MACS 205 : interpolation et quadrature
Étude de Cas

1 Introduction

L'étude de cas consiste à utiliser les résultats théoriques et les méthodes numériques vus en cours et TP pour l'étude de la densité et de la fonction de répartition de la loi Gamma de paramètres $(k,\theta) \in (\mathbb{R}_+^*)^2$ notée $\mathrm{Gamma}_{k,\theta}$, de densité

$$\phi_{k,\theta}(x) = \frac{x^{k-1}e^{-x/\theta}}{\Gamma(k)\theta^k} \tag{1}$$

où $\Gamma(k)$ est la fonction Gamma d'Euler, $\Gamma(k) = \int_0^\infty t^{k-1}e^{-t}\,\mathrm{d}t$, $k > 0$. La fonction de répartition de la loi est

$$F_{k,\theta}(x) = \frac{1}{\Gamma(k)}\mathrm{G}(k, x/\theta)$$

où G est la fonction Gamma incomplète

$$\mathrm{G}(k,y) = \int_0^y t^{k-1}e^{-t}\,\mathrm{d}t, \tag{2}$$

qui n'a pas d'expression explicite. Avec ces notations $\Gamma(k) = \lim_{y\to\infty}\mathrm{G}(k,y)$.

Dans toute la suite, on prendra

$$k = 2.1, \theta = 0.5$$

La première partie du sujet vous amènera à utiliser les méthodes d'interpolation vues en cours sur la densité $\phi_{k,\theta}$ sur l'intervalle $[a,b] = [0,14]$. La deuxième partie consiste à exploiter les méthodes d'intégration numériques pour approcher la quantité

$$p_{a,b} = \mathbb{P}(a \le X_1 \le b) = \int_a^b \phi_{k,\theta}(t)\,\mathrm{d}t. \tag{3}$$

avec $a = 0, b = 3.5$.

Dans `R` les valeurs de $p_{0,b}$ sont tabulées numériquement avec une grande précision grâce à la fonction de répartition de la loi gamma, `pgamma`.

Sur le site se trouve un petit fichier `ini.R` contenant le code de la fonction de densité (fonction `densite`) et les paramètres globaux du problème k, θ codés en majuscules par convention.

NB la fonction `densite` peut prendre un vecteur en argument, elle renvoie alors un vecteur de même taille. Faites le test.

Consignes

L'ensemble de votre travail devra être rassemblé dans un seul dossier portant votre nom, et archivé au format .zip ou tar.gz. Ce dossier doit contenir

- Un Notebook R compilé sous forme html ou pdf, obtenu avec Rstudio ou jupyter notebook, contenant les réponses rédigées au questions, les figures et résultats numériques. Tous les résultats doient êre commentés et interprétés.
- le code du notebook, exécutable sans modification de la part du correcteur (mettez des chemins d'accès relatifs pour les fichiers auxiliaires)
- Toutes les fonctions utilisées dans le notebook (y compris les fonctions codées en cours) rassemblées dans un fichier `functions_ETC2021.R`

Les dossiers doivent être déposés sur le site pédagogique ecampus avant le dimanche 11 avril 2021, 23h59. Chaque heure de retard coûte deux points (sur vingt).

En cas de problème d'ordre technique, vous pouvez envoyer un mail à l'adresse : `anne.sabourin@telecom-paris.fr`, avec en objet l'intitulé “EDC MACS205”.

2 Interpolation polynomiale

1. Autour de quel degré d'interpolation le phénomène de Runge apparaît-il lorsque l'on interpole $\phi_{k,\theta}$ sur l'intervalle $[a = 0, b = 14]$, spectivement pour l'interpolation aux noeuds équidistants et aux noeuds de Tchebychev ? Affichez les graphiques correspondants.
2. on s'autorise une interpolation par morceaux :
   (a) Quelle est le nombre d'évaluations de la fonction 'densite' nécéssaires pour appliquer l'interpolation de lagrange de degré n par morceaux avec M sous intervalles, (i) avec noeuds équidistants (ii) avec noeuds de Tchebychev ?
   (b) Ecrire un algorithme permettant de choisir automatiquement la paire (M, n) comme ci-dessus avec un budget maximal de 150 évaluations de la fonction 'densite' en minimisant l'erreur d'interpolation. On considérera séparément le cas des noeuds équi distants et des noeuds de Tchebychev.Pour chaque degré n on prendra le nombre de sous intervalle le plus grand possible sous la contrainte de budget. On approchera la norme sup de l'erreur par le maximum de l'erreur sur un vecteur de 2000 points équirepartis sur $[a = 0, b = 14]$.
   Donnez la paire optimale (M, n) et l'erreur associée, pour les deux types de noeuds (equidistants et de Tchebychev). Tracez également les deux courbes de l'erreur en fonction du degré n. Commentez.

3 Méthodes de quadrature

On veut ici estimer la valeur de l'intégrale

$$I = \int_a^b \phi_{k,\theta}.$$

avec $a = 0, b = 3.5$.

3.1 Méthode de Simpson

Utilisons la méthode de quadrature composite de Cavalieri-Simpson On appelle $\hat{I}_M^{simp}$ le résultat obtenu par une méthode de simpson composite avec M intervalles de même taille.

1. On vous donne un budget de 81 évaluations de la fonction `densite`.
   (a) À combien de sous-intervalles d'intégration ce budget vous donne-t-il droit ?
   (b) Donnez une approximation a posteriori de l'erreur commise avec ce budget ainsi que la valeur de $\hat{I}_M^{simp}$ associée, au niveau de précision correspondant à l'approximation de l'erreur.
   (Attention, la fonction codée en TP renvoie une erreur correspondant au double du paramètre M passé en argument).
   (c) Quelle est l'erreur relative commise sur l'erreur d'intégration, c'est à dire $|\widehat{E}(M)-E(M)|/E(M)$, où $\widehat{E}(M)$ est l'approximation a posteriori de la question précédente, et $E(M)$ est la vraie erreur, calculable en comparant $\hat{I}_M^{simp}$ et la fonction de répartition de la loi Gamma (fonction `pgamma`) ?

3.2 Méthode de Romberg

Dans cette partie, on cherche à évaluer à nouveau I par la méthode des trapèzes, avec une étape d'extrapolation de Romberg.

Estimez la valeur de I avec la méthode de Romberg : Passez en argument de la fonction `romberg` un nombre initial d'intervalles $M = 2$ et un nombre de raffinements successifs $n = 20$.

1. Nommez `estRomberg` le vecteur renvoyé par la fonction `Romberg` avec les paramètres indiqués, et affichez-le grahiquement.
2. On prendra comme "vraie" valeur de l'intégrale la quantité calculée par la fonction `pgamma`. Tracez sur le même graphique : le vecteur `0:n` en abscisses ; et en ordonnées : le logarithme de l'erreur $\log |$ `estRomberg` $-I|$ et le logarithme de l'erreur des estimateurs "naïfs" correspondants, c'est à dire les $\log |\hat{I}_m - I|$, pour $m = M, 2M, \ldots, 2^n M$. Que constatez-vous ?