

## T.D. V - Estimation

### I - Construction d'estimateurs

**Exercice 1.** Soit  $n \in \mathbb{N}^*$ ,  $p \in [0, 1]$  et  $X$  suivant une loi de Bernoulli de paramètre  $p$ . On note  $(X_1, \dots, X_n)$  un  $n$ -échantillon de  $X$  et  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ .

1. Déterminer la loi de la variable aléatoire  $Y_n = n\bar{X}_n$ .
2. Exprimer  $\mathbf{E}[Y_n]$ ,  $\mathbf{V}(Y_n)$  et  $\mathbf{E}[Y_n^2]$  en fonction de  $n$  et de  $p$ .
3.  $\bar{X}_n^2$  est-il un estimateur sans biais de  $p^2$ ? Sinon, proposer un estimateur sans biais de  $p^2$ .

**Exercice 2. (Estimation de la variance)** Soit  $X$  une variable aléatoire qui admet une espérance  $m$  et une variance  $\sigma^2$ . Soit  $n \in \mathbb{N}^*$  et  $(X_1, \dots, X_n)$  un  $n$ -échantillon de  $X$ . On pose  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  et  $s_n = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ .

1. Montrer que  $\bar{X}_n$  est un estimateur sans biais de  $m$ .
2. Montrer que  $\mathbf{V}(\bar{X}_n) = \frac{\sigma^2}{n}$ .
3. Montrer que  $s_n = \frac{1}{n} \left( \sum_{i=1}^n X_i^2 \right) - \bar{X}_n^2$  puis que  $\mathbf{E}[s_n] = \frac{n-1}{n} \sigma^2$ .
4. La variable aléatoire  $s_n$  est-elle un estimateur sans biais de  $\sigma^2$ ?
5. Montrer que  $S_n = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  est un estimateur sans biais de  $\sigma^2$ .

**Exercice 3.** Soit  $N \geq 2$ . Une urne contient  $N$  boules numérotées de 1 à  $N$ . On effectue  $n$  tirages indépendants avec remise. Pour tout  $k \in \llbracket 1, n \rrbracket$ , on note  $X_k$  le numéro de la boule tirée au  $k^{\text{e}}$  tirage. Enfin, on pose :

$$M_n = \max \{X_1, \dots, X_n\}.$$

1. Montrer que, pour tout  $i \in \llbracket 1, N \rrbracket$ ,

$$\mathbf{P}([M_n \leq i]) = \left(\frac{i}{N}\right)^n.$$

2. Déterminer  $\mathbf{P}([M_n = N])$  puis  $\lim_{n \rightarrow +\infty} \mathbf{P}([M_n = N])$ .

3. Montrer que si  $Y$  est une variable aléatoire à valeurs dans  $\llbracket 1, N \rrbracket$ , alors  $\mathbf{E}[Y] = \sum_{i=1}^N \mathbf{P}([Y \geq i])$ .

4. En déduire que  $\mathbf{E}[M_n] = N - \sum_{k=0}^{N-1} \left(\frac{k}{N}\right)^n$ .

5. Montrer que  $N - N \left(1 - \frac{1}{N}\right)^n \leq \mathbf{E}[M_n] \leq N$  puis déterminer la limite de  $\mathbf{E}[M_n]$  lorsque  $n$  tend vers l'infini. L'estimateur  $M_n$  est un estimateur *asymptotiquement sans biais* de  $N$ .

### II - Comparaison d'estimateurs

**Exercice 4.** Soit  $\lambda > 0$  et  $X$  une variable aléatoire suivant une loi de Poisson de paramètre  $\lambda$ . Pour estimer ce paramètre, on considère un  $n$ -échantillon  $(X_1, \dots, X_n)$  de  $X$ . On considère les deux estimateurs de  $\lambda$  suivants :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \text{ et } T_n = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

1. Ces estimateurs sont-ils biaisés?

2. Calculer le risque quadratique de  $\bar{X}_n$ .
3. On admet que le risque quadratique de  $T_n$  vaut  $\frac{\lambda}{n} + \frac{2\lambda^2}{n-1}$ . Lequel de ces deux estimateurs vous semble préférable ?

**Exercice 5.** Soit  $\theta \neq 0$ ,  $X$  une variable aléatoire discrète telle que  $\mathbf{E}[X] = \theta$  et  $\mathbf{V}(X) = 1$ . On dispose d'un  $n$ -échantillon  $(X_1, \dots, X_n)$  de  $X$  et on pose  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ .

1. Montrer que  $\bar{X}_n$  est un estimateur sans biais de  $\theta$  et calculer son risque quadratique.
2. Soit  $(\alpha_1, \dots, \alpha_n) \in (\mathbb{R}^*)^n$ . On note  $Y_n = \sum_{i=1}^n \alpha_i X_i$ . Donner une condition nécessaire et suffisante pour que  $Y_n$  soit un estimateur sans biais de  $\theta$ . On suppose ensuite cette condition vérifiée.
3. Calculer  $\text{Cov}(\bar{X}_n, Y_n)$ . En déduire que  $\mathbf{V}(\bar{X}_n) \leq \mathbf{V}(Y_n)$ . Que dire en cas d'égalité ?
4. Interpréter les résultats obtenus.