$, on en déduit alors que

$$\begin{aligned}\sup_{K \geq 100} \mathbb{P}_{\lambda, K}(T = 1) &= \sup_{K \geq 100} \mathbb{P}(\mathcal{B}(n, p(\lambda, K)) \geq s_\alpha) \\ &= \mathbb{P}(\mathcal{B}(n, p(\lambda, 100)) \geq s_\alpha).\end{aligned}$$

Si on choisit s_α comme $q_{1-\alpha}(\mathcal{B}(n, p(\lambda, 100))) + 1$, on a un test de niveau α .

14. Dans le cas où λ est inconnu, la statistique de test et la forme restent inchangées, le seuil lui est plus délicat à calibrer. On va encore se servir de la majoration de λ donnée avec forte probabilité par $\hat{\lambda}_+$. Choisissons

$$s_\alpha = q_{1-\alpha}(\mathcal{B}(n, p(\hat{\lambda}_+, 100))) + 1.$$

On a alors

$$\begin{aligned}\sup_{K \leq 100, \lambda} \mathbb{P}_{\lambda, K}(T = 1) &= \sup_{\lambda} \mathbb{P}\left(\mathcal{B}(n, p(\lambda, 100)) \geq q_{1-\alpha}(\mathcal{B}(n, p(\hat{\lambda}_+, 100))) + 1\right) \\ &\leq \sup_{\lambda} \mathbb{P}\left([\mathcal{B}(n, p(\lambda, 100)) \geq q_{1-\alpha}(\mathcal{B}(n, p(\hat{\lambda}_+, 100))) + 1] \cap \{\lambda < \hat{\lambda}_+\}\right) \\ &\quad + \sup_{\lambda} \mathbb{P}\left([\mathcal{B}(n, p(\lambda, 100)) \geq q_{1-\alpha}(\mathcal{B}(n, p(\hat{\lambda}_+, 100))) + 1] \cap \{\lambda \geq \hat{\lambda}_+\}\right) \\ &\leq 5\% + \sup_{\lambda} \mathbb{P}\left([\mathcal{B}(n, p(\lambda, 100)) \geq q_{1-\alpha}(\mathcal{B}(n, p(\hat{\lambda}_+, 100))) + 1] \cap \{\lambda < \hat{\lambda}_+\}\right).\end{aligned}$$

Comme p est croissante en λ (ordre stochastique des $\mathcal{P}(\lambda(T_0 + T_1))$), l'ordre stochastique sur les $\mathcal{B}(n, p)$ donne

$$q_{1-\alpha} \left(\mathcal{B}(n, p(\hat{\lambda}_+, 100)) \right) \mathbb{I}_{\lambda < \hat{\lambda}_+} \geq q_{1-\alpha} \left(\mathcal{B}(n, p(\lambda, 100)) \right) \mathbb{I}_{\lambda < \hat{\lambda}_+},$$

et donc

$$\begin{aligned} & \sup_{\lambda} \mathbb{P} \left([\mathcal{B}(n, p(\lambda, 100)) \geq q_{1-\alpha} \left(\mathcal{B}(n, p(\hat{\lambda}_+, 100)) \right) + 1] \cap \{\lambda < \hat{\lambda}_+\} \right) \\ & \leq \sup_{\lambda} \mathbb{P} \left([\mathcal{B}(n, p(\lambda, 100)) \geq q_{1-\alpha} \left(\mathcal{B}(n, p(\lambda, 100)) \right) + 1] \cap \{\lambda < \hat{\lambda}_+\} \right) \\ & \leq \alpha. \end{aligned}$$

On en d duit que $T = 1_{S \geq s_\alpha}$ est de niveau 90%.