Méthodes statistiques Durée : 3 Heures

1- ECUE 1 : Lois de probabilités et comportements asymptotiques. (STAT2 & PLAN 2)

1.1 Let A and B be two events. Express the following events in words, and display them in a Venn diagram:

a) $A \cap B$ b) $A \cap B^*$ c) $A^* \cap B^*$

d) Prove (using a Venn diagram) the rule

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

e) Prove (using a Venn diagram) de Morgan's rules:

$(A \cup B)^* = A^* \cap B^*$ and $(A \cap B)^* = A^* \cup B^*$

1.2 At the production of a certain item, two types of defects, A and B, can occur. We know that $P(A) = 0.1$, $P(B) = 0.2$ and $P(A \cap B) = 0.05$. Compute the probability that a produced unit has

a) at least one of the defects

b) defect A but not defect B

c) none of the defects

d) precisely one of the defects A and B

1.3 We take at random three cards (without replacement) from an ordinary deck of cards (of 52 cards). Compute (using the classical probability definition) the probabilities for the events

a) all three are hearts

b) none of the cards is hearts

c) all three are aces.

1.4 (Challenging!) Compute the probability that out of 23 randomly selected individuals, at least two of them have the same birthday. (Assume that all birthdays are equally common, and that there are 365 days in a year.)

STAT 2 & PLAN 2

1.5. Déterminer les paramètres m et σ de loi normale suivie par X, sachant que $P\{X<2\} = 0,10$ et $P\{X>5\} = 0,30$.

1.6. Soit $u_1, u_2, \ldots$ une suite de v.a indépendantes de même loi définie par $P\{u_n=1\}=p$, $p\in]0,1[$, et $P\{u_n=-1\}=1-p=q$. On définit la suite de v.a $V_n=U_1U_2\ldots U_n$. Déterminer $E(V_n)$ et en déduire la loi de V_n puis sa loi limite.

1.7. Afin d'établir le profil statistique de certains malades d'un hôpital, on prélève au hasard et avec remise 100 dossiers médicaux. Malheureusement, on constate que 20 dossiers d'entre eux sont incomplets et donc inexploitables. Si on considère qu'il faut en prélever au moins 1000 dossiers, combien faudra-t-il en prélever pour que cette condition soit réalisée avec une probabilité égale à 0,95 ?

1.8. Soit X une variable aléatoire de densité f continue et de fonction de répartition F strictement monotone. Exprimer à l'aide de f et F la densité de probabilité et la fonction de répartition de chacune des variables aléatoires suivantes :

1. $Y = ax + b$, $a \in \mathbb{R}_+$ et $b \in \mathbb{R}_+$
2. $Z = |X|$
3. $T = X^2$
4. $U = ln|X|$
5. $V = F(X)$
6. $W = [X]$

1.9. Soient $X_1, \ldots, X_n$ n v.a. réelles indépendantes suivant une même loi d'espérance m et de variance σ^2. Soit S_n la v.a définie par : $S_n = \frac{1}{n}\sum_{i=1}^{n} x^2$.

a) Étudier la convergence en probabilité de S_n quand $X_1, \ldots, X_n$ suivent une loi normale N(0,1). En déduire une approximation de la loi de Student T_n quand n tend vers l'infini.

b) Vers quoi converge S_n dans le cas général ?

2 - ECUE N°2 : Estimations et tests

2.1. Le P.D.G d'une chaine de magasins à grande surface voudrait savoir s'il doit entreprendre une campagne publicitaire à l'échelon national pour augmenter la vente de ses produits. Pour prendre une décision, il fait ne campagne publicitaire dans cinq villes et observe les accroissements $x_i (1 \leq i \leq 5)$ de bénéfices correspondants. On suppose que $x_i (1 \leq i \leq 5)$ suit une loi $N(m,\sigma)$ d'écart-type connu $\sigma=2$ et $m=0$ si la campagne est inefficace et $m=2$ si elle est efficace. Toutes les données sont exprimées en million de francs (MF) et le coût par ville de la campagne publicitaire est évaluer à 0,5. Les accroissements de bénéfices observés dans les cinq villes sont : 1 ; 1,5 ; 0 ; 1 ; 0,5.

1. Quelle décision prendra ce P.D.G s'il pense qu'une telle campagne a une chance sur deux de réussir ?

2. a) Quelle serait sa décision s'il souhaitait avant tout limiter le risque de ne pas entreprendre une campagne qui serait efficace ?

2. b) Même question si la campagne est considérée comme efficace dès que $m>0$.

2.2. On étudie la circulation en un point fixe d'une autoroute en comptant, pendant deux heures, le nombre de voitures passant par minute devant un observateur. Le tableau suivant résume les données obtenues :

Débit en voiture par minute	Fréquences observées
0	4
1	9
2	24
3	25
4	22
5	18
6	6
7	5
8	3
9	2
10	1
11	1
	120

Tester l'adéquation de la loi empirique observée à une loi théorique simple ; on prendra $\alpha=0,10$.

STAT2 & PLAN2

2.3. Un examen est ouvert à des étudiants d'origines différentes : économie, informatique, mathématique. Le responsable de l'examen désire savoir si la formation initiale d'un étudiant influe sur sa réussite. A cette fin, il construit le tableau ci-dessous à partir des résultats obtenus par les 286 candidats.

Origine / Résultat	Économie	Informatique	Mathématiques	Total
Réussite	41	59	54	154
Échec	21	36	75	132
Total	62	95	129	286

Quelle est sa conclusion ?

2.4. Afin d'établir le profil statistique de certains malades d'un hôpital, on prélève au hasard et avec remise 100 dossiers médicaux. Malheureusement, on constate que 20 dossiers d'entre eux sont incomplets et donc inexploitables. Si on considère qu'il faut en prélever au moins 1000 dossiers, combien faudra-t-il en prélever pour que cette condition soit réalisée avec une probabilité égale à 0,95 ?

2.5. Soit $(X_1, \ldots, X_n)$ un n échantillon d'une loi de normale N(m, σ^2). Déterminer l'estimateur du maximum de vraisemblance T_n du paramètre m et étudier ses propriétés.

Le fait que σ soit connu ou non modifie-t-il le résultat ?

FIN

STAT2 & PLAN2