EPFL

Physique Statistique PHYS 338
2024-2025

Homework 1: Un peu de probabilités

Consignes de rendu : Pour les questions numériques (indiqué par un astérisque), veuillez soumettre un notebook Jupyter qui peut être exécuté de bout en bout sans erreur et qui reproduit tous les résultats demandés dans ce document.Assurez-vous que votre code soit clair, bien commenté et que toutes les dépendances nécessaires soient clairement indiquées. Pour les questions théoriques, vous avez deux options :

- Vous pouvez inclure vos réponses directement dans le notebook Jupyter, en utilisant des cellules de texte Markdown pour une présentation claire et structurée.
- Alternativement, vous pouvez soumettre un fichier PDF séparé contenant vos réponses aux questions théoriques. Dans ce cas, assurez-vous que vos réponses soient bien organisées et correspondent clairement aux numéros des questions.

Dans les deux cas, veillez à ce que vos explications soient claires, concises et rigoureuses. N'hésitez pas à inclure des schémas ou des équations lorsque cela est pertinent pour illustrer vos raisonnements.

Cette série d'exercices est conçue pour approfondir votre compréhension des concepts fondamentaux en théorie des probabilités et en statistique. Le devoir est structuré de manière séquentielle, vous guidant à travers une progression logique des concepts. Nous commencerons par explorer la loi des grands nombres, pierre angulaire de la théorie des probabilités. Ensuite, nous aborderons le théorème central limite, qui décrit le comportement asymptotique des sommes de variables aléatoires. Nous examinerons ensuite les corrections au théorème central limite, affinant notre compréhension de ses limites et applications. Enfin, on va traite la théorie des grandes déviations, qui s'intéresse aux événements rares mais significatifs.

Chaque partie comprend à la fois des exercices théoriques et des exercices numériques. Les premiers pour consolider votre compréhension conceptuelle, et les deuxièmes pour tester et appliquer ces concepts dans des situations pratiques.

Bon travail!

Exercise 1 Échauffement et Loi de Grand Nombre

Q1. On note $Q(t) := \mathbb{P}[T > t] = 1 - F(t)$ la probabilité qu'un événement ne se soit pas encore produit après un temps t ($F(t)$ désigne la fonction de distribution cumulative). On suppose $t \geq 0$. Sachant que le taux de changement de cette probabilité est constant et égal à $-\lambda$, c'est-à-dire

$$\frac{\mathrm{d}Q(t)}{\mathrm{d}t} = -\lambda Q(t)\,. \tag{1}$$

a) Dérivez la forme de la distribution $Q(t)$.

b) Fixez la constante indéterminée dans la question précédente de telle sorte que la probabilité totale sur tous les temps possibles est bien égale à 1. Quelle est la forme finale de la **densité** de probabilité $p_\lambda(t)$ pour la distribution?

c) Calculer le valeur moyenne (espérance) et la variance de la variable aléatoire correspondante en fonction du paramètre λ.

Q2. Supposons que vous ne disposiez que d'un générateur de nombres aléatoires uniformes sur l'intervalle [0,1]. Montrez comment, grâce à un changement de variable judicieux, vous pouvez utiliser cette distribution uniforme pour générer des échantillons suivant la distribution de paramètre λ obtenue en **Q1**.

* **Q3.** Le but de cette question est de créer un générateur d'échantillons de cette distribution avec un paramètre λ donné.

a) Créez dans un notebook la fonction `generator(M: int, lam:  float) -> np.ndarray` qui prend en entrée le nombre M d'échantillons à générer et le paramètre λ de la distribution, et renvoie un tableau numpy contenant les échantillons générés.

b) Fixez $N = 10^5$ et $\lambda = 2$ puis créez un histogramme, normalisé pour que ce soit un densité [1], des échantillons générés et superposez cet histogramme avec la courbe de densité théorique de la distribution obtenue en **Q1**.

Q4. Soit $X_1, X_2, \ldots, X_N \overset{\text{iid}}{\sim} p_\lambda$ des variables aléatoires indépendantes et identiquement distribuées selon la distribution de la **Q1**. On définit la somme S_N et la moyenne empirique $\hat{\mu}_N$ comme

$$S_N = \sum_{i=1}^{N} X_i\,, \quad \hat{\mu}_N = \frac{S_N}{N}\,. \tag{2}$$

a) Vers quelle valeur $\hat{\mu}_N$ converge-t-elle lorsque N tend vers l'infini? Justifiez votre réponse.

b) Calculez la variance de la variable aléatoire $\hat{\mu}_N$. Comment cette variance évolue-t-elle lorsque N augmente ? Que pouvez-vous en conclure sur la précision de l'estimation de $\mathbb{E}[X]$ par $\hat{\mu}_N$?

c) On définit la variance **empirique** $\hat{\sigma}_N^2$ comme suit :

$$\hat{\sigma}_N^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \hat{\mu}_N)^2 \tag{3}$$

Calculez $\mathbb{E}[\hat{\sigma}_N^2]$. Cette estimation de la variance est-elle biaisée[2] ? Si oui, expliquez la source du biais.

d) Proposez une correction à l'estimateur $\hat{\sigma}_N^2$ permettant d'obtenir un estimateur non biaisé de la variance. Justifiez votre réponse.

* **Q5.** Écrivez un code Python qui calcule $\hat{\mu}_N$ pour différentes valeurs de N avec le code créé dans **Q3**. Pour $\lambda = 2$, tracez deux graphiques: le premier est un graphique montrant la moyenne empirique $\hat{\mu}_N$ en fonction de N, pour N entre 10 et 10^7. Le second est un graphique montrant la variance de $\hat{\mu}_N$ en fonction de N, pour N entre 100 et 10^5. Commentez les résultats obtenus. Ces résultats sont-ils cohérents avec la théorie?

Solution of Exercise 1

Q1. a) On peut modeliser le temp quand l'evenment se passe comme un variable aleatoire que on appelle T. La densite de probabilite que on nous demande est

$$Q(t) = \mathbb{P}[T > t] \tag{4}$$

ou on peut seulement considérée de temps qui sont positif. On reconnaît que $Q(t)$ est liée a la fonction de répartition de la variable T parce que

$$F_T(t) = \mathbb{P}[T \leq t] = 1 - \mathbb{P}[T > t] = 1 - Q(t) \tag{5}$$

[1] avec matplotlib cela se fait en mettant l'option `True` à l'option densité: `matplotlib.pyplot.hist( ... , density=True)`.

[2] Un estimateur $\hat{\lambda}$ est dit non-biaisé si et seulement si $\mathbb{E}[\hat{\lambda}] = \lambda$.

