TD, Feuille 6 - Intervalles de confiance

EXERCICE 1. — Pile ou Face

On joue n fois à pile ou face, avec une probabilité de succès (pile) $p \in]0, 1[$ inconnu, et on observe le nombre de pile obtenus.

1. Donner un modèle pour cette expérience.
2. Proposer un estimateur de p basé sur la méthode des moments ou du maximum de vraisemblance. On le notera $\hat{p}_n$.
3. Montrer que $\hat{p}$ est consistant.
4. Montrer que $\hat{p}$ est asymptotiquement normal.
5. En déduire un intervalle de confiance asymptotique bilatère sur p (au niveau de confiance 98%).
6. Montrer que $E_p((\hat{p} - p)^2) \leq 1/(4n)$.
7. En déduire un intervalle de confiance non-asymptotique bilatère sur p (au niveau de confiance 98%).
8. Application numérique: combien de lancers faut-il prévoir pour obtenir une précision de 1% dans les intervalles construits précédemment? On admettra que le quantile d'ordre 99% d'une loi normale standard vaut à peu près 2,33. Les comparer sur cette base.

EXERCICE 2. — Pile ou Face - Bis

On joue une infinité de fois à pile ou face, avec une probabilité de succès (pile) $p \in]0, 1[$, et on observe le nombre d'essais nécessaires à l'obtention d'un pile. On répète cette expérience n fois.

1. Donner un modèle pour cette expérience.
2. Proposer un estimateur de p basé sur la méthode des moments ou du maximum de vraisemblance. On le notera $\hat{p}_n$.
3. Montrer que $\hat{p}$ est consistant.
4. Montrer que $\hat{p}$ est asymptotiquement normal.
5. En déduire un intervalle unilatère de niveau de confiance asymptotique 97% de la forme $[0, \hat{p}_-$ sur p .

EXERCICE 3. — Capture

On pêche n poissons (sans remise) dans un étang comprenant N_0 carpes et N_1 brochets (ces deux quantités étant inconnues, toutes deux supposées plus grandes que n). On observe le nombre de carpes pêchées, et on suppose le nombre total de poissons N connu.

1. Donner un modèle pour cette expérience.
2. Proposer un estimateur de N_0 basé sur la méthode des moments.

EXERCICE 4. — Capture - Recapture

On a déjà marqué n_0 poissons toutes espèces confondues au cours d'une pêche précédente, que l'on a ensuite relâchés. On pêche successivement n poissons avec remise dans l'étang qui est censé contenir N poissons (inconnu), et on observe le nombre de poissons déjà vus au cours de la première pêche.

1. Donner un modèle pour cette expérience.
2. Proposer un estimateur de N basé sur la méthode des moments, on le notera $\hat{N}_n$.
3. Montrer que $\hat{N}_n$ est consistant.
4. Montrer que $\hat{N}_n$ est asymptotiquement normal.
5. En déduire un intervalle de confiance de niveau asymptotique 95% sur N , du type $[\hat{N}_-; +\infty[$.

EXERCICE 5. — Tanks allemands

Au cours de la seconde guerre mondiale (du moins au début), les allemands numérotaient leurs tanks produit pendant un mois de 1 à N (inconnu). Un espion se place sur une route près de l'usine, et observe les numéros des n tanks qui passent devant lui (un même tank peut passer plusieurs fois).

1. Proposer un modèle pour cette expérience.
2. Donner un estimateur $\hat{N}_{n,1}$ par la méthode des moments.
3. Montrer que $\hat{N}_{n,1}$ est consistant.
4. Montrer que $\hat{N}_{n,1}$ est asymptotiquement normal.
5. En utilisant la méthode Delta, donner un intervalle de niveau de confiance asymptotique 98% sur $\log(N)$ de la forme $[0; \hat{L}_+]$.
6. En déduire un intervalle de niveau de confiance asymptotique 98% sur N de la forme $[1; \hat{N}_+]$.
7. Donner un estimateur $\hat{N}_{n,2}$ par la méthode du maximum de vraisemblance.
8. Montrer que $\hat{N}_{n,2}$ est consistant.
9. Montrer que, pour toute suite positive $r_n \rightarrow +\infty$, $r_n(N - \hat{N}_{n,2}) \rightsquigarrow 0$.
10. Construire un intervalle de niveau de confiance 98% sur N , de la forme $[1, \hat{N}_+[$, basé sur $\hat{N}_{n,2}$.

EXERCICE 6. — Stylométrie de discours

Un chercheur en sciences sociales s'intéresse aux occurrences du mot 'sécurité' dans les discours de J. Chirac. Ce nombre d'occurrences dans un discours de 5000 mots est assez bien représenté par une loi de Poisson de paramètre λ . Le chercheur compte le nombre de fois où 'sécurité' apparaît dans une succession de n discours composés chacun de 5000 mots.

