Math 701 exercices – Probabilités

Nicolas Gast

8 décembre 2020

Exercise 1 Règle de Bayes

Selon https://www.caminteresse.fr/sciences/quel-est-le-pourcentage-de-daltoniens-en-france-119268 8% des hommes sont daltoniens contre 0.5% des femmes. Selon https://www.insee.fr/fr/statistiques/2381474 il y a $34.4M$ de femmes et $32.4M$ d'hommes en France.

1. On choisit un daltonien au hasard dans la population. Quel est la probabilité que ce soit un homme ?

Exercise 2 Intervales de confiance

Soit $X_1, X_2, X_3, \dots$ une suite de variables aléatoires indépendantes et de même loi. On définit $S_n = \frac{1}{n}\sum_{i=1}^n X_i$ et $\sigma_n = \sqrt{\frac{1}{n}\sum_{i=1}^n (X_i - S_n)^2}$.

1. Un simulateur nous montre que $S_{100} = 43.5$ et $\sigma_n = 5.1$. Calculer un intervalle de confiance pour $\mathbb{E}[X_1]$.
2. Combien de simulation faudrait-il faire pour que l'intervalle de confiance sur $\mathbb{E}[X_1]$ soit plus petit que 0.1.

Exercise 3 Simulation et ruine d'un joueur

On cherche à simuler un joueur qui démare avec 100 euros. À chaque tour il paie 1 euro. Il gagne 2 euros avec probabilité p et ne gagne rien avec probabilité $1-p$. Le joueur s'arrête s'il est ruiné ou s'il atteint 200 euros. On se donne le code suivante qui simule son argent au cours du temps :

```python
def simulation_joueur(p):
    """
    Simule un joueur jusqu'a sa ruine (et au maximum Tmax pas de temps).

    Retourne: son argent au cours du temps.
    """
    x = 10
    X = [x]
    while(x>0 and x<1000):
        if (np.random.rand()<p):
            x += 1
        else:
            x -= 1
        X.append(x)
    return(X)
# Par exemple pour
x = simulation_joueur(0.1,1000)[0]
```

En utilisant le simulateur, calculer une estimation (et un intervale de confiance) du :

1. Temps moyen que le joueur jour (pour $p \in \{0.3, 0.5, 0.6\}$)
2. Argent moyen avec lequel il ressort (pour $p \in \{0.3, 0.5, 0.6\}$).

Exercise 4 Rejet

1. Écrire un algorithme qui génère un point uniformément sur l'intérieur de l'ellipse définie par $x^2 + 2y^2 \leq 1$.
   a) Quelle est sa complexité moyenne ? (c'est-à-dire combien de tirages doit-il faire pour tirer une paire de point).
   b) Faire tourner votre algorithme et générer un plot 2D pour illustrer votre algorithme.
2. Écrire un algorithme qui génère un point uniformément sur l'ellipse définie par $x^2 + 2y^2 = 1$.

Exercise 5 Le collecteur : approche théorique

Vous collectionnez des autocollants. Il y en a n différents. Tous les jours, vous achetez un nouveau paquet qui contient un autocollant (qui est tiré uniformément au hasard parmi les n autocollants). On chercher à calculer le temps qu'il faut pour collectionner les n autocollants distincts. Pour $i \in \{1 \dots n\}$, on note T_i le temps qu'il vous faut pour collectionner T_i compons différents, avec la convention que $T_0 = 0$.

1. Que vaut T_1 ?
2. On note $X_i = T_i - T_{i-1}$.
   a) Quelle est la loi de X_i ? (indice : c'est une loi géométrique, mais de quelle paramètre ?)
   b) Calculer $\mathbb{E}[X_i]$.
   c) Les variables X_i sont elles indépendantes ?
3. Calculer l'espérance de T_n et montrer que $\mathbb{E}[T_n] = \sum_{k=1}^{n} \frac{n}{k} \approx n \log n$.
4. Calculer la variance de T_n. En déduire que T_n est proche de sa moyenne.

Exercise 6 Le cllecteur : approche par simulation

Vous collectionnez des autocollants. Il y en a n différents. Tous les jours, vous achetez un nouveau paquet qui contient un autocollant. Cette fois on suppose que la probabilité d'obtenir le coupon i est p_i.

On chercher à calculer le temps qu'il faut pour obtenir k collections de n autocollants distincts.

1. Programmer un simulateur qui prend en entrée un vecteur `p` (qui est une distribution de probabilité sur $\{1 \dots n\}$) et un entier `k` et qui calcule le temps qu'il faut pour obtenir k collections de n coupons.
2. On considère la distribution uniforme : $p_i = 1/n$ pour tout i.
   a) Calculer le temps moyen pour obtenir une collection de 100 autocollants distincts.
   b) Comparer avec le temps moyen pour obtenir 2, 3 ou 5 colletions.
   c) Pouvez vous expliquer pourquoi il vaut mieux se mettre à plusieurs pour collectionner ?
3. On suppose maintenant que $p_i = 2i/(n(n+1))$.
   a) Expliquer pourquio p est une distribution de probabilité.
   b) Le temps moyen pour obtenir une collection complète est-il plus grand ou plus petit que dans le cas uniforme ?