Pour dériver la forme de la distribution $P(t)$, nous résolvons l'équation différentielle donne dans la consigne. L'équation est

$$\frac{\mathrm{d}Q(t)}{\mathrm{d}t} = -\lambda Q(t) \tag{6}$$

On veut procéder avec la méthode de séparation des variables, donc

$$\frac{\mathrm{d}Q}{Q} = -\lambda dt \tag{7}$$

En intégrant des deux côtés

$$\ln Q(t) = -\lambda t + C \tag{8}$$

où C est une constante d'intégration. En prenant l'exponentielle

$$Q(t) = Ae^{-\lambda t} \tag{9}$$

où $A = e^C$ est une constante à déterminer.

b) Pour normaliser la distribution on peut bien imposer que $F_T(0) = 0$, ou $Q(0) = 1$ ou normaliser la densité de probabilité. On procède avec $Q(0) = 1$, ce qui implique $A = 1$. Ainsi

$$Q(t) = e^{-\lambda t} \tag{10}$$

La densité de probabilité $p(t)$ est la dérivée négative de $Q(t)$ parce que

$$p(t) = \frac{\mathrm{d}F_T(t)}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}[1 - Q(t)] = -\frac{\mathrm{d}Q(t)}{\mathrm{d}t} = \lambda e^{-\lambda t} \tag{11}$$

On peut vérifier que aussi la densité est normalisée

$$\int_0^\infty p(t)dt = \int_0^\infty \lambda e^{-\lambda t}dt = 1 \tag{12}$$

Ainsi, la forme finale de la densité de probabilité pour la distribution exponentielle est

$$p(t) = \begin{cases} \lambda e^{-\lambda t} & \text{pour } t \geq 0 \\ 0 & \text{sinon} \end{cases}. \tag{13}$$

c) Calculons l'espérance et la variance de un variable aléatoire T avec cette distribution de probabilité. La moyenne μ est donnée par

$$\mu = \mathbb{E}[T] = \int_0^\infty tp(t)\,\mathrm{d}t = \int_0^\infty t\lambda e^{-\lambda t}\,\mathrm{d}t \tag{14}$$

En intégrant par parties on peut obtenir l'integral indefini

$$\int t\lambda e^{-\lambda t}\,\mathrm{d}t = -e^{-\lambda t}t + \int e^{\lambda t}\,\mathrm{d}t = -\frac{e^{-\lambda t}(\lambda t + 1)}{\lambda} + C \tag{15}$$

donc on peut directement ecrire que

$$\mu = -\frac{e^{-\lambda t}(\lambda t + 1)}{\lambda}\bigg|_{t=0}^{\infty} = 0 + \frac{1}{\lambda} = \frac{1}{\lambda} \tag{16}$$

On continue avec la variance. La variance σ^2 est donnée par $\mathbb{E}[T^2] - \mu^2$. Calculons d'abord $\mathbb{E}[T^2]$.

$$\begin{aligned}\mathbb{E}[T^2] &= \int_0^\infty t^2\lambda e^{-\lambda t}\,\mathrm{d}t = -t\frac{e^{-\lambda t}(\lambda t + 1)}{\lambda}\bigg|_{t=0}^{\infty} + \int_0^\infty \frac{e^{-\lambda t}(\lambda t + 1)}{\lambda}\,\mathrm{d}t \\ &= -t\frac{e^{-\lambda t}(\lambda t + 1)}{\lambda}\bigg|_{t=0}^{\infty} + \int_0^\infty e^{-\lambda t}t\,\mathrm{d}t + \frac{1}{\lambda}\int_0^\infty e^{-\lambda t}\,\mathrm{d}t \\ &= 0 + \frac{1}{\lambda^2} + \frac{1}{\lambda^2} = \frac{2}{\lambda^2}\end{aligned} \tag{17}$$

Ou on a intégrée par parties deux fois et l'integral d'avant on nous aide a faire seulement un integration par partie. Maintenant, nous pouvons calculer la variance :

$$\sigma^2 = \mathbb{E}[T^2] - \mu^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} \tag{18}$$

On peut noter que l'écart-type $\sigma = \frac{1}{\lambda}$ est égal à la moyenne, ce qui est une propriété caractéristique de la distribution exponentielle.

Q2. Pour générer des échantillons suivant une distribution exponentielle à partir d'une distribution uniforme, nous pouvons utiliser *la méthode de la transformation inverse*. Soit U une variable aléatoire uniformément distribuée sur $[0, 1]$, i.e. $U \sim \text{Unif}([0, 1])$. L'idée est de trouver une fonction g telle que $X = g(U)$ suive une distribution exponentielle. Pour faire ça, on peut constater que de manière générale

$$F_X(x) = \mathbb{P}[X \leq x] = \mathbb{P}[g(U) \leq x] = \mathbb{P}[U \leq g^{-1}(x)] = g^{-1}(x) \tag{19}$$

et donc on peut prendre $g(\cdot) = F_X^{-1}(\cdot)$. Ça veut dire $U = F_X(X)$, et donc

$$U = 1 - e^{-\lambda X} \tag{20}$$

En résolvant pour X, on obtient :

$$1 - U = e^{-\lambda X} \tag{21}$$

$$-\ln(1 - U) = \lambda X \tag{22}$$

$$X = -\frac{1}{\lambda} \ln(1 - U) \tag{23}$$

Ainsi, si U est uniformément distribué sur [0,1], alors $X = -\frac{1}{\lambda} \ln(1 - U)$ suivra une distribution exponentielle de paramètre λ. Comme U est uniformément distribuée sur $[0, 1]$, $1-U$ l'est aussi. Donc, en pratique, on peut simplifier l'expression en

$$X = -\frac{1}{\lambda} \ln(U) \tag{24}$$

Donc l'algorithme a trois étapes

1. Générer U uniformément sur [0,1]
2. Calculer $X = -\frac{1}{\lambda} \ln(U)$

Q3. Voir le Notebook.

Q4. a) D'après la loi des grands nombres, la moyenne empirique $\hat{\mu}_N$ converge en probabilité vers l'espérance mathématique de la distribution sous-jacente lorsque n tend vers l'infini. Pour la distribution exponentielle de paramètre λ, nous savons que $\mathbb{E}[X] = \frac{1}{\lambda}$. Donc

$$\lim_{N \to \infty} \hat{\mu}_N = \frac{1}{\lambda} \tag{25}$$

b) Pour calculer la variance de $\hat{\mu}_N$, nous utilisons les propriétés de la variance et le fait que les X_i sont indépendantes

$$\text{Var}(\hat{\mu}_N) = \text{Var}\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) = \frac{1}{N^2}\text{Var}\left(\sum_{i=1}^{N} X_i\right) = \frac{1}{N^2}\sum_{i=1}^{N}\text{Var}(X_i) = \frac{1}{N^2} \cdot N \cdot \text{Var}(X) \tag{26}$$

Sachant que pour une distribution exponentielle, $\text{Var}(X) = \frac{1}{\lambda^2}$, nous obtenons

$$\text{Var}(\hat{\mu}_N) = \frac{1}{N^2} \cdot N \cdot \frac{1}{\lambda^2} = \frac{1}{N\lambda^2} \tag{27}$$

La diminution de la variance avec l'augmentation de N indique que l'estimateur $\hat{\mu}_N$ devient de plus en plus précis pour estimer $\mathbb{E}[X]$. Cette analyse montre que $\hat{\mu}_N$ est un estimateur consistant et non biaisé de $\mathbb{E}[X]$ pour la distribution exponentielle, avec une précision qui s'améliore à mesure que la taille de l'échantillon augmente.

c) Pour calculer $\mathbb{E}[\hat{\sigma}_N^2]$, nous allons utiliser la décomposition suivante

$$\hat{\sigma}_N^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \hat{\mu}_N)^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i^2 - 2X_i\hat{\mu}_N + \hat{\mu}_N^2) \tag{28}$$

$$= \frac{1}{N}\sum_{i=1}^{N} X_i^2 - 2\hat{\mu}_N^2 + \hat{\mu}_N^2 = \frac{1}{N}\sum_{i=1}^{N} X_i^2 - \hat{\mu}_N^2 \tag{29}$$