1. Proposer un modèle pour cette expérience.
2. Donner un estimateur $\hat{\lambda}_n$ de λ basé sur la méthode des moments ou du maximum de vraisemblance.
3. Montrer que $\hat{\lambda}_n$ est consistant.
4. Montrer que $\hat{\lambda}_n$ est asymptotiquement normal.
5. En déduire que $\sqrt{n}(\sqrt{\hat{\lambda}_n} - \sqrt{\lambda}) \rightsquigarrow \mathcal{N}(0, 1/4)$.
6. Déduire de la question précédente un intervalle de confiance asymptotique, de niveau 96%, bilatère, sur $\sqrt{\lambda}$.
7. En déduire un intervalle de confiance bilatère et asymptotique sur λ , au niveau 96%. On le notera I_1 .

8. Montrer que $E_\lambda((\hat{\lambda}_n - \lambda)^2) = \lambda/n$.
9. En déduire un intervalle de confiance, bilatère, de niveau 96% sur λ . On le notera I_2 .

EXERCICE 7. — Temps d'attente devant les ascenseurs

Pour un usager, le temps que met un ascenseur à arriver est bien modélisé par une loi exponentielle de paramètre λ . Un usager se trouve devant trois ascenseurs, supposés indépendants, et observe le temps que met le plus rapide à arriver. Il répète l'expérience n fois.

1. Proposer un modèle pour cette expérience.
2. Donner un estimateur de λ par la méthode du maximum de vraisemblance ou des moments, on le notera $\hat{\lambda}_n$.
3. Montrer que $\hat{\lambda}_n$ est consistant.
4. Montrer que $\hat{\lambda}_n$ est asymptotiquement normal.
5. Construire un intervalle asymptotique de niveau de confiance 90% unilatère (de type $] - \infty, \hat{\lambda}_+]$).

EXERCICE 8. — Notes

La note d'un étudiant à un examen dans une matière est modélisée par une loi normale de moyenne μ et variance $v = \sigma^2$, les 2 sont supposées inconnues. On suppose que les différentes matières sont indépendantes, qu'il y a n matières, et que l'étudiant observe ses notes.

1. Donner un modèle pour cette expérience.
2. Proposer un estimateur de μ et de v par la méthode des moments ou la méthode du maximum de vraisemblance (attention il faut aussi optimiser en v). On les notera $\hat{\mu}$ et $\hat{v}$.
3. Montrer que $\hat{\mu}$ est consistant.
4. Écrire $\hat{v}$ en fonction de $(\sum_{i=1}^n (X_i - \mu)^2)/n$ et $(\hat{\mu} - \mu)^2$.
5. En déduire que $\hat{v}$ est consistant.
6. Montrer que $\hat{\mu}$ est **non-asymptotiquement** normal.
7. Montrer que $\sqrt{n}(\bar{X}_n - \mu)^2 \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0$ (on pourra utiliser l'inégalité de Markov).
8. Montrer que $\hat{v}$ asymptotiquement normal. On admettra que si $X \sim \mathcal{N}(0, 1)$, $\text{Var}(X^2) = 2$.
9. On suppose que v est connue. Donner un intervalle bilatère de niveau de confiance 96% sur μ .
10. Même question mais en supposant v inconnu et pour un intervalle de niveau de confiance asymptotique 96% sur μ .
11. On suppose que μ et v sont inconnus. Donner un intervalle de niveau de confiance asymptotique 80% pour v de type $[\hat{v}_-, +\infty[$.

EXERCICE 9. — Age de l'univers

Des astrologues observent l'âge de n corps célestes arrivés sur Terre, l'âge d'un corps céleste étant modélisé par une loi uniforme entre 0 et A , où A est l'âge de l'univers (inconnu).

1. Donner un modèle pour cette expérience.

2. Donner un estimateur de A par la méthode des moments. On le notera $\hat{A}_{n,1}$.
3. Montrer que $\hat{A}_{n,1}$ est consistant.
4. Montrer que $\hat{A}_{n,1}$ est asymptotiquement normal.
5. Montrer que $E_A((\hat{A}_{n,1} - A)^2) = A^2/(3n)$.
6. Donner un estimateur de A par la méthode du maximum de vraisemblance. On le notera $\hat{A}_{n,2}$.
7. Montrer que $\hat{A}_{n,2}$ est consistant.
8. Montrer que $n(A - \hat{A}_{n,2}) \rightsquigarrow \mathcal{E}(a)$, où a est un paramètre à déterminer.
9. Montrer que $E_A((\hat{A}_{n,2} - A)^2) = 2A^2/((n+1)(n+2))$. Comparer avec le risque quadratique de $\hat{A}_{n,1}$.
10. Donner un intervalle de niveau de confiance 95% sur A basé sur $\hat{A}_{n,1}$. On le notera I_1 .
11. Donner un intervalle de niveau de confiance 95% sur A basé sur $\hat{A}_{n,2}$. On le notera I_2 .
12. Comparer I_1 et I_2 sur la base de leurs longueurs lorsque $n \rightarrow +\infty$.

