

Devoir Maison 1

lagache@biologie.ens.fr

A rendre le mardi 23 Novembre 2011/vendredi 26 Novembre 2011

Un directeur de grande surface souhaite étudier la distribution statistique des tickets de caisse les mercredi et samedi. Il souhaite en effet ajuster le nombre de caisses C pour chaque jour. Au cours du mois d'Octobre, sur les 6000 tickets de caisse observés le mercredi, il obtient la répartition suivantes (le ticket minimum est de 5 euros et le ticket maximum est de 300 euros):

	montant du ticket en euros	proportion des tickets
i=1	de 5 à moins de 30	0.09
i=2	de 30 à moins de 60	0.12
i=3	de 60 à moins de 100	0.22
i=4	de 100 à moins de 150	0.25
i=5	de 150 à moins de 200	0.19
i=6	de 200 à 300	0.13
	Total	1

alors que le samedi, les 6000 tickets se répartissent suivant la distribution suivante:

	montant du ticket en euros	proportion des tickets
i=1	de 5 à moins de 30	0.20
i=2	de 30 à moins de 60	0.14
i=3	de 60 à moins de 100	0.16
i=4	de 100 à moins de 150	0.19
i=5	de 150 à moins de 200	0.13
i=6	de 200 à 300	0.18
	Total	1

1. Pour ces deux distributions, quelle est la population? Quelle est la variable? De quel type de variable s'agit-il?
2. Tracer les histogrammes des deux échantillons (en fréquence relative) avec les cellules pré-définies par l'énoncé.

On suppose à présent que les montants des tickets dans chaque intervalle de prix sont uniformément répartis.

1. En déduire, pour chaque jour, la moyenne, l'écart-type et les différents quartiles. Discuter.

2. Tracer les fonctions de répartition des 2 échantillons et représenter graphiquement la distance de Kolmogorov-Smirnov.

Pour un montant de caddie $5 \leq m \leq 300$, on modélise le temps t nécessaire pour passer en caisse par $t = t_0 + \alpha m + e^{\frac{m-m_0}{\beta}}$ avec $t_0 = 5$ minutes, $\alpha = \frac{1}{30}$, $m_0 = 150$ et $\beta = 150/\ln(10)$.

1. Selon vous, quelle est l'origine de chacun des trois termes t_0 , αm et $e^{\frac{m-m_0}{\beta}}$.
2. Calculer, pour chaque intervalle de montant, le temps moyen $\bar{t}_i$, $1 \leq i \leq 6$ (on supposera encore que le montant des caddies est uniformément réparti dans chaque intervalle).
3. En déduire, pour chaque jour le temps de caisse moyen total $\bar{t} = \sum_{i=1}^6 \bar{t}_i$ et le nombre de caisse $C = \bar{t}/400$ nécessaire. Conclure.