Prenons l'espérance des deux côtés

$$\mathbb{E}[\hat{\sigma}_N^2] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} X_i^2\right] - \mathbb{E}[\hat{\mu}_N^2] = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[X_i^2] - \mathbb{E}[\hat{\mu}_N^2] \tag{30}$$

En substituant ces valeurs obtenu a la question **Q1.c**. De plus, $\mathbb{E}[\hat{\mu}_N^2] = \text{Var}(\hat{\mu}_N) + (\mathbb{E}[\hat{\mu}_N])^2 = \frac{1}{N\lambda^2} + \frac{1}{\lambda^2}$.

$$\mathbb{E}[\hat{\sigma}_N^2] = \frac{2}{\lambda^2} - \left(\frac{1}{N\lambda^2} + \frac{1}{\lambda^2}\right) = \frac{2}{\lambda^2} - \frac{1}{N\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} - \frac{1}{N\lambda^2} = \frac{N-1}{N\lambda^2} \tag{31}$$

L'estimation de la variance est biaisée car $\mathbb{E}[\hat{\sigma}_N^2] \neq \text{Var}(X) = \frac{1}{\lambda^2}$.
Le biais provient de l'utilisation de la moyenne empirique $\hat{\mu}_N$ dans le calcul de $\hat{\sigma}_N^2$. Cette moyenne est elle-même une variable aléatoire, ce qui introduit une dépendance entre les termes $(X_i - \hat{\mu}_N)^2$. Cette dépendance conduit à une sous-estimation systématique de la vraie variance.

d) Pour obtenir un estimateur non biaisé de la variance, nous pouvons multiplier $\hat{\sigma}_N^2$ par un facteur correctif

$$\hat{\sigma}_{N,\text{corrigé}}^2 = \frac{N}{N-1}\hat{\sigma}_N^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \hat{\mu}_N)^2 \tag{32}$$

Avec le resultat de la question precedent nous obtenons :

$$\mathbb{E}[\hat{\sigma}_{N,\text{corrigé}}^2] = \frac{N}{N-1}\mathbb{E}[\hat{\sigma}_N^2] = \frac{N}{N-1} \cdot \frac{N-1}{N\lambda^2} = \frac{1}{\lambda^2} \tag{33}$$

Cette valeur est exactement égale à $\text{Var}(X)$ pour la distribution exponentielle. Cet estimateur corrigé est connu sous le nom de variance empirique non biaisée ou variance empirique corrigée. Il est couramment utilisé en statistiques pour estimer la variance d'une population à partir d'un échantillon.
Cette correction est valable non seulement pour la distribution exponentielle, mais pour toute distribution avec une variance finie. Pour de grandes valeurs de n, la différence entre l'estimateur biaisé et non biaisé devient négligeable.

Q5. Voir le notebook

Exercise 2 Fonction caractéristique et TCL

Soit $i = \sqrt{-1}$ le nombre imaginaire unité, la fonction caractéristique, de la densité de probabilité p_X est définie de la façon suivante

$$\varphi_X(t) = \mathbb{E}\big[e^{itX}\big] = \int_{-\infty}^{\infty} p_X(x) e^{itx}\,\mathrm{d}x \ . \tag{34}$$

Par développement de l'exponentielle en série entière et interversion de la série et de l'espérance, dans un voisinage de l'origine, on a le développement **en série formelle**:

$$\varphi_X(t) = 1 + \sum_{k=1}^{\infty} m_{k,X} \frac{(it)^k}{k!} \tag{35}$$

où $m_{k,X} := \mathbb{E}[X^k]$ est le k^{eme} moment de la variable aléatoire X quand ce dernier existe. On définit le log de la fonction caractéristique comme la fonction $H_X(t) := \log \varphi_X(t)$ et on définit $n$ème cumulant de X, $\kappa_{n,X}$ comme le coefficient de la série de Taylor de cette fonction

$$H_X(t) = \log \varphi_X(t) = \sum_{l=1}^{\infty} \kappa_{l,X} \frac{(it)^l}{l!} \tag{36}$$

Dans la première partie on veut utiliser cette definition pour redémontrer le TCL. Si il est pas explicitement dit, considérée que la variable aléatoire X a les trois premières moments fini.

Q1. Montrer que $\varphi_X(t)$ existe pour toute distribution p_X et que $|\varphi_X(t)| \leq 1$.

Q2. Comment obtient-on $p_X(x)$ à partir de $\varphi_X(t)$?

Q3. Si $\varphi_X(t)$ a un rayon de convergence $R > 0$ a l'origine, que peut-on en déduire sur les moments de la variable aléatoire X ?

Q4. Comment s'expriment les quatre premiers cumulants en fonction des moments de X?

Q5. Quelle est la transformée de Fourier d'une Gaussienne ? En déduire que les cumulants d'une variable Gaussiene, *i.e.* $\kappa_{n,X}$ pour $X \sim \mathcal{N}(\mu, \sigma^2)$, sont nuls pour $n > 2$, i.e. $\kappa_{n,X} = 0\, \forall n > 2$.

Q6. On considère $S_N = X_1 + \cdots + X_N$, où les X_i sont N variables aléatoires indépendantes distribuées selon la loi de densité p_X. Que vaut la fonction génératrice de S_N en fonction de celle de X? Que vaut son logarithme $H_{S_N}(t)$ en fonction de $H_X(t)$? Que valent les cumulants de S_N en fonction de ceux de X?

Q7. Dans la suite, on note par

$$G_N = \frac{1}{\sqrt{\kappa_{2,X}}} \cdot \left(\frac{S_N}{\sqrt{N}} - \sqrt{N} \cdot m_{1,X} \right) . \tag{37}$$

Montrez que le logarithme de la fonction caractéristique de G_N — la fonction $H_{G_N}(t)$ — satisfait

$$H_{G_N}(t) = N H_{\frac{X-\mu}{\sigma}} \left(\frac{t}{\sqrt{N}} \right) . \tag{38}$$

Quelle est la limite des fonctions $H_{G_N}(t)$ (et donc de la fonction caractéristique $\varphi_{G_N}(t)$ de G_N) dans la limite ou N est très grand? En déduire que la distribution de G_N converge vers une distribution Gaussienne de moyenne *nulle* et de variance *unité* dans la limite.

Q8. Tentons de répondre à la question suivante : Est-il nécessaire que tous les X_i aient la même distribution pour obtenir la convergence vers la Gaussienne ? On va supposer d'avoir N variable aléatoire $\{X_i\}_{i=1}^N$ indépendants avec $\mathbb{E}\big[X_i^2\big] < \infty$ (mais pas nécessairement identiquement distribuées).

a) Pour des variables indépendantes (mais pas nécessairement identiquement distribuées), que vaut l'espérance et la variance de leur somme ? En déduire la construction de la variable $\tilde{G}_N$ de moyenne nulle et de variance un.
b) Exprimez la fonction $H_{\tilde{G}_N}(t)$ en tant que fonction des fonctions $H_{X_i}(t)$ des variables aléatoires individuelles.
c) En utilisant le théorème de Taylor avec reste de Peano, développez les fonctions H_{X_i} et calculez la fonction $H_{\tilde{G}_N}(t)$. Que peut-on dire du terme résiduel correspondent ? Pourquoi tend-il vers 0 lorsque N augmente ?
d) Prouvez donc la généralisation du résultat de la **Q7** pour ce cas.

Q9. Calculez la fonction caractéristique $\varphi_X(t)$ de la distribution obtenue à la **Q1** de l'**Exercise 1**.

