

Exercice 1 :

On suppose que le nombre K d'absences par semaine dans une école d'ingénieurs X suit une loi de Poisson de paramètre λ . On a observé l'échantillon suivant où n_k est le nombre de semaines pour lesquelles on a relevé k absences.

k	0	1	2	3	4	5	6	7
n_k	7	18	16	35	12	7	4	1

(4p) Déterminer $\hat{\lambda}_{MV}$ - l'estimateur du maximum de vraisemblance de λ . Est-il efficace ? Déterminer son espérance, sa variance et sa loi asymptotique.

(4p) Le conseil du département émet l'hypothèse (de référence) qu'il y a 3 absences en moyenne par semaine, alors que les délégués des étudiants disent eux qu'il y a seulement 2,6. Tester ces hypothèses au niveau de signification $\alpha^*=0,05$. Justifier le choix de la région critique. Cela correspond-t-il à votre intuition ? Déterminer la puissance du test.

Exercice 2 :

La principale cause de l'absentéisme dans l'école X était, à cause de son emplacement dans un climat humide, le rhume. 6 étudiants se sont portés volontaires pour une étude sur la sensibilité au vaccin anti-rhume.

Cette sensibilité était mesurée pour chacun d'entre eux par le nombre de rhumes par an. Le tableau ci-dessous fournit les résultats de mesures - sans vaccin - pour l'an 199x et - avec vaccin - pour l'année suivante.

(4p) La sensibilité au vaccin devrait se manifester par une diminution du nombre de rhumes par an. Tester la validité de cette hypothèse au niveau de signification $\alpha=0,1$.

X (Sans vaccin)	2	3	5	2	6	7
Y (Avec vaccin)	1	4	2	2	4	10

Exercice 3 :

Dans l'école X , un professeur enseigne la même matière à 2 classes, A et B , de 16 et 25 élèves respectivement. Au cours d'un même examen, bien qu'il n'y ait aucune différence significative entre les moyennes, l'écart-type de la classe A est 9, alors que celui de la classe B est 12. Nota : l'on suppose que A et B suivent des lois normales.

(4p) Peut-on conclure qu'au niveau de signification (a) 0,01 et (b) 0,05, il y ait plus de dispersion dans la classe B que dans la classe A .

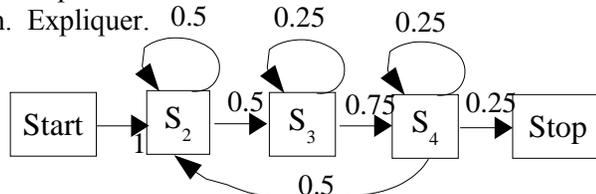
(1p) Est-ce que les conclusions tirées précédemment changeraient s'il y avait une différence significative entre les moyennes des classes ?

Questions :

(1p) Qu'est qu'une p -value ? Connaissant la valeur de la p -value, quelle est la règle d'acceptation de l'hypothèse H_0 au niveau de signification α ? Que peut-on dire si la valeur de p -value est très proche de 1 ?

(1p) Quel est l'objectif d'une Analyse en Composantes Principales ? Quel est l'indice utilisé pour mesurer la liaison entre les variables observées pour chaque individu de l'échantillon ? Quel est l'objectif d'une Analyse Discriminante ? Donner un exemple d'utilisation.

(1p) Calculer la probabilité de générer la séquence d'observations $\{a, b, a, c\}$ avec le modèle de Markov caché (MMC) suivant. Préciser quel chemin avez-vous choisi pour la génération de cette séquence. Pensez-vous qu'il soit le meilleur. !! Aucune observation ne pourra être générée dans les états muets *Start* et *Stop*. Est-ce que ce MMC pourrait être utilisé dans le cadre de la reconnaissance de la parole pour modéliser une syllabe ou un son. Expliquer.



Observations = $\{a, b, c, d\}$
 $P(a | S_2) = 1$
 $P(a | S_3) = 0.25 ; P(b | S_3) = 0.25 ; P(c | S_3) = 0.5$
 $P(b | S_4) = 0.75 ; P(d | S_4) = 0.25$

NOTA : L'examen est noté sur 20 points.