HEC 1986 MII

L'objet du problème est de décrire et comparer deux méthodes de détection de pannes. On se place dans la situation suivante : on considère un ordinateur comprenant un ensemble C de n circuits intégrés, où $n \geqslant 2$, et on suppose qu'une panne a endommagé un circuit et un seul. On note c_1, c_2, …, c_n ces circuits, où c_n est le circuit défectueux.

Dans la partie I on étudie une méthode de tests par paquets quelconques de circuits. Dans la partie II, on évalue l'espérance du nombre de tests nécessaires pour détecter la panne par cette méthode et on compare celle-ci avec une méthode de tests où les circuits sont pris un à un.

PARTIE I : Tests par tirages de paquets de circuits, avec remise

On note Ω l'ensemble des parties de C. On tire au hasard et de façon équiprobable, par un procédé adéquat, des parties de C, c'est à dire des éléments de Ω (y compris la partie vide). Pour tout élément G de Ω, on note $\overline{G}$ le complémentaire de G dans C.

Q1 Soit A une partie de C de cardinal r, où $0 \leqslant r \leqslant n$.

a) Déterminer le nombre de parties de C ne rencontrant pas A.

b) Déterminer le nombre de parties de C contenant A.

Q2 On tire un élément G de Ω.

a) Calculer la probabilité d'obtenir une partie **donnée** B de C.

b) Soit A une partie de C de cardinal r, où $0 \leqslant r \leqslant n$. Calculer la probabilité pour que G contienne A (*JF* $\frac{1}{2^r}$).

c) Soit c_j un élément de C distinct de c_n. On suppose que c_n appartient à G ; calculer la probabilité conditionnelle pour que c_j appartienne à G (*JF* *on cherche donc la probabilité pour que c_j appartienne à G sachant que c_n est dans G ; revenir à la définition d'une proba conditionnelle*).

Q3 Soit Ω_n l'ensemble des parties de C contenant l'élément c_n. Soit h un nombre entier naturel non nul.

On tire, successivement et avec remise, h éléments B_1, B_2, …, B_h de Ω.

Une technique permet de tester un ensemble de circuits et de savoir si le circuit défectueux se trouve parmi eux.

Pour chaque entier i appartenant à l'intervalle $[\![1, h]\!]$, on teste le paquet B_i. On pose $G_i = B_i$ si B_i contient le circuit défectueux ; dans le cas contraire, on pose $G_i = \overline{B_i}$. Ainsi, G_i est un élément de Ω_n. On désigne enfin par D_h l'intersection des parties G_i.

a) Montrer que, pour tout nombre entier naturel j tel que $1 \leqslant j \leqslant n-1$, la probabilité pour que c_j appartienne à D_h est égale à $\dfrac{1}{2^h}$ (*JF* *on commencera par calculer proprement et simplement $P(c_j \in G_i)$)*.

b) Plus généralement, soit E une partie de C de cardinal r ne contenant pas l'élément c_n. Calculer la probabilité pour que D_h contienne E.

c) En déduire que les événements $c_j \in D_h$, où $1 \leqslant j \leqslant n-1$, sont mutuellement indépendants et qu'il en est de même pour les événements $c_j \notin D_h$ ((*JF* *on commencera par donner la définition de l'indépendance mutuelle de n événements ; le second point est du cours*).

d) Prouver enfin que : $P(D_h = \{c_n\}) = \left(1 - \dfrac{1}{2^h}\right)^{n-1}$.

PARTIE II : Étude du nombre de tests nécessaires pour détecter la panne

Q1 Pour tout nombre entier naturel k, on pose : $a_k = 1 - \left(1 - \dfrac{1}{2^k}\right)^{n-1}$.

a) Montrer que l'application qui à tout nombre entier naturel non nul k associe $a_{k-1} - a_k$ définit une loi de probabilité sur $\mathbb{N}^*$.

b) Déterminer la fonction de répartition associée à cette loi.

Facultatif ! Construire sa courbe représentative lorsque $n = 3$.

Q2 On conserve la procédure de la question I.3, mais le nombre de tirages n'est pas fixé ; pour tout nombre entier naturel non nul k, on note D_k l'intersection des parties $G_1, G_2, \ldots, G_k$.