* **Q10.** En utilisant le générateur de cette même distribution, que vous avez implémenté dans l'exercice précédent, tracez pour différentes valeurs de $N \in \{4, 16, 64, 256, 1024, 4096\}$ l'histogramme (normalisé pour que cela soit une densité de probabilité) de $\rho_N := \sqrt{N}(\hat{\mu}_N - \mathbb{E}[X])$, où $\hat{\mu}_N$ est la moyenne empirique de N échantillons, et comparez-le à une distribution gaussienne ayant la moyenne et variance donnée par le TCL.

Solution of Exercise 2

Q1. On commence avec l'existence de $\varphi_X(t)$. $\varphi_X(t)$ est définie comme une espérance mathématique. Elle existe si l'intégrale converge absolument. Vu que $|e^{itx}| = 1$ pour tout x réel on a que

$$\int_{-\infty}^{\infty} |p_X(x)e^{itx}| \, \mathrm{d}x = \int_{-\infty}^{\infty} p_X(x) \, \mathrm{d}x = 1 < \infty \tag{39}$$

L'intégrale converge car $p_X(x)$ est une densité de probabilité.
On procède pour montrer l'inégalité. On a que

$$\begin{aligned}
|\varphi_X(t)| &= \left| \int_{-\infty}^{\infty} p_X(x)e^{itx} \, \mathrm{d}x \right| && (40)\\
&\leq \int_{-\infty}^{\infty} |p_X(x)e^{itx}| \, \mathrm{d}x \quad \text{(par l'inégalité triangulaire pour les intégrales)} && (41)\\
&= \int_{-\infty}^{\infty} p_X(x)|e^{itx}| \, \mathrm{d}x \quad \text{(car } p_X(x) \text{ est réelle et non-négative)} && (42)\\
&= \int_{-\infty}^{\infty} p_X(x) \, \mathrm{d}x \quad \text{(car } |e^{itx}| = 1 \text{ pour tout } x \text{ réel)} && (43)\\
&= 1 \quad \text{(car } p_X(x) \text{ est une densité de probabilité)} && (44)
\end{aligned}$$

Donc, $|\varphi_X(t)| \leq 1$.

Q2. On peut obtenir $p_X(x)$ à partir de $\varphi_X(t)$ en utilisant la transformée de Fourier inverse. La fonction génératrice $\varphi_X(t)$ est en fait la transformée de Fourier de $p_X(x)$. Donc

$$p_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_X(t)e^{-itx} \, \mathrm{d}t \tag{45}$$

Cette formule permet de retrouver la densité de probabilité à partir de la fonction génératrice, à condition que $\varphi_X(t)$ soit intégrable.

Q3. if $R > 0$ then $\phi(t)$ is analytical on $\mathbb{B}(R)$ and in particular $\phi^{(k)}(0)$ is well defined for any k. Since the latter is also equal to $i^k \mathbb{E}X^k$, it implies the existence of all moments.

Q4. Pour exprimer les cumulants en fonction des moments, nous utilisons le développement en série de Taylor de $\log \varphi_X(t)$ et le comparons avec celui de $\varphi_X(t)$. Soit $m_{n,X} = \mathbb{E}[X^N]$ le n-ième moment de X.

$$\varphi_X(t) = \int_{-\infty}^{\infty} p_X(x) e^{itx}\, \mathrm{d}x \tag{46}$$

$$= \int_{-\infty}^{\infty} p_X(x) \left[1 + itx - \frac{t^2x^2}{2} - \frac{1}{6} it^3x^3 + \frac{t^4x^4}{24} + \mathcal{O}(t^5)\right] \mathrm{d}x \tag{47}$$

$$:= 1 + itm_{1,X} - \frac{t^2}{2!} m_{2,X} - \frac{it^3}{3!} m_{3,X} + \frac{t^4}{4!} m_{4,X} + \mathcal{O}(t^5) \tag{48}$$

on a donc pour la fonction $H_X(t)$ que on peut aussi développer le log et obtenir que

$$H_X(t) = im_{1,X}t + \frac{1}{2}t^2\left(m_{1,X}^2 - m_{2,X}\right) - \frac{1}{6}it^3\left(2m_{1,X}^3 - 3m_{1,X}m_{2,X} + m_{3,X}\right) \tag{49}$$

$$+ \frac{1}{24}t^4\left(-6m_{1,X}^4 + 12m_{1,X}^2 m_{2,X} - 4m_{1,X}m_{3,X} - 3m_{2,X}^2 + m_{4,X}\right) + \mathcal{O}(t^5) \tag{50}$$

$$:= \kappa_{1,X}\frac{it}{1!} + \kappa_{2,X}\frac{(it)^2}{2!} + \kappa_{3,X}\frac{(it)^3}{3!} + \kappa_{4,X}\frac{(it)^4}{4!} + \mathcal{O}(t^5) \tag{51}$$

En comparant ces deux expressions, nous obtenons

- Premier cumulant : $\kappa_1 = \mu_1 = \mathbb{E}[X]$ (moyenne)
- Deuxième cumulant : $\kappa_2 = \mu_2 - \mu_1^2 = \mathrm{Var}(X)$ (variance)
- Troisième cumulant : $\kappa_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3$

Ces relations montrent que les cumulant sont des combinaisons de moments centrés, offrant une description alternative des propriétés de la distribution.

Q5. Pour démontrer cela, nous allons utiliser la fonction characteristique de la distribution gaussienne et la définition des cumulants. Pour $X \sim \mathcal{N}(\mu, \sigma^2)$ on a que

$$\begin{aligned}
\varphi_X(t) = \mathbb{E}[e^{itX}] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2} + itx\right) \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}(x^2 - 2\mu x + \mu^2) + itx\right) \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}x^2 + \left(\frac{\mu}{\sigma^2} + it\right)x - \frac{\mu^2}{2\sigma^2}\right) \mathrm{d}x \\
&= \frac{e^{-\frac{\mu^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left[x^2 - 2\sigma^2\left(\frac{\mu}{\sigma^2} + it\right)x\right]\right) \mathrm{d}x \\
&= \frac{e^{-\frac{\mu^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left[x^2 - 2\sigma^2\left(\frac{\mu}{\sigma^2} + it\right)x + \sigma^4\left(\frac{\mu}{\sigma^2} + it\right)^2 - \sigma^4\left(\frac{\mu}{\sigma^2} + it\right)^2\right]\right) \mathrm{d}x \\
&= \frac{e^{-\frac{\mu^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} e^{\frac{\sigma^2}{2}\left(\frac{\mu}{\sigma^2} + it\right)^2} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left[x - \sigma^2\left(\frac{\mu}{\sigma^2} + it\right)\right]^2\right) \mathrm{d}x \\
&= e^{-\frac{\mu^2}{2\sigma^2}} e^{\frac{\sigma^2}{2}\left(\frac{\mu^2}{\sigma^4} + \frac{2\mu it}{\sigma^2} - t^2\right)} = e^{-\frac{\mu^2}{2\sigma^2} + \frac{\mu^2}{2\sigma^2} + \mu it - \frac{\sigma^2 t^2}{2}} = e^{\mu it - \frac{\sigma^2 t^2}{2}}
\end{aligned} \tag{52}$$

On peut bien prendre le logaritme et avoir

$$H_X(t) = \log \varphi_X(t) = \log\left(e^{i\mu t - \frac{1}{2}\sigma^2 t^2}\right) = i\mu t - \frac{1}{2}\sigma^2 t^2 \tag{53}$$

En comparant cette série avec l'expression de de la question precedent, nous voyons que

- $\kappa_1 = \mu$ (le premier cumulant est la moyenne)
- $\kappa_2 = \sigma^2$ (le deuxième cumulant est la variance)
- $\kappa_n = 0 \ \forall n > 2$ (il n'y a pas de termes d'ordre supérieur à 2 dans l'expression de $H_X(t)$)

Ainsi, nous avons démontré que pour une distribution gaussienne, tous les cumulants d'ordre supérieur à 2 sont nuls.

