TD 7. Tests d'hypothèses

1 Heures de vol

Le temps de vol d'un certain type d'avion sur un trajet fixé possède une moyenne de 16,25 heures depuis la mise en service de ces avions. La distribution du temps de vol a un écart type de 1,5 heures. Des mesures récentes couvrant 120 vols donnent une moyenne de 15 heures 56 minutes.

1. Donner un intervalle de confiance pour le temps moyen de vol (risque 5%).
2. Y-a-t-il une différence significative avec la moyenne annoncée ?
3. Mêmes questions si l'on ne suppose plus que l'écart-type est connu, mais que 1,5 est sa valeur estimée sur les 120 vols.

2 Loi exponentielle

Soit $X_1, \dots, X_n$ un n-échantillon d'une loi de fonction de répartition F. On note $\mathcal{F}$ l'ensemble des fonctions de répartition $F_\theta,\ \theta > 0$ où $F_\theta(x) = (1 - \exp(-x/\theta))\,\mathbf{1}_{\mathbb{R}_+}(x)$. Soit $\alpha \in]0,1[$ et $\theta_o > 0$. On suppose que $F = F_\theta$.

1. Calculer l'estimateur du maximum de vraisemblance T de θ. Que dire de l'estimation de θ par la méthode des moments ?
2. Prouver que $\frac{2nT}{\theta}$ suit la loi $\chi^2(2n)$. En déduire un intervalle de confiance bilatère pour θ de risque α.
3. Bâtir un test de "$\theta = \theta_o$" contre "$\theta \neq \theta_o$".

3 Puissance

Quelle est l'effectif nécessaire n, pour qu'avec une erreur de première espèce de 5%, on détecte un différence de moyenne égale à 0.6 fois l'écart type, avec une erreur de deuxième espèce infèrieur à 1%. ?

4 Test de Student

Soient $a,\ b \in \mathbb{R}$ et $\sigma \in \mathbb{R}^+$. Soient $(X_j)_{j=1\dots n}$ et $(Y_j)_{j=1\dots n}$ $2n$ variables aléatoires indépendantes. On suppose que, pour $j = 1\dots n$, X_j (respectivement Y_j) suit la loi normale de moyenne a et d'écart type σ (respectivement de moyenne b et d'écart type σ).

1. Quelle est la vraisemblance associée aux observations $Z := (X_1, \dots, X_n, Y_1, \dots, Y_n)$? Calculer les estimateurs du maximum de vraisemblance $\hat{a},\ \hat{b},\ \hat{\sigma^2}$.
2. Montrer que $\hat{a}$ et $\hat{b}$ peuvent s'écrire sous la forme

$$\hat{a} = \frac{1}{n}\langle v_1, Z\rangle \qquad\qquad \hat{b} = \frac{1}{n}\langle v_2, Z\rangle$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire sur $\mathbb{R}^{2n}$ et $v_1,\ v_2$ sont des vecteurs de $\mathbb{R}^{2n}$ à préciser. Montrer que $\hat{\sigma^2}$ est un estimateur biaisé. Proposer un estimateur sans biais de σ^2. Dans la suite on appelle S^2 cet estimateur.

3. Vérifier que $S^2 = \frac{1}{2n-2}\|Z - \hat{a}v_1 - \hat{b}v_2\|^2$, où $\|\cdot\|$ désigne la norme euclidienne sur $\mathbb{R}^{2n}$. Quelle est la loi de :
$$T = \frac{\sqrt{n}(\hat{a} - \hat{b} - (a-b))}{\sqrt{2}S}.$$
On justifiera sa réponse
4. Construire un test d'hypothèse de H_0 $a = b$ contre H_1 $a \neq b$ basé sur T. Application numérique :
$$n = 100; \qquad \sum_{i=1}^{n} x_i^2 = 150,4 \quad \sum_{i=1}^{n} y_i^2 = 160,2$$
$$\sum_{i=1}^{n} x_i = 120,11 \quad \sum_{i=1}^{n} y_i = 110,33.$$

5 Fruitier

Un producteur de fruits a constaté sur des récoltes de plusieurs années, qu'en moyenne le poids d'une pomme d'une variété A, est supérieur de 12 grammes au poids d'une pomme d'une variété B. Il veut savoir si cette différence subsiste si l'on utilise un nouveau type d'engrais. On admet que dans tous les cas le poids d'une pomme suit une loi normale de variance 66. Le producteur pèse 25 pommes de la variété A et de la variété B. Ces 50 pommes ont toutes été traitées avec le nouvel engrais. Il observe alors un poids moyen de 116 grammes pour la variété A et de 109 grammes pour la variété B. Tester si le nouvel engrais améliore de façon identique le rendement pour les variétés A et B.