



DISVE
Pôle Scolarité

Année universitaire 2010 - 2011
Session 1 d'automne



Parcours/Etape: CSB3 UE: INF 162
Epreuve: Probabilités pour l'informatique
Date: 3 janvier 2011 Heure: 8h30 Durée:
1h30
Documents non autorisés
Epreuve de M. Melançon

Examen blanc – Epreuve préparatoire

Exercice 1. “Vrai ou faux / Choix multiples” : une bonne réponse vaut 1 point, une mauvaise réponse vaut -1 point

Note : les réponses données ici sont justifiées – aucune justification n'était cependant demandée.

1. On désigne par $U = \text{random}()$ l'appel à la fonction random qui retourne un nombre réel aléatoire uniformément sur $[0, 1)$. L'instruction $\lfloor \frac{\log U}{\log 1-p} \rfloor + 1$ permet de simuler :
 - la loi géométrique
 - la loi binomiale
 - la loi exponentielle

La loi géométrique. En effet, l'expression donnée $X = \lfloor \frac{\log U}{\log 1-p} \rfloor + 1$

calcule bien un entier, et on a

$$\begin{aligned}P(X = k) &= P\left(\lfloor \frac{\log U}{\log 1-p} \rfloor + 1 = k\right) \\&= P\left(k-1 \leq \frac{\log U}{\log 1-p} < k\right) \\&= P\left((1-p)^{k-1} \leq U < (1-p)^k\right) \\&= (1-p)^{k-1} - (1-p)^k = (1-p)^{k-1}(1 - (1-p)) = (1-p)^{k-1}p\end{aligned}$$

2. Soit X_1 une loi binomiale de paramètres N_1 et p , et X_2 une loi binomiale de paramètres N_2 et p . La variable aléatoire $X = X_1 + X_2$ est :

- une loi binomiale de paramètres $N_1 + N_2, 2 \cdot p$
- une loi binomiale de paramètres N_1, N_2 et p
- une loi binomiale de paramètres $N_1 + N_2, p$
- aucune de ces réponses

Loi binomiale de paramètres $N_1 + N_2, p$. En effet, on a

$$\begin{aligned}P(X_1 + X_2 = k) &= \sum_{i=0}^k P(X_1 = i)P(X_2 = k-i) \\&= \sum_{i=0}^k \binom{N_1}{i} p^i (1-p)^{N_1-i} \binom{N_2}{k-i} p^{k-i} (1-p)^{N_2-(k-i)} \\&= p^k (1-p)^{N_1+N_2-k} \sum_{i=0}^k \binom{N_1}{i} \binom{N_2}{k-i} \\&= \binom{N_1 + N_2}{k} p^k (1-p)^{N_1+N_2-k}\end{aligned}$$

3. La moyenne et la variance d'une variable X de Bernoulli de paramètre p sont :

- $\mathbb{E}(X) = p, \mathbb{V}(X) = 1$
- $\mathbb{E}(X) = 0, \mathbb{V}(X) = 1$
- $\mathbb{E}(X) = p, \mathbb{V}(X) = 1 - p$
- $\mathbb{E}(X) = 0, \mathbb{V}(X) = p$

$$\square \mathbb{E}(X) = p, \mathbb{V}(X) = p(1 - p)$$

L'espérance d'une variable de Bernouilli X est $\mathbb{E}(X) = 0 \cdot (1 - p) + 1 \cdot p = p$. La variance est $\mathbb{V}(x) = (0 - p)^2 \cdot (1 - p) + (1 - p)^2 \cdot p = p(1 - p)(p + (1 - p)) = p(1 - p)$.

Exercice 2. On veut calculer une bonne estimation du nombre π . On tire au hasard des points dans le carré $[0, 1) \times [0, 1)$. Pour chacun des tirages, soit le point tombe à l'intérieur du cercle de diamètre 1 inscrit dans le carré, soit il est à l'extérieur. On considère la variable aléatoire X qui donne le nombre total de points qui tombent à l'intérieur du cercle (après N tirages). Justifiez chacune de vos affirmations.

1. Peut-on voir *chacun* des tirages comme un évènement aléatoire ? de quel type ? quel paramètre(s) ?
2. Quelle la distribution de la variable X ? quels sont ses paramètres ?
3. On souhaite obtenir une estimation du nombre π avec deux décimales exactes. Combien de tirage faut-il effectuer pour obtenir cette estimation avec une probabilité d'au moins 95% ?