Q6. Soit $\varphi_X(t)$ la fonction génératrice de X. Pour S_N, nous avons :

$$\varphi_{S_N}(t) = \mathbb{E}[e^{itS_N}] = \mathbb{E}[e^{it(X_1+\cdots+X_N)}] = \mathbb{E}[e^{itX_1}\cdots e^{itX_N}] \tag{54}$$
$$= \mathbb{E}[e^{itX_1}]\ldots\mathbb{E}[e^{itX_N}] \quad \text{(par indépendance)} \tag{55}$$
$$= \varphi_X(t)^N \tag{56}$$

Donc, la fonction génératrice de S_N est la N-ième puissance de la fonction génératrice de X. Soit $\kappa_{n,X}$ le n-ième cumulant de X et κ_{n,S_N} le n-ième cumulant de S_N, donc avec un calcul direct

$$\log \varphi_{S_N}(t) = \log\big(\varphi_X(t)^N\big) = N \log \varphi_X(t) = N \sum_{n=1}^{\infty} \kappa_{n,X} \frac{(it)^n}{n!} \tag{57}$$

En comparant cela avec la définition des cumulants pour S_N

$$\log \varphi_{S_N}(t) = \sum_{n=1}^{\infty} \kappa_{n,S_N} \frac{(it)^n}{n!} \tag{58}$$

Nous concluons que

$$\kappa_{n,S_N} = N\kappa_{n,X} \quad \forall n, N \in \{1, 2, 3\ldots\} \tag{59}$$

Cette relation montre que les cumulants de la somme de variables aléatoires indépendantes et identiquement distribuées sont simplement N fois les cumulants de la variable individuelle.

Q7. Commençons par exprimer la fonction génératrice de G_N en termes de celle de X. Comme l'exercise precedent, nous avons:

$$\varphi_{G_N}(t) = \mathbb{E}[e^{i\frac{t}{\sqrt{\kappa_2}}\left(\frac{\sum_i X_i}{\sqrt{N}} - \sqrt{N}m_1\right)}] = \mathbb{E}[e^{i\frac{t}{\sqrt{N}}\left(\frac{X_1-m_1}{\sqrt{\kappa_2}}\right)}]\ldots\mathbb{E}[e^{i\frac{t}{\sqrt{N}}\left(\frac{X_N-m_1}{\sqrt{\kappa_2}}\right)}] \tag{60}$$
$$= \left[\varphi_{\frac{X-m_1}{\sqrt{\kappa_2}}}\left(\frac{t}{\sqrt{N}}\right)\right]^N \tag{61}$$

L'étape clé consiste à examiner le comportement de cette expression pour N grand. Si on reprend l'expression calculé dans la **Q3** de cette exercise, on a que

$$H_{G_N}(t) = NH_{\frac{X-m_1}{\sqrt{\kappa_2}}}\left(\frac{t}{\sqrt{N}}\right) = N\left[0 + \frac{1}{2}\left(i\frac{t}{\sqrt{N}}\right)^2 + o(N^{-1})\right] = -\frac{t^2}{2} + o_N(1)\,. \tag{62}$$

Cela signifie que, pour N grand, on a $H_{G_N} = -\frac{t^2}{2}$; c'est exactement l'expression que l'on a trouvée pour une gaussienne de moyenne nulle et variance un.

Q8. Non, il n'est pas nécessaire que tous les X_i aient la même distribution pour que le théorème central limite s'applique. Soit $\{X_1, X_2, ..., X_N\}$ une suite de variables aléatoires indépendantes, pas nécessairement identiquement distribuées. Supposons que chaque X_i a une espérance μ_i et une variance σ_i^2 finies et aussi le troisième moment qui est fini $\mathbb{E}[X_i^3] < \infty$ pour toutes $i = 1, \ldots, N$.
Pour généraliser l'approximation gaussienne, considérons la somme normalisée suivante

$$G_N = \frac{1}{\sqrt{\kappa_2}}\left(\frac{\sum_{i=1}^N X_i}{\sqrt{N}} - \frac{\sum_{i=1}^N \mu_i}{\sqrt{N}}\right), \quad \text{ou } \kappa_2 = \frac{1}{N}\sum_i \sigma_i^2 \tag{63}$$

On peut continuée comme la question precedent:

$$\varphi_{G_N}(t) = \mathbb{E}[e^{i\frac{t}{\sqrt{\kappa_2 N}}(\sum_i X_i - \sum_i \mu_i)}] = \prod_i \mathbb{E}[e^{i\frac{t}{\sqrt{N}}\left(\frac{X_i-\mu_i}{\sqrt{\kappa_2}}\right)}] = \prod_i \left[\varphi_{\frac{X_i-\mu_i}{\sqrt{\kappa_2}}}\left(\frac{t}{\sqrt{N}}\right)\right] \tag{64}$$

On peut d'abord utiliser le théorème de Taylor avec le reste de Peano pour écrire

$$\varphi_{\frac{X_i-\mu_i}{\sqrt{\kappa_2}}}(t/\sqrt{N}) = 1 + i\frac{t}{\sqrt{N}} m_{1,\frac{X_i-\mu_i}{\sqrt{\kappa_2}}} - \frac{t^2}{2N} m_{2,\frac{X_i-\mu_i}{\sqrt{\kappa_2}}} + \frac{t^2}{N} m_{2,\frac{X_i-\mu_i}{\sqrt{\kappa_2}}} h(t/\sqrt{N}) \tag{65}$$

ou $h(t/\sqrt{N}) \to 0$ pour $N \to \infty$. Si on prend le log, on a que

$$H_{G_N}(t) = \sum_i H_{\frac{X_i-\mu_i}{\kappa_2}}\left(\frac{t}{\sqrt{N}}\right) = \sum_i \left[0 - \frac{\sigma_i^2}{2\kappa_2}\frac{t^2}{N} + \frac{\sigma_i^2}{\kappa_2}\frac{t^2}{N} h(t/\sqrt{N})\right] \xrightarrow[N\to\infty]{} -\frac{t^2}{2} . \tag{66}$$

Dans la dernière étape, nous avons utilisé la définition de κ_2 et le fait que $h(t/\sqrt{N}) \to 0$ pour $N \to \infty$. Encore une fois, ça c'est exactement l'expression que l'on a trouvée pour une gaussienne de moyenne nulle et variance un.

Q9. De la définition de fonction caractéristique on a que

$$\varphi_X(t) = \int_0^\infty e^{itx} \lambda e^{-\lambda x}\, \mathrm{d}x = \lambda \int_0^\infty e^{-(\lambda - it)x}\, \mathrm{d}x = \lambda \left[-\frac{1}{\lambda - it} e^{-(\lambda-it)x}\right]_0^\infty = \frac{\lambda}{\lambda - it} \tag{67}$$

La fonction caractéristique existe pour tout $t \in \mathbb{R}$. On a aussi que avec un calcul

$$M_X(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x}\, \mathrm{d}x = \frac{\lambda}{\lambda - t} \tag{68}$$

la fonction génératrice des moments n'existe que pour $t < \lambda$.

