

Théorèmes limites et tests

Dans ce TP, on utilisera la bibliothèque `scipy.stats` avec l'alias `ss`. On rappelle que les commandes suivantes permettent d'obtenir la densité, la fonction de répartition et sa bijection réciproque de la loi normale centrée réduite :

```
ss.norm.pdf(x,0,1) # probability distribution function
ss.norm.cdf(x,0,1) # cumulative distribution function
ss.norm.ppf(x,0,1) # percent point function
```

Exercice 1. Dans cet exercice, on pourra utiliser `rd.randint`.

1. Créer une fonction `Échantillon(n)` qui renvoie une liste de n nombres dans $\llbracket 1;6 \rrbracket$ tirés de façon uniforme.
2. Pour un tel 100-uplet, afficher le graphe du nuage de points (k, M_k) où M_k désigne la moyenne empirique des k premiers nombres. Quel résultat avez-vous illustré ?
3. Écrire une fonction `SimulerMnstar(n,taille)` qui renvoie un échantillon de taille `taille` de M_n^* .
4. Afficher l'histogramme de cet échantillon. Quel résultat avez-vous illustré ?
5. Rajouter sur le graphique la densité de la loi limite.

Il est possible d'afficher l'histogramme cumulé (ce qui revient à la fonction de répartition empirique), pour cela il suffit de rajouter l'option `cumulative=True` dans la commande `plt.hist`

6. Afficher l'histogramme cumulé d'un échantillon de M_n^* ainsi que la fonction de répartition de la loi limite. Est-ce cohérent ?

Exercice 2. 1. Écrire une fonction `FoncRepartitionBinomial(n,p,x)` qui renvoie $F(x)$ où F est la fonction de répartition de $X \sim \mathcal{B}(n,p)$.

2. Toujours si $X \sim \mathcal{B}(n,p)$, afficher la fonction de répartition de X^* .
3. Quel résultat du cours de mathématiques avez-vous illustré ?
4. Afficher la fonction de répartition de la loi limite.

Exercice 3. 1. Afficher la loi de $X \sim \mathcal{B}(n,p)$ avec p petit et n grand ainsi que celle de $Y \sim \mathcal{P}(np)$.

2. Quel résultat du cours de mathématiques avez-vous illustré ?

Exercice 4. On considère la liste suivante : $[7, 7, 1, 7, 5, 2, 3, 8, 8, 6, 3, 6, 6, 3, 3, 1, 6, 4, 2, 7, 7, 6, 3, 8, 5, 7, 2, 7, 6, 8, 5, 6, 3, 4, 7, 2, 4, 1, 3, 8, 6]$.

1. Sans utiliser les commandes implémentées dans Python, écrire des fonctions qui calculent, la moyenne empirique, l'écart-type empirique et l'écart-type empirique corrigé de cette liste.
2. En utilisant les fonctions dans Python sur la fonction de répartition (ou de sa bijection réciproque) de la loi normale centrée réduite. Déterminer α tel que $\mathbb{P}(-\alpha \leq Z \leq \alpha) = 0.99$ pour $Z \sim \mathcal{N}(0,1)$.
3. Ces données sont des réalisations d'un n -uplet de variables aléatoires indépendantes et de même loi que X . En utilisant le résultat de la question précédente, l'hypothèse $\mathbb{E}(X) = 4$ peut-elle être rejetée ?

Exercice 5. Soit $(X_1, X_2, \dots, X_n)$ une famille de VA indépendantes suivant une loi normale de paramètres μ (inconnue) et σ^2 (connue). On note M_n la moyenne empirique et M_n^* la variable aléatoire centrée réduite associée à M_n .

1. Quelle est la loi de M_n ? Et celle de M_n^* ?
2. On considère la liste de valeurs $[-1.6705680249696178, 7.232741859688569, 1.8216579211943107, 9.925103912950288, 4.365864941800229, 1.6425548593097299, 5.786656176385351, 5.887092676215338, 2.1266838865241127, 8.512226145970482]$, on suppose qu'il s'agit de réalisations de loi normales de paramètre μ et $\sigma^2 = 4$. Déterminer un intervalle tel que la probabilité que μ soit dans cet intervalle soit plus grand que 99%.
3. L'hypothèse d'avoir un nombre grand d'observations est-elle nécessaire ici ?
4. L'hypothèse $\mu = 4$ peut-elle être rejetée ?

Exercice 6. Soit $(X_1, X_2, \dots, X_n)$ une famille de VA indépendantes suivant toute une loi normale de paramètres μ (inconnue) et σ^2 (inconnue). On note M_n la moyenne empirique et M_n^* la variable aléatoire centrée réduite associée à M_n .

1. Quelle est la loi de M_n ?
2. Comme σ n'est pas connu, on le remplace par l'écart-type corrigé, ainsi, la variable aléatoire $\frac{M_n - \mu}{\frac{S'_n}{\sqrt{n}}}$ suit une loi dite de Student à $n-1$

degrés de liberté, avec M_n la moyenne empirique et S'_n l'écart-type empirique corrigé.

Les commandes suivantes permettent respectivement de calculer la densité, la fonction de répartition, et la bijection réciproque de la loi de Student à d degré de liberté :

```
ss.t.pdf(x,d) # probability distribution function
ss.t.cdf(x,d) # cumulative distribution function
ss.t.ppf(x,d) # percent point function
```

3. Tracé la densité de la loi de Student à k degrés de liberté ainsi que la densité la loi normale centrée réduite, puis faire augmenter le nombre de degrés de liberté, que peut-on en conclure ?
4. Déterminer $\alpha \in \mathbb{R}$ tel que $\mathbb{P}(-\alpha \leq T \leq \alpha) = 0.99$ si T suit une loi de Student à 9 degrés de liberté, sachant que la densité d'une loi de Student est paire (comme la loi normale), à l'aide de Python.
5. On considère la liste de valeurs $[-7.743313632342028, -2.3404945016285206, 9.191871085318914, 11.98324209371975, 5.192986514516299, 4.610009953443223, 4.07441486300174, 4.023946630215239, 2.4809648235475104, 4.312927240240231]$, on suppose qu'il s'agit de réalisations de loi normales de paramètres μ et σ^2 . Déterminer un intervalle tel que la probabilité que μ soit dans cet intervalle soit plus grand que 99%.
6. L'hypothèse d'avoir un nombre grand d'observations est-elle nécessaire ici ?
7. L'hypothèse $\mu = 4$ peut-elle être rejetée ?

Remarque 1. Le test dans le dernier exercice utilise la loi de Student et s'appelle donc le test de Student, il permet de s'affranchir de l'hypothèse que n est grand contrairement au théorème central limite.