1 Panini

Votre petit cousin collectionne les vignettes autocollantes des footballeurs célèbres. Il achète au marchand de journaux des pochettes de $k = 5$ vignettes à un prix de 2 euros. L'album de la collection complète contient $M = 200$ places pour les vignettes autocollantes (prix d'achat de l'album 4 euros). Il souhaite savoir combien de pochettes il devra acheter pour avoir toute la collection. Saurez-vous lui répondre ?

On **simplifie le modèle** en supposant que l'on achète les vignettes **une par une** ($k = 1$) et on note T le nombre total de vignettes achetées pour avoir toute la collection.

On suppose également que l'éditeur des vignettes est honnête et qu'il imprime les vignettes dans les mêmes proportions (il n'y a pas de vignettes plus rares que d'autres).

1.1 Modélisation

On note $X_1, \ldots X_n, \ldots$ la séquence aléatoire de vignettes obtenues successivement à chaque achat.

Lorsque la collection comporte déjà $i - 1$ vignettes distinctes, on note Y_i le nombre de vignettes à acheter pour avoir une nouvelle vignette distincte.

Question 1.1 : Quelles hypothèses statistiques raisonnables peut-on faire sur les X_i ? Donner, selon ces hypothèses, la loi des X_i, c'est-à-dire $\mathbb{P}[X_i = k]$, $1 \leqslant k \leqslant M$.

Question 1.2 : Calculer la loi de Y_1.

Question 1.3 : Lorsque l'on possède $i - 1$ vignettes distinctes, et que l'on achète une nouvelle vignette : quelle est la probabilité p_i pour que cette nouvelle vignette ne soit pas déjà dans la collection ? (exprimer cette probabilité en fonction de i et de M).

Question 1.4 : Donner, en fonction de p_i, la loi de Y_i.

Question 1.5 : Les variables Y_i, pour $1 \leqslant i \leqslant M$, sont-elles indépendantes ?

Question 1.6 : En déduire, en fonction de i et de M, l'espérance de Y_i.

Question 1.7 : Exprimer le coût total T (nombre de vignettes achetées) en fonction des Y_i.

Question 1.8 : En déduire une expression pour l'espérance de T et, si possible, un équivalent asymptotique $\mathcal{O}()$ de cette dernière.

1.2 Simulation

Question 1.9 : Proposer une fonction `vignette(M)` qui utilise un générateur pseudo-aléatoire uniforme sur $]0, 1[$ noté `random()` pour simuler le tirage de la prochaine vignette X_i (un entier entre 1 et M).

Question 1.10 : Utiliser les nombres pseudo-aléatoires fournis en annexe pour donner les 5 premières images obtenues par votre fonction `vignette`.

Question 1.11 : Écrire une fonction `album(M)` en R qui simule une collection de M vignettes et renvoie le nombre d'images qu'il aura fallu acquérir pour compléter la collection.

1.3 Tables de hachage (bonus)

Question 1.12 : Expliquer le lien entre l'exercice ci-dessus et les tables de hachage.

Annexe :Réels (float) pseudo-aléatoires

Des appels successifs à un générateur pseudo-aléatoire donnent les nombres suivants (lecture en *ligne*, de gauche à droite).

```
0.327010 0.057128 0.994553 0.214157 0.825574 0.795653 0.068671 0.667426 0.755272 0.461837
0.788446 0.411315 0.905150 0.781532 0.794132 0.095405 0.647180 0.548351 0.271737 0.638842
0.723094 0.464648 0.332958 0.886690 0.764691 0.604677 0.390348 0.213932 0.135788 0.528952
0.155550 0.462798 0.586080 0.150103 0.676956 0.411654 0.945757 0.745627 0.079080 0.701028
0.207464 0.867526 0.112343 0.112614 0.649058 0.906475 0.208019 0.296238 0.454826 0.479756
0.935080 0.177919 0.944403 0.268038 0.064609 0.709094 0.872715 0.454958 0.923026 0.008503
0.983909 0.078576 0.471301 0.569990 0.228680 0.148257 0.981644 0.174436 0.893884 0.060724
0.875465 0.101348 0.928250 0.987808 0.213961 0.577309 0.894283 0.421980 0.873546 0.349109
0.901736 0.808627 0.527028 0.846139 0.076665 0.591637 0.555233 0.949380 0.046595 0.478259
0.957883 0.030504 0.556835 0.429184 0.600494 0.785515 0.577441 0.582138 0.959951 0.471325
```