Q10. Voir le Notebook

Exercise 3 Correction au CLT

On note par

$$\gamma_{\mu,\sigma^2}(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{69}$$

la densité d'une variable aléatoire Gaussienne de moyenne μ et variance σ^2.

Q1. a) Reprenez le calcul de **Q7** de l'**Exercise 2**, et montrez que pour N grand, la fonction H_{G_N} peut s'écrire comme

$$H_{G_N}(t) = -\frac{t^2}{2} - \frac{i\kappa_3 t^3}{6\sigma^3\sqrt{N}} + O(N^{-1}) . \tag{70}$$

b) En utilisant la relation entre la fonction caractéristique et la densité de probabilité, montrez que

$$p_{G_N}(x) = \gamma_{0,1}(x) \left[1 + \frac{\kappa_{3,X}}{6\sigma^3}\left(x^3 - 3x\right)\frac{1}{\sqrt{N}} + O(N^{-1})\right] \tag{71}$$

c) En déduire que $S_N = \sum_{i=1}^N X_i$ peut être approximée par une loi Gaussienne de largeur $\sigma\sqrt{N}$ dans un intervalle de largeur d'ordre N^α et donner la valeur maximale possible du paramètre α.

Q2. On note $F_{G_N}(x) := \int_{-\infty}^{x} p_{G_N}(y)\mathrm{d}y$ la cumulative de la variable G_N et $\Phi(x) := \int_{-\infty}^{x} \gamma_{0,1}(y)\mathrm{d}y$ celle d'une standard Gaussienne. On rappelle que la norme sup d'une fonction est donnée par $\|f\|_\infty := \sup_t |f(t)|$. Montrez que la différence entre les 2 cumulatives est donnée par

$$\|F_{G_N} - \Phi\|_\infty \le \frac{C}{\sqrt{N}} . \tag{72}$$

pour une constante C indépendante de N, c'est à dire que la vitesse de convergence est d'ordre $O(N^{-1/2})$.

* **Q3.** Écrivez un code python pour calculer la différence entre F_{G_N} et Φ puis tracez cette distance empirique en fonction de N en échelle logarithme, pour des valeurs de $N \in \{4, 16, 64, 256, 1024, 4096\}$. Montrez que le scaling est bien en $N^{-1/2}$ sur cette figure.

Solution of Exercise 3

Q1. a) Pour le troisième moment de G_N on peut voir que

$$m_{3,G_N} = \int p_{G_N}(g) g^3 dg = \int p(x) \frac{(x-\mu)^3}{\sigma^3} dx = \frac{\kappa_3}{\sigma^3} \tag{73}$$

où nous avons utilisé la définition du troisième cumulant.
Considérons la fonction caractéristique de G_N que comme dans l'autre reponse et développons $\log(1+x)$ autour de $x = 0$

$$\varphi_{G_N}(t) = \prod_i \varphi_{\frac{X_i - \mu}{\sqrt{\kappa_2}}}\left(\frac{t}{\sqrt{N}}\right) = \exp N \log\left(\varphi_{\frac{X-\mu}{\sqrt{\kappa_2}}}\left(\frac{t}{\sqrt{N}}\right)\right) \tag{74}$$

$$= \exp N \log\left(1 + 0 - \frac{m_2}{2\kappa_2}\left(\frac{t}{\sqrt{N}}\right)^2 - \frac{i}{6}\frac{\kappa_3}{\kappa_2^{3/2}}\left(\frac{t}{\sqrt{N}}\right)^3 + O(1/N^2)\right) \tag{75}$$

$$= \exp N\left(-\frac{t^2}{2N} - \frac{i}{6}\frac{\kappa_3}{\sigma^3}\frac{t^3}{N\sqrt{N}} + O(1/N^2)\right) = \exp\left(-\frac{t^2}{2} - \frac{i\kappa_3 t^3}{6\sigma^3\sqrt{N}} + O(1/N)\right) . \tag{76}$$

b) Définissons maintenant $k' = \frac{\kappa_3}{6\sigma^3\sqrt{N}}$ et utilisons la relation entre la fonction caractéristique et la fdp pour écrire:

$$p_{G_N}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-t^2/2 - itx} e^{-it^3 k'} dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-t^2/2-itx}(1 - it^3 k' + O(1/N))dt \tag{77}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-t^2/2-itx} dt - \frac{ik'}{2\pi}\int_{-\infty}^{\infty} t^3 e^{-t^2/2-itx} dt + O(1/N) \tag{78}$$

Pour le premier terme, nous pouvons remarquer qu'il s'agit de la transformée de Fourier d'une gaussienne standard. Alternativement, nous pouvons compléter les carrés au niveau de l'exposant et écrire

$$p_{G_N}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(t-ix)^2 - \frac{x^2}{2}} dt + \frac{ik'}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(t-ix)^2 - \frac{x^2}{2}} t^3 dt + O(1/N) \tag{79}$$

$$\approx \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\left(1 - \frac{ik'}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(t-ix)^2} t^3 dt\right) \tag{80}$$

En effectuant le changement de variables $t' = t + ix$, on peut écrire

$$p_{G_N}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(1 - \frac{ik'}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t'^2}{2}} \left(t'^3 - 3ixt'^2 + 3x^2 t' + ix^3\right) dt'\right) \tag{81}$$

$$= \gamma_{0,1}(x) \left(1 - ik'(-3ix + ix^3)\right) = \gamma_{0,1}(x) \left(1 + k'(x^3 - 3x)\right) . \tag{82}$$

Si nous substituons l'expression de k', nous retrouvons le résultat souhaité.

c) L'approximation a une signification si $\frac{x^3}{\sqrt{N}} \ll 1$. Ça veut dire $x \ll N^{1/6}$.

Q2. En intégrant Eq.(71) par rapport à x, on obtient

$$F_{G_N}(x) := \Phi(x) + \frac{f(x)}{\sqrt{N}} + O(N^{-1}) \tag{83}$$

avec $f(x) := \frac{\kappa_{3,X}}{6\sigma^3} \int^x He_3(y)\gamma_{0,1}(y)\mathrm{d}y$ et donc $||F_{G_N}(x) - \Phi(x)||_\infty = O(\|f(x)\|_\infty/\sqrt{N})$ et on obtient le résultat désiré avec $C = \|f(x)\|_\infty$.

Q3. Voir le Notebook

Exercise 4 Théorème de Cramer

Pour la variable aléatoire S_N, nous venons de voir qu'il existe un intervalle de taille N^α (avec un certain $\alpha \in (0,1)$ calculée en **Q1.b** de l'**Exercise 3**) dans laquelle l'approximation Gaussienne s'applique pour N très grand. La taille de cet intervalle est donc petite par rapport à la valeur typique (espérance) de S_N qui est d'ordre N. On s'intéresse ici à comprendre le probabilité que S_N prennent des valeurs dans un intervalle de taille $N \gg N^\alpha$. De manière équivalente, on cherche à comprendre la probabilité que $\hat{\mu}_N$ dévie de $O(1)$ de sa valeur typique. On rappelle que $H_X(t) := \log \varphi_X(t)$.

