Initiation aux statistiques
Feuille de révisions

Exercice 1

Nous observons deux caractères statistiques sur un même individu (une moule). Il s'agit du poids brut et du poids utile. Le tableau suivant représente 10 observations de ces deux caractères.

Poids brut	3.75	7.05	5.83	9.58	1.31	8.58	7.16	4.80	6.32	6.90
Poids utile	1.51	2.60	1.33	3.49	0.36	3.57	2.98	1.74	3.12	3.08

1) Calculer les paramètres de position (moyenne, médiane) des deux séries.
2) Calculer les paramètres de dispersion (variance, écart-type, écart inter-quartile) des deux séries.
3) Calculer le coefficient de corrélation entre les deux séries. Peut-on dire qu'il y a une dépendance entre les deux caractères statistiques?

Exercice 2

Dans un concours, les candidats passent deux épreuves de mathématiques, l'une en probabilités, l'autre en statistiques.
On note X la note obtenue à l'épreuve de probabilités, et Y celle obtenue à l'épreuve de statistiques. Les résultats obtenus pour 104 candidats sont donnés dans le tableau ci-dessous :

X \ Y	[0 ; 4[	[4 ; 8[	[8 ; 12[	[12 ; 16[	[16 ; 20[	Total
[0 ; 4[	1	1	0	0	0	2
[4 ; 8[	1	3	5	11	0	20
[8 ; 12[	2	10	10	28	0	50
[12 ; 16[	0	1	3	9	11	24
[16 ; 20[	0	0	2	4	2	8
Total	4	15	20	52	13	104

1) Calculer la moyenne et l'écart type de X et de Y.
2) Calculer la fréquence conditionnelle de $X \in [12 ; 16[$ sachant $Y \in [4 ; 8[$.
3) Calculer la covariance de X et Y, et le coefficient de corrélation.

Exercice 3

Nous nous intéressons à un élément du système de navigation de bord. Le tableau suivant représente les instants de premières défaillances de ces éléments, observés sur 10 avions.

Numéro de l'avion	1	2	3	4	5	6	7	8	9	10
Temps de défaillance	6	8*	35	65	30	80*	52	88	20	36*

Temps de premières défaillances, exprimés en mois. Les astérisques représentent des éléments suspendus.

1) Donner une estimation de la fiabilité $R(t)$, en utilisant la méthode de Kaplan-Meier.
2) Peut-on justifier d'une loi exponentielle pour modéliser le temps de bon fonctionnement des éléments étudiés ? Si oui, donner une estimation du MTTF et du paramètre λ.

Exercice 4
Un biologiste étudie un type d'algue qui attaque les plantes marines. La toxine contenue dans cette algue est obtenue sous forme de solution organique. Il mesure la quantité de toxine par gramme de solution. Il a obtenu les mesures suivantes, exprimées en milligrammes :

Quantité	0,5	0,6	0,7	0,8	0,9	1	1,1	1,2	1,3	1,4	1,5
Fréquence	1	4	6	5	9	8	12	10	7	5	1

On suppose que ces mesures sont les réalisations de variables aléatoires indépendantes et identiquement distribuées suivant une loi normale d'espérance *m*, et d'écart type σ.

1) Donner une estimation ponctuelle de l'espérance *m* et de l'écart type σ de la quantité de toxine par gramme de solution.
2) Déterminer un intervalle de confiance à 95% pour l'espérance *m* de la quantité de toxine par gramme de solution (on prendra comme écart type l'estimation ponctuelle de σ).

Exercice 5

Soit X une variable aléatoire réelle de densité : $f_X(x) = \frac{2}{\theta} e^{-\frac{2}{\theta}x} 1_{\mathbb{R}_+}(x)$.

On cherche à estimer le paramètre θ. Pour cela, on crée un estimateur $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i$, où $(X_1, X_2, \ldots, X_n)$ est un échantillon de même loi que X.

1) Calculer, en utilisant une intégration par partie, $IE\left(\hat{\theta}\right)$. Cet estimateur est-il sans biais de θ ?
2) Proposer un estimateur sans biais de θ, fonction de $\hat{\theta}$. Ce nouvel estimateur est-il convergent vers θ ?

Exercice 6
La presse affirme que parmi les Français regardant la télévision plus de 4 heures par jour, on a la même proportion de personnes dans chaque tranche d'âge. Sur un échantillon de 30 personnes, on a observé les données suivantes :

Age	Moins de 20 ans	De 20 à 35 ans	De 35 à 50 ans	Plus de 50 ans
Effectif	11	4	5	10

1) Que penser de l'opinion de la presse au seuil de 5% ?
2) Que penser de l'opinion selon laquelle un quart des Français regardant la télévision plus de 4 heures par jour a moins de 20 ans, au seuil 5% ?

Exercice 7
Une étude américaine effectuée sur 106 patients a donné lieu au tableau suivant :

	A un cancer du poumon	N'a pas de cancer du poumon
Est fumeur	60	32
N'est pas fumeur	3	11

Peut-on considérer au risque de 5% qu'il existe un lien entre le fait de fumer et avoir un cancer du poumon ?