EXERCICE 10. — Centre d'appel (*) - Loi Gamma

Le temps mis par un opérateur d'un centre d'appel à traiter une demande est modélisé par une loi exponentielle de paramètre λ (inconnu). Durant n jours, un superviseur observe le temps mis par un employé à répondre à 10 appels quotidiens.

1. Donner un modèle pour cette expérience.
2. Donner un estimateur de λ par la méthode du maximum de vraisemblance ou des moments. On le notera $\hat{\lambda}_n$.
3. Montrer que $\hat{\lambda}_n$ est consistant.
4. Montrer que $\hat{\lambda}_n$ est asymptotiquement normal.

EXERCICE 11. — Fourmis et fourmilières (*) - Mélange Gaussien

On suppose que dans une clairière vivent trois espèces de fourmis, chacune ayant sa fourmilière, et que, sachant le type de fourmi, la position de la fourmi suit une loi normale centrée en μ_j , de variance σ_j^2 (où j est le numéro de l'espèce). Les proportions de chaque espèce π_j sont supposées connues. Un entomologue collecte n fourmis (avec remise) dans la clairière, et observe leurs espèces et positions.

1. Donner un modèle pour cette expérience.
2. Pour chaque j , donner un estimateur de μ_j et σ_j^2 (par la méthode des moments ou du max de vraisemblance). On les notera $\hat{\mu}_j$ et $\hat{\sigma}_j^2$.
3. Pour chaque j , montrer que $\hat{\mu}_j$ et $\hat{\sigma}_j^2$ sont consistants.
4. Pour chaque j , montrer que $\hat{\mu}_j$ est **non-asymptotiquement** normal, et $\hat{\sigma}_j^2$ est asymptotiquement normal.

On suppose maintenant les proportions du mélange π_j inconnues.

5. Donner le nouveau modèle.

6. Donner des estimateurs des π_j par la méthode de votre choix (moments ou max de vraisemblance). On les notera $\hat{\pi}_{j,n}$.
7. Montrer que les $\hat{\pi}_{j,n}$ sont consistants.
8. Montrer que les $\hat{\pi}_{j,n}$ sont asymptotiquement normaux.
9. Donner de nouveaux estimateurs pour les μ_j (notés $\hat{\mu}_{j,2}$).
10. Montrer que les $\hat{\mu}_{j,2}$ sont consistants.
11. Montrer que les $\hat{\mu}_{j,2}$ sont asymptotiquement normaux.

EXERCICE 12. — Filtre à Poisson (*)

Un probabiliste commence par tirer un entier N suivant une loi de Poisson de paramètre λ (inconnu), puis réalise une expérience de N piles ou face avec probabilité de succès $p \in]0, 1[$ connu. Le statisticien n'observe que le nombre de succès obtenu, et cette expérience se répète n fois.

1. Donner un modèle pour cette expérience.
2. Donner un estimateur de λ par la méthode des moments ou du maximum de vraisemblance. On le notera $\hat{\lambda}_n$.
3. Montrer que $\hat{\lambda}_n$ est consistant.
4. Montrer que $\hat{\lambda}_n$ est asymptotiquement normal.
5. Montrer que $E_\lambda((\hat{\lambda}_n - \lambda)^2) = \lambda/(np)$.

EXERCICE 13. — Logs d'uniformes (*)

Un probabiliste commence par tirer un n -échantillon de loi uniforme sur $]0, 1[$, calcule leur maximum, en prend le log, additionne un θ inconnu et envoie le résultat à son collègue statisticien.

1. Donner un modèle pour cette expérience.
2. Donner un estimateur de θ par la méthode du maximum de vraisemblance. On le notera $\hat{\theta}_n$.
3. Montrer que $\hat{\theta}_n$ est consistant.
4. Donner la loi de $\theta - \hat{\theta}_n$.

EXERCICE 14. — Répartition du patrimoine (*)

Une modélisation historique de la répartition des richesses consiste à supposer que le patrimoine d'un individu tiré au hasard suit une loi de Pareto, de densité donnée par

$$f(x) = (r - 1)x^{-r} \mathbb{1}_{x \geq 1},$$

pour un $r > 2$ inconnu. Un chercheur en sciences sociales observe le patrimoine de n individus tirés au hasard (avec remise).

1. Donner un modèle pour cette expérience.
2. Donner un estimateur de r par la méthode du maximum de vraisemblance. On le notera $\hat{r}_n$.
3. Montrer que $\hat{r}_n$.
4. Donner une condition sur r pour que $\hat{r}_n$ soit asymptotiquement normal.

EXERCICE 15. — Pile ou face - Fin (*)

Un statisticien tire une infinité de fois à pile ou face, avec proba p de tomber sur pile, et observe le nombre de face qu'il a fallu tirer jusqu'à obtenir n piles (où $n \geq 1$ est connu).

1. Donner un modèle pour cette expérience.
2. Donner un estimateur pour p , par la méthode des moments ou du maximum de vraisemblance. On le notera $\hat{p}_n$.
3. Montrer que $\hat{p}_n$ est consistant.
4. Montrer que $\hat{p}_n$ est asymptotiquement normal.