

Problème du collecteur de coupon

Auteur: Nicolas Gast

Le problème du collecteur de coupon intervient souvent dans l'analyse d'algorithmes randomisés (protocoles de communication, table de hachage,...). Dans ce TD, on se propose de simuler un tel système.

Votre petite soeur est fan de foot et a acheté un album panini. Pour le compléter, elle doit réussir à collectionner un sticker de chacun des n footballeurs qui composent l'album. Les stickers peuvent être achetés par paquets. Chaque paquet contient k stickers différents et il est impossible de savoir quels stickers sont contenus dans un paquet. Dans les deux premiers exercices, on suppose que l'éditeur est honnête et donc que les paquets sont remplis de manière aléatoire et indépendante et que chaque paquet contient k stickers différents, tirés uniformément parmi l'ensemble des combinaisons de k stickers différents.

Exercice 1: Panini tout seul

1. En utilisant uniquement l'intuition, essayer de deviner combien de paquets faut-il en moyenne acheter pour compléter une collection de 100 stickers ?
2. Écrire une fonction `panini_seul = fonction(n,k)` qui simule une collection de n stickers et rend le nombre de paquets que votre petite soeur aura du acheter pour compléter son album.
3. Simuler 100 scénarios pour $n = 100$ et $k \in \{1, 5\}$ et tracer l'histogramme du nombre de stickers à acheter dans chaque cas.
4. Calculer le nombre moyen de stickers à acheter pour $n = 100$ et $k \in \{1, 5\}$.
5. Le résultat dépend-t-il de k ? Est-il conforme à l'intuition ?

Exercice 2: Panini à plusieurs

Les m enfants de la classe de votre petite soeur cherchent maintenant à compléter leur m albums panini. Pour cela, ils décident de mettre leur argent en commun et de remplir les albums dans l'ordre.

1. Écrire une fonction `panini_plusieurs = fonction(n,k,m)` qui rend un vecteur $(X_1 \dots, X_m)$ où X_1 est le nombre de stickers qu'il aura fallu acheter pour compléter le premier album panini, X_2 le nombre de stickers supplémentaires qu'il aura fallu pour compléter le deuxième panini, etc.
2. Calculer la moyenne de X_i pour $n = 100$, $k = 5$ et $m = 20$. Quelle différence y a-t-il entre la moyenne de X_1 et les moyennes des X_i ($i \geq 2$).
3. Que penser de la collection à plusieurs par rapport à la collection seul ?

Exercice 3: Panini non uniformes

Afin de gagner plus d'argent, l'éditeur a décidé que la vignette d'Ibrahimovic apparaîtrait 10 fois moins souvent que les autres.

1. Écrire une fonction `panini_non_uniforme = fonction(n,p)` qui rend le nombre de stickers à acheter lorsque la vignette 1 apparaît p fois moins souvent que les autres.
2. Calculer le nombre moyen de stickers à acheter pour $n = 100$ et $p = 10$ et $p = 100$.
3. Commenter l'effet de p .