Q1. Écrivez la densité de probabilité de $\hat{\mu}_N$ en fonction de sa fonction caractéristique puis celles des X_i. En déduire en utilisant la méthode du col [3], montrez que $p_{\hat{\mu}_N}(u) \asymp \exp[-N I_X(u)]$, où la "fonction des grandes déviations" $I_X(u)$ est donnée par

$$I_X(u) = iz^\star u - H_X(z^\star) , \tag{85}$$

$z^\star$ étant la solution de $H'_X(z) = iu$.

Q2. On définit la fonction génératrice des cumulants comme

$$K_X(t) = \log\left(\mathbb{E}\left[e^{tX}\right]\right) \tag{86}$$

et l'on pose maintenant $z^* = -it^*$. Justifier pourquoi, si tout les moments existent, on trouve bien que $I_X(u) = ut^* - \kappa_X(t^*)$ (avec $\kappa_X(t) = \log M_X(t)$) avec t^* tel que $\partial_t(ut - \log M_X(t))|_{t^*} = 0$, ce qui est bien le théorème de Cramer.

Q3. Supposons que la distribution de X soit une gaussienne de largeur 1, *i.e.* $p_X(x) = \gamma_{0,1}(x)$. Calculer $H_X(t)$ et, en utilisant l'equation (85), calculer $I_X(u)$. Est-ce que dans ce cas on peut calculer $I_X(u)$ d'une facon plus simple?

Q4. On veut déterminer la loi de la moyenne empirique $\hat{\mu}_N$, où les X_i sont des variables aléatoires indépendantes et identiquement distribuées selon une loi exponentielle de paramètre λ.

[3] la méthode du col est l'analogue complexe de la méthode de Laplace. Si on considere un integrale sur un chemin C

$$\int_C g(z) e^{\lambda f(z)} \mathrm{d}z \approx g(z^*) \, e^{\lambda f(z^*)} \tag{84}$$

ou l'on a $f'(z^*) = 0$.

Distribution, Symbol	Forme densité $p_X(x)$	Fonction Caractéristique
Bernoulli, $\mathcal{B}(p)$	$p\delta(x-1)+(1-p)\delta(x+1)$	$1-p+pe^{it}$
Geometric, $\mathrm{Geom}(p)$	$\sum_{k=0}^{\infty}(1-p)^{k-1}p\delta(x-k)$	$\frac{pe^{it}}{1-(1-p)e^{it}}$
Uniform continuous, $\mathrm{Unif}([a,b])$	$\frac{1}{b-a}\mathbf{1}[x\in[a,b]]$	$\frac{e^{itb}-e^{ita}}{it(b-a)}$
Laplace, $\mathrm{Laplace}(a,b)$	$\frac{1}{2b}\exp\left(-\frac{\lvert x-\mu\rvert}{b}\right)$	$\frac{e^{it\mu}}{1+b^2t^2}$
Gaussian, $\mathcal{N}(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	$e^{it\mu-\frac{1}{2}\sigma^2t^2}$
Gamma, $\Gamma(\alpha,\beta)$	$\frac{x^{\alpha-1}e^{-\beta x}\beta^\alpha}{\Gamma(\alpha)}\ \forall\alpha,\beta>0$	$\left(1-\frac{it}{\beta}\right)^{-\alpha}$
Chi squared, $\chi^2(k)$	$\frac{x^{k/2-1}e^{-x/2}}{2^{k/2}\Gamma(k/2)}\ k\in\{1,2,3,\dots\}$	$(1-2it)^{-\frac{k}{2}}$
Binomial, $B(n,p)$	$\sum_{k=0}^{N}\binom{n}{k}p^k(1-p)^{n-k}\delta(x-k)$	$\left(1-p+pe^{it}\right)^N$
Cauchy, $\mathrm{Cauchy}(x_0,\gamma)$	$\frac{1}{\pi}\left[\frac{\gamma}{(x-x_0)^2+\gamma^2}\right]$	$e^{it\mu-\theta\lvert t\rvert}$

a) Trouvez la fonction caractéristique du $\varphi_{\hat{\mu}_N}(t)$ du somme de variables aléatoires indépendantes. Quelle loi est $\hat{\mu}_N$?

b) Calculez la moyenne et la variance de la loi trouvée.

Q5. En utilisant le théorème de Cramér, calculez la fonction de grandes déviations $I_X(x)$ pour la moyenne empirique de N variables aléatoires indépendentes suivant la distribution calculée en **Q1** de l'**Exercise 1**.

* **Q6.** *(Simulations numériques)*

a) Comparez graphiquement la fonction de grandes déviations obtenue à la question précédente avec les résultats de simulations pour des valeurs de $N\in\{4,32,256\}$ (nombre d'échantillons). Discutez de la convergence observée.

b) Commentez les différences entre ces approximations et la fonction de grandes déviations, en particulier pour les événements rares.

Solution of Exercise 4

Q1. Commençons par exprimer la probabilité $P_{\hat{\mu}_N}(u)$ en termes de la fonction caractéristique

$$P_{\hat{\mu}_N}(u)=P(S_N/N=u)=\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{2\pi}e^{-itu}\mathbb{E}[e^{itS_N/N}]=\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{2\pi}e^{-itu}(\varphi_X(t/N))^N \tag{87}$$

$$=\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{2\pi}\exp[-itu+NH_X(t/N)] \tag{88}$$

ou on a utilisez la définition $H_X(t)=\log\varphi_X(t)$.
Pour N grand, on peut appliquer la méthode du col. On va definir $z=t/N$ e on cherche le point $z^\star$ qui maximise l'exposant

$$\frac{\mathrm{d}}{\mathrm{d}z}[-izu+H_X(z)]=-iu+H_X'(z)=0 \tag{89}$$

Ce qui donne la condition $H_X'(z^\star)=iu$.
En développant l'exposant autour de $z^\star$ et en gardant les termes dominants pour N grand, on obtient

$$P_{\hat{\mu}_N}(u)\asymp C_Ne^{-N(iz^\star u-H_X(z^\star))} \tag{90}$$

où C_N est un facteur de normalisation. En identifiant avec l'expression donnée, on trouve bien

$$I_X(u)=iz^\star u-H_X(z^\star) \tag{91}$$

Ce qui démontre le résultat demandé.

Q2. On a le relation

$$H'_X(z) = \frac{\phi'(z)}{\phi(z)}\,, \tag{92}$$
$$H'_X(z) = \mathrm{i} \cdot \frac{M'(\mathrm{i}z)}{M(\mathrm{i}z)} \tag{93}$$

et donc en posant $z = -it$, l'équation du point col se réécrit:

$$\mathrm{i} \cdot \frac{M'(t^*)}{M(t^*)} = \mathrm{i}u \tag{94}$$

c'est à dire

$$K'_X(t^*) = u\,, \tag{95}$$

on peut donc réécrire

$$I_X(u) = t^*u - H_X(-it^*)\,, \tag{96}$$
$$I_X(u) = t^*u - K_X(t^*)\,, \tag{97}$$

ce qui est le résultat souhaité.

