TD 4 : tests élémentaires

Les questions marquées d'un astérisque **(*)** *sont facultatives.*

Dans tous les exercices, on détaillera avec autant de soin que possible toutes les étapes des tests mis en oeuvre.

Exercice 1. Soient $X_1, \dots, X_n$ des v.a. i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$ pour (μ, σ^2) inconnu. On admet que la matrice de variance-covariance de (X_1, X_1^2) est

$$\begin{pmatrix} \sigma^2 & 2\mu\sigma^2 \\ 2\mu\sigma^2 & 2\sigma^4 + 4\mu^2\sigma^2 \end{pmatrix}$$

1. Énoncer le théorème central limite pour le couple $(\frac{1}{n}\sum_{i=1}^n X_i, \frac{1}{n}\sum_{i=1}^n X_i^2)$.
2. En utilisant la delta-méthode, en déduire la normalité asymptotique de l'estimateur de la variance $\frac{n-1}{n}S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ où $\bar{X}_n = \frac{1}{n}\sum_{j=1}^n X_j$.
3. En déduire la normalité asymptotique de S_n^2. *Indication pour vérifier vos calculs : la variance de la limite dépend-elle de μ ?*
4. En déduire un intervalle de confiance de σ^2.
5. **(*)** Démontrer la formule de la matrice de variance-covariance.

Exercice 2. Vous venez d'être recruté dans un laboratoire pharmaceutique et on vous charge de concevoir un test pour détecter une maladie génétique rare. Cette maladie concerne une personne sur 10 000. On souhaite que le test sache reconnaître une personne malade avec 9 chances sur 10 et qu'au plus la moitié des personnes positives à ce test soient saines.

1. Formulez l'hypothèse nulle et l'hypothèse alternative de ce test.
2. Traduire l'énoncé : quelle est le niveau souhaité du test ? Sa puissance ? Son risque de première espèce ? Son risque de deuxième espèce ? On donnera la valeur exacte de chacune de ces quantités sous forme de fraction.

Concrètement, ce test consiste à estimer la concentration d'un marqueur biologique dans un échantillon. Si la personne est saine, cette concentration vaut exactement μ_0 ; si elle est malade, exactement $\mu_0 + \Delta$ avec $\Delta > 0$. On suppose que la mesure de cette concentration suit une loi normale d'écart-type σ. En investissant davantage pour augmenter la qualité du procédé de mesure, il est possible de réduire cet écart-type autant que vous le souhaitez.

3. En admettant que si Z suit la loi normale centrée réduite, alors pour tout $x > 0$,
$$\mathbb{P}(Z \geq x) \leq e^{-x^2/2},$$
donner un majorant du quantile d'ordre $1 - \alpha$ de la loi normale $\mathcal{N}(a, b^2)$ en fonction de $a \in \mathbb{R}$, $b > 0$ et $\alpha \in [0, 1/2[$.
4. Quelle valeur de σ devez-vous fixer pour satisfaire le cahier des charges ?

Exercice 3. Vous participez à une tombola. A chaque tentative, l'opérateur secoue une urne contenant un certain nombre de billes blanches et une bille noire jusqu'à ce qu'une bille sorte. Si la bille noire sort, vous avez gagné. Par contre, il est incapable de se souvenir du nombre de billes blanches, même s'il vous assure qu'il y en a 19 ou moins. Méfiant, vous décidez d'observer ceux qui tentent leur chance avant de décider s'il est intéressant de participer.

1. Au bout de 30 tentatives, vous n'avez toujours pas vu la bille noire. Avez-vous des raisons de mettre en doute l'affirmation de l'opérateur ?
2. Au bout de 300 tentatives, vous avez vu 9 fois la bille noire. Avez-vous des raisons de mettre en doute l'affirmation de l'opérateur ? Fournir la p-valeur du test effectué.

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X = k)$	0	0,0001	0,0005	0,0016	0,0042	0,0094	0,0181	0,0309	0,0473	0,0657	0,0832
$\mathbb{P}(X \leq k)$	0	0,0002	0,0007	0,0023	0,0066	0,016	0,0341	0,065	0,1123	0,178	0,2612

k	13	14	15	16	17	18	19	20	21	22
$\mathbb{P}(X = k)$	0,0971	0,1047	0,1051	0,0985	0,0866	0,0717	0,056	0,0414	0,0291	0,0194
$\mathbb{P}(X \leq k)$	0,3583	0,463	0,5681	0,6666	0,7533	0,825	0,881	0,9224	0,9514	0,9708

FIGURE 1 – Fonction de masse et de répartition d'une variable aléatoire $X \sim \mathrm{Bin}(300; 0,05)$.

Exercice 4. Des chercheurs ont développé une nouvelle espèce de sardines d'élevage avec pour objectif de maximiser la masse de viande obtenue par poisson. La masse produite par une sardine adulte de l'ancienne espèce était de 0,113 kg. Sur 25 sardines de la nouvelle espèce, on a mesuré une masse moyenne produite de 0,145 kg et un écart-type de 0,073 kg. Faut-il, au vu de ces résultats, favoriser l'élevage de la nouvelle espèce ?

Voici les valeurs approchées de quelques quantiles d'ordre 95%*, c'est-à-dire l'infimum des $x \in \mathbb{R}$ tels que $\mathbb{P}(X \leq x) \geq 0,95$ pour X suivant les lois suivantes. Précisez lequel vous avez utilisé et pourquoi :*

— Bin(25 ; 0,113) : 5

— Bin(25 ; 0,145) : 6

— $\mathcal{N}(0,1)$: 1,645

— $\mathcal{T}_{24}$: 1,711

— $\mathcal{T}_{25}$: 1,708

— $\mathcal{T}_{26}$: 1,706