Le cercle inscrit dans le carré est de rayon $1/2$ et donc d'aire $\pi/4$, c'est donc la probabilité qu'on a qu'un point (x, y) choisi aléatoirement dans le carré tombe dans le cercle. On désigne par X_i ($i = 1, \dots, N$) la variable aléatoire de Bernouilli correspondant au tirage du i ème point, où un succès correspond au tirage d'un point *dans* le cercle avec probabilité $p = \pi/4$. Le nombre total de points qui tombe effectivement dans le cercle est $S_N = X_1 + \dots + X_N$, que l'on compare à N pour obtenir une approximation de $\pi/4$ (de laquelle on déduit une approximation de π). La variable S_N suit donc une loi binomiale de paramètres N et p .

On aura une approximation de $\pi/4$ à deux décimales exactes si $S_N/N - \pi/4 < 0.01$. Or, on a :

$$P\left(\left|\frac{S_N}{N} - \frac{\pi}{4}\right| < \epsilon\right) = 1 - P\left(\left|\frac{S_N}{N} - \frac{\pi}{4}\right| \geq \epsilon\right) \geq 1 - \frac{\mathbb{V}(X_1)}{N\epsilon^2}$$

puisque selon la loi des grands nombres on a $P\left(\left|\frac{S_N}{N} - \mathbb{E}(X_1)\right| \geq \epsilon\right) \leq \frac{\mathbb{V}(X_1)}{N\epsilon^2}$. Par ailleurs, $1 - \frac{\mathbb{V}(X_1)}{N\epsilon^2}$ si et seulement si $\frac{\mathbb{V}(X_1)}{N\epsilon^2} < 0.05$. Or, ici $\mathbb{V}(X_1) = \pi/4(1 - \pi/4)$ et par conséquent, en prenant même une "mauvaise" approximation de π à 3.14, on trouve : $\frac{\mathbb{V}(X_1)}{N\epsilon^2} \sim \frac{0.168775}{10^{-4}} N$. On doit donc

prendre $N > \frac{0.168775}{10^{-40.05}} = 33755$. Il faut donc utiliser près de 35000 points pour obtenir une approximation à deux décimales près de π .

Exercice 3. Imaginez que vous collectionnez des objets offerts à chaque achat au supermarché. La collection compte N objets différents et vous vous mettez en tête de continuer à fréquenter ce supermarché pour faire vos courses, une fois chaque semaine, jusqu'à ce que vous ayez la collection complète. Pendant combien de semaines aurez-vous à faire vos courses avant d'avoir obtenu la collection complète (en moyenne) ?

Modélisez la situation ainsi. Vous considérez les variables X_1, \dots, X_N qui indique le temps d'attente (en nombre de semaines) avant d'obtenir une i ème carte.

1. Que vaut $X_1 = 1$?
2. Calculer $P(X_2 = 1), P(X_2 = 2)$.
3. Calculer $P(X_3 = 1), P(X_3 = 2)$.
4. De manière générale, quelle est la loi de probabilité de la variable X_k ?
5. Quel est la moyenne de la variable X_k ?
6. On considère la variable aléatoire $X = X_1 + X_2 + \dots + X_N$. Que permet de calculer cette variable ?
7. Calculer sa moyenne.

On a $X_1 = 1$ puisque la première carte de la collection est forcément "nouvelle". La collection compte N cartes, on a donc une probabilité $P(X_2 = 1) = \frac{N-1}{N}$ d'avoir une "nouvelle" carte dès le second achat. On a en revanche une probabilité $\frac{1}{N}$ de tomber sur la carte que l'on possède déjà. Ainsi, on a une probabilité $P(X_2 = 2) = \frac{1}{N} \frac{N-1}{N}$ de n'obtenir une seconde carte de la collection qu'à la troisième semaine. On raisonne de la même manière pour trouver $P(X_3 = 1) = \frac{N-2}{N}$ (puisque'il reste $N - 2$ cartes à collectionner), et $P(X_3 = 2) = \frac{2}{N} \frac{N-2}{N}$ (puisque l'on possède déjà deux cartes de la collection).

De manière générale, la variable X_k suit une loi géométrique de paramètre $p_k = \frac{N-(k-1)}{N}$ puisque l'on possède déjà $k - 1$ cartes de la collection. Par conséquent, sa moyenne est $\mathbb{E}(X_k) = 1/p_k = N/(N - k + 1)$. La variable $X = X_1 + X_2 + \dots + X_N$ permet de calculer le nombre de semaines nécessaires pour collectionner la totalité des cartes. Les variables étant indépendantes, on a :

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(X_1 + X_2 + \cdots + X_N) = \sum_{k=1}^N \mathbb{E}(X_k) \\ &= \sum_{k=1}^N \frac{N}{N-k+1} = N \sum_{k=1}^N \frac{1}{k}\end{aligned}$$

(on a $\lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{1}{k} \sim \log(N)$, ainsi le nombre de semaines requis en moyenne pour obtenir toute la collection est de l'ordre de $N \log N$.)