Q3. Pour une distribution gaussienne standard $X \sim \mathcal{N}(0,1)$, la fonction caractéristique et son log sont

$$\varphi_X(t) = e^{-t^2/2}\,, \quad H_X(t) = \log \varphi_X(t) = -\frac{t^2}{2}\,. \tag{98}$$

Pour calculer $f_X(u)$, on doit d'abord trouver $z^\star$ tel que $H'_X(z^\star) = iu$ et il est $z^\star = -iu$. En substituant dans l'équation pour $I_X(x)$

$$I_X(u) = i(-iu)u - H_X(-iu) = u^2 + \frac{(-iu)^2}{2} = \frac{u^2}{2} \tag{99}$$

On peut en effet calculer $I_X(u)$ d'une façon plus simple dans ce cas. Comme S_N/N est la moyenne de N variables aléatoires gaussiennes indépendantes de moyenne 0 et variance 1, sa distribution est une gaussienne de moyenne 0 et variance $1/N$. Donc

$$P_{u_N}(u) = \frac{\sqrt{N}}{\sqrt{2\pi}} e^{-N\frac{u^2}{2}} \tag{100}$$

En identifiant avec la forme $e^{-NI_X(u)}$, on retrouve bien

$$I_X(u) = \frac{u^2}{2} \tag{101}$$

Q4. Sous l'hypothèse que tous les cumulants de X existent, on peut exprimer $H_X(z)$ comme une série de Taylor

$$H_X(z) = \sum_{n=1}^{\infty} \kappa_{n,X} \frac{(iz)^n}{n!} \tag{102}$$

Notons que tous les $\kappa_{n,X}$ sont réels. La condition $H'_X(z^*) = iu$ devient donc

$$\sum_{n=1}^{\infty} \frac{\kappa_{n,X} i^n}{(n-1)!} (z^\star)^{n-1} = \sum_{n=1}^{\infty} \frac{\kappa_{n,X}}{(n-1)!} i^n w(\mu^2 + \zeta^2)^{\frac{n-1}{2}} e^{i(n-1)\theta} = iu \tag{103}$$

ou on pose $z^\star = \mu + i\zeta$ et $\theta = \arctan(\zeta/\mu)$. On peut d'abord diviser la partie reelee

$$\begin{aligned}0 &= \sum_{n=1}^{\infty} \frac{\kappa_{2n,X}}{(2n-1)!}(-1)^n(\mu^2+\zeta^2)^{\frac{2n-1}{2}} \cos\left((2n-1)\theta\right) \quad \text{(paire)} \\ &+ \sum_{n=1}^{\infty} \frac{\kappa_{2n-1,X}}{(2n-2)!}(-1)^{n-1}(\mu^2+\zeta^2)^{\frac{2n-2}{2}} \sin\left((2n-2)\theta\right) \quad \text{(impaire)}\end{aligned}$$

et la partie imaginaire

$$\begin{aligned}u &= \sum_{n=1}^{\infty} \frac{\kappa_{2n,X}}{(2n-1)!}(-1)^n(\mu^2+\zeta^2)^{\frac{2n-1}{2}} \sin\left((2n-1)\theta\right) \quad \text{(paire)} \\ &+ \sum_{n=1}^{\infty} \frac{\kappa_{2n-1,X}}{(2n-2)!}(-1)^{n+1}(\mu^2+\zeta^2)^{\frac{2n-2}{2}} \cos\left((2n-2)\theta\right) \quad \text{(impaire)} \qquad (104)\end{aligned}$$

La première équation est satisfaite si $\mu = 0$, ce qui implique $\theta = \pi/2$. La deuxième équation détermine alors ζ en fonction de u. Ainsi, on a bien montré que $z^\star = \mu + i\zeta$ avec ζ réel.
La convexité de $I_X(u)$ découle directement de cette relation de Legendre comme on avait vu dans la premiere serie. En effet, la transformée de Legendre d'une fonction convexe est toujours convexe.

Q5. a) Nous allons procéder par étapes pour déterminer la loi de la moyenne empirique $\hat{\mu}_N$. Dans l'exercice precedent **Q8** on a trouve la fonction characteristique du varaible exponetiel avec parameter λ. On peut suivre le raisonment dans la question **Q6** du meme exo pour dire que

$$\begin{aligned}\varphi_{\hat{\mu}_N}(t) &= \mathbb{E}[e^{it\hat{\mu}_N}] = \mathbb{E}[e^{it(\frac{1}{N}\sum_{i=1}^N X_i)}] = \mathbb{E}[e^{\frac{it}{N}\sum_{i=1}^N X_i}] \quad \text{(independance)} \qquad (105) \\ &= \prod_{i=1}^N \mathbb{E}[e^{\frac{it}{N}X_i}] = \prod_{i=1}^N \varphi_X\left(\frac{t}{N}\right) = \left(\frac{\lambda}{\lambda - \frac{t}{N}}\right)^N \qquad (106) \\ &= \left(\frac{N\lambda}{N\lambda - t}\right)^N = \left(1 - \frac{t}{N\lambda}\right)^{-N} \qquad (107)\end{aligned}$$

Cette forme correspond à la fonction génératrice des moments d'une distribution Gamma de paramètres $\alpha = N$ et $\beta = \frac{1}{N\lambda}$. Donc, la moyenne empirique $\hat{\mu}_N$ suit une loi Gamma de paramètres $\alpha = N$ et $\beta = \frac{1}{N\lambda}$. La loi Gamma a un densité de probabilité

$$X \sim \Gamma(\alpha, \beta)\,, \quad p_X(x) = \frac{x^{\alpha-1}e^{-\beta x}\beta^\alpha}{\Gamma(\alpha)} \quad \text{ou } \Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\,\mathrm{d}t \qquad (108)$$

On note que il ne peut pas être un loi chi squared parce que nous avons deux paramètres.

b) On veut démontrer que l'espérance et la variance d'une variable aléatoire $X \sim \Gamma(\alpha, \beta)$

$$\mathbb{E}[X] = \frac{\alpha}{\beta}\,, \quad \mathrm{Var}(X) = \frac{\alpha}{\beta^2} \qquad (109)$$

On pourrait faire les integrales exact mais il est plus facile de developer en taylor la forme de la fonction characteristique. On a que

$$H_X(t) = -\alpha \log\left(1 - \frac{it}{\beta}\right) = \frac{i\alpha t}{\beta} - \frac{\alpha t^2}{2\beta^2} + \mathcal{O}(t^2)\,. \qquad (110)$$

La distribution de $\hat{\mu}_N$ a donc $\mathbb{E}[\hat{\mu}_N] = \frac{1}{\lambda}$ et $\mathrm{Var}(\hat{\mu}_N) = \frac{1}{N\lambda^2}$.
Remarque : Lorsque N tend vers l'infini, cette distribution converge vers une distribution normale centrée sur $\frac{1}{\lambda}$, conformément au théorème central limite.

Q6. On a deja trouve dans **EX2** la fonction generatrice de moments d'une variable exponentielle de paramètre λ. Selon le théorème de Cramér, la fonction de grandes déviations $I(x)$ est la transformée de Legendre-Fenchel de la fonction generatrice de cumulants $H_X(t)$. Donc on a que

$$I(x) = \sup_{t\in\mathbb{R}}\{tx - H_X(t)\} = \sup_{t\in\mathbb{R}}\{tx + \log\left(1 - \frac{t}{\lambda}\right)\} \tag{111}$$

Pour trouver le supremum, nous dérivons par rapport à t et égalisons à zéro

$$\frac{\mathrm{d}}{\mathrm{d}t}(tx - H_X(t)) = x - \frac{1}{\lambda - t} = 0 \tag{112}$$

et donc

$$x = \frac{1}{\lambda - t} \quad \Rightarrow \quad t = \lambda - \frac{1}{x} \tag{113}$$

Donc, la fonction de grandes déviations pour la moyenne empirique de n variables exponentielles de paramètre λ est

$$I(x) = \lambda x - \log(\lambda x) - 1, \quad \text{pour } x > 0 \tag{114}$$

Q7. Voir le notebook