

SÉRIE 12

Exercice 1. * Tests unilatéraux. Parfois, l'hypothèse nulle concerne un continuum de valeurs. Si T est la statistique de test et α est le niveau de signification du test, il faut alors trouver la valeur critique c qui vérifie

$$\max_{\theta \in H_0} \mathbb{P}_\theta(T > c) = \alpha,$$

où $\mathbb{P}_\theta$ désigne la probabilité associée à la valeur θ . Pour répondre à la question ci-dessous, il faut considérer un **test unilatéral** : $H_0 : \theta \leq \theta^*$ et $H_1 : \theta > \theta^*$. Nous sommes donc dans le cas venant d'être mentionné où H_0 concerne un continuum de valeurs.

Pour estimer le résultat d'un possible référendum sur l'indépendance du Québec, un sondage d'opinion a été effectué. Parmi les 800 personnes interrogées, 55 % se sont prononcés en faveur de l'indépendance du Québec. Ce sondage montre-t-il que la majorité de la population est favorable à l'indépendance ?

NB : les données sont inventées.

Indication : Utilisez le théorème central limite. Puis, pour un a donné, étudiez l'évolution de $\mathbb{P}_\theta(T > a)$, où T est une statistique appropriée et θ est le vrai pourcentage de personnes favorables à l'indépendance.

Exercice 2. Le tableau suivant donne le nombre de naissances dans un hôpital pour quatre trimestres consécutifs :

Trimestre	Janv-Mars	Avr-Juin	Juil-Sept	Oct-Déc	Total
Nombre de naissances	110	57	53	80	300

- Tester à un niveau de signification de 1 % si la natalité du premier trimestre est deux fois plus élevée que celle de chacun des autres trimestres, qui sont toutes égales.
- Tester à un niveau de signification de 1 % si la natalité du premier trimestre est égale à celle du quatrième trimestre, et celle du deuxième trimestre est égale à celle du troisième trimestre.

Indication : dans ce cas il va falloir estimer un paramètre.

Exercice 3. Les données ci-dessous représentent le taux d'oxygénation de cours d'eau ayant la même température dans une certaine région :

taux d'oxygénation (par jour)	fréquence
< 0.1	12
$(0.1, 0.15]$	20
$(0.15, 0.20]$	23
$(0.20, 0.25]$	15
> 0.25	13

La moyenne et l'écart-type calculées pour cet échantillon de taille 83 sont $\bar{x} = 0.173$ et $s_x = 0.066$, respectivement. Testez à un niveau de signification de 5 % la qualité de l'ajustement d'une loi normale aux données ci-dessus.

Exercice 4. Dans l'étude des défaillances d'un composant électronique, on a relevé les retours en fabrique de 200 composants dont les défaillances ont été classées par type (T1, T2) et par localisation (L1, L2, L3). Sur la base du tableau ci-dessous, peut-on conclure qu'il y a une dépendance entre le type et la localisation du défaut ? Utiliser un niveau de signification de $\alpha = 5\%$.

	L1	L2	L3	Total
T1	50	16	31	97
T2	61	26	16	103
Total	111	42	47	200

Exercice 5. Cet exercice porte sur la puissance de tests, un critère pour juger si un test est performant ou non.

Définition : La puissance β d'un test est la probabilité de rejeter H_0 quand elle est fausse. Donc $\beta = 1 - \Pr_\theta(\text{erreur de type II})$ pour θ qui correspond à H_1 .

Soient $Y_1, \dots, Y_n \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$ où μ est inconnu mais $\sigma^2 > 0$ est connu. Considérons $H_0 : \mu = \mu_0$ et $H_1 : \mu > \mu_0$ (paire unilatéral). Utiliser la statistique de test $T_n = \sqrt{n}(\bar{Y}_n - \mu_0)/\sigma$ avec niveau de significativité $\alpha \in (0, 1)$.

- (i). Quelle est la distribution de T si H_0 est vraie ?
- (ii). Pour quelles valeurs est-ce qu'on rejette H_0 ?
- (iii). Donner une formule pour la puissance de test, en fonction de α , la vraie valeur μ , et la taille de l'échantillon n . La puissance croît ou décroît-elle en fonction de α , de n et de μ et de σ . Interpréter. *Indication.* Considérez le pivot $W_n = \sqrt{n}(\bar{Y}_n - \mu)/\sigma$.