On considère une variable aléatoire X_n à valeurs dans $\mathbb{N}^*$, qui prend la valeur 1 si l'événement $D_1 = \{c_n\}$ est réalisé, la valeur k, où $k \geqslant 2$, si l'événement $D_k = \{c_n\}$ est réalisé et si l'événement $D_{k-1} = \{c_n\}$ ne l'est pas.

(Une telle variable aléatoire représente donc le nombre de tests nécessaires pour détecter le circuit défectueux).

a) Pour tout nombre entier naturel non nul k, calculer la probabilité de l'événement $\{X_n \leqslant k\}$.

b) En déduire que la loi de probabilité de X_n est celle qui a été définie dans la question 1.a.

Q3 a) Montrer que, pour tout nombre entier naturel non nul s fixé, la série de terme général $\left(\dfrac{1}{2^{ks}}\right)$ est convergente.

En déduire que la série de terme général (a_k) est convergente.

b) Montrer que la variable aléatoire X_n admet une espérance et que : $E(X_n) = \displaystyle\sum_{k=0}^{+\infty} a_k$.

À cet effet, on pourra calculer : $\displaystyle\lim_{q\to+\infty} \left(\sum_{k=1}^{q} k(a_{k-1} - a_k) - \sum_{k=0}^{q-1} a_k\right)$.

Q4 On se propose d'évaluer $E(X_n)$ en comparant la série de terme général (a_k) à une intégrale. À cet effet, pour tout nombre entier naturel non nul p, on pose :

$$I_p(x) = \int_0^x f_p(t)\,\mathrm{d}t \qquad \text{où} \qquad f_p(t) = 1 - (1 - e^{-t})^p$$

a) Montrer que l'intégrale $I_p = \displaystyle\int_0^{+\infty} f_p(t)\,\mathrm{d}t$ est convergente (on pourra développer $1-(1-e^{-t})^p$ $\boxed{JF}$ *Bof, tu crois ?*).

b) Calculer une primitive de la fonction $f_{p+1} - f_p$ sur l'intervalle $[0, +\infty[$. En déduire la valeur de $I_{p+1} - I_p$. Montrer finalement que :

$$I_{n-1} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1}$$

c) Montrer que pour tout nombre entier naturel $m \geqslant 2$: $\displaystyle\int_m^{m+1} \frac{\mathrm{d}t}{t} \leqslant \frac{1}{m} \leqslant \int_{m-1}^{m} \frac{\mathrm{d}t}{t}$.

En déduire que : $\ln(n) \leqslant I_{n-1} \leqslant 1 + \ln(n-1)$. Calculer $\displaystyle\lim_{n\to+\infty} \frac{I_{n-1}}{\ln(n)}$

d) Soit g_n la fonction numérique définie sur $[0; +\infty[$ par la relation : $g_n(u) = 1 - \left(1 - \dfrac{1}{2^u}\right)^{n-1}$.

En étudiant la variation de g_n, montrer que, pour tout nombre entier naturel non nul k :

$$a_k \leqslant \int_{k-1}^{k} g_n(u)\,\mathrm{d}u \leqslant a_{k-1}$$

e) Montrer que, pour tout nombre entier naturel $q \geqslant 2$: $\displaystyle\sum_{k=1}^{q} a_k \leqslant \int_0^q g_n(u)\,\mathrm{d}u \leqslant \sum_{k=0}^{q-1} a_k$.

En effectuant le changement de variable $t = u\ln(2)$ dans cette intégrale et en passant à la limite dans l'encadrement précédent lorsque q tend vers $+\infty$, montrer que :

$$E(X_n) - 1 \leqslant \frac{I_{n-1}}{\ln(2)} \leqslant E(X_n)$$

f) Déduire des résultats des questions c) et e) un encadrement de $E(X_n)$ et déterminer la limite de $\dfrac{E(X_n)}{\ln(n)}\ln(2)$

Q5 On teste maintenant les circuits c_j un par un, en les tirant de manière équiprobable et sans remise. On désigne par Y_n le nombre de tests nécessaires pour détecter le circuit défectueux.

a) Déterminer la loi de la variable aléatoire Y_n.

b) Calculer l'espérance de Y_n.

Q6 a) Calculer la limite de $\dfrac{E(X_n)}{E(Y_n)}$ lorsque n tend vers $+\infty$.

b) Grâce à l'encadrement obtenu dans la question 4 f), comparer $E(X_n)$ et $E(Y_n)$ lorsque $n = 100$, puis lorsque $n = 1000$.
