TD, Feuille 7 - Intervalles de confiance et tests

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.
9. On cherche à prouver que la pièce est truquée, et tombe sur pile avec probabilité strictement plus grande que 1/2. Donner les hypothèses de test associées.
10. Construire un test d'erreur de première espèce 5%, basé sur $\hat{p}_n$, pour en avoir le coeur net.
11. **Application**: On observe une proportion de piles de 3/4 sur $n = 50$ lancers. Que pouvez-vous en conclure? Un tableau standard indique que $\mathbb{P}(\mathcal{B}(50, 1/2) \leq 30) \approx 0.9405$ et $\mathbb{P}(\mathcal{B}(50, 1/2) \leq 31) \approx 0.9675$.

Solution 1. —

1. En notant X_i le résultat du i-ème tirage (avec par convention 1 pour pile, 0 pour face), on peut supposer les X_i i.i.d., de loi $\mathcal{B}(p)$, pour p inconnu. On observe $S = \sum_{i=1}^n X_i$ qui suit alors une loi $\mathcal{B}(n, p)$. Le modèle s'écrit alors
$$\left([\![0, n]\!], (\mathcal{B}(n, p))_{p \in]0,1[} \right).$$
La densité s'écrit, pour $x \in [\![0, n]\!]$, $f_p(x) = \binom{n}{x} p^x (1-p)^{n-x}$.
2. Par la méthode du maximum de vraisemblance. Pour $x \in [\![0, n]\!]$ (dans l'espace d'observation), la vraisemblance s'écrit
$$V_x(p) = \binom{n}{x} p^x (1-p)^{n-x}.$$

Pour $x = 0$, le maximum est atteint en $p = 0$, pour $x = n$, il est atteint en $p = 1$. Sinon $V_x(p) > 0$ pour tout $p \in [0,1]$, et on peut se contenter de maximiser

$$\ell_x(p) = \log\left(\binom{n}{x}\right) + x\log(p) + (n-x)\log(1-p).$$

Cette fonction étant strictement concave, annuler ℓ'_x fournira un maximum. On a alors

$$\begin{aligned}\ell'_x(p) = 0 \quad &\Leftrightarrow \quad \frac{x}{p} - \frac{n-x}{1-p} = 0\\ &\Leftrightarrow \quad p = \frac{x}{n}.\end{aligned}$$

Un estimateur par maximum de vraisemblance est alors donné par $\hat{p}_n = \frac{S}{n}$.

3. Comme $S = \sum_{i=1}^n X_i$, et $\mathbb{E}(|X_1|) \leq 1$, la loi des grand nombres donne

$$\hat{p}_n \xrightarrow[n\to+\infty]{\mathbb{P}} p,$$

donc $\hat{p}_n$ est consistant.

4. Comme $\mathbb{E}(X_1^2) \leq 1$, la loi des grands nombres donne

$$\sqrt{n}(\hat{p}_n - p) \rightsquigarrow \mathcal{N}(0, \mathrm{Var}_p(X_1)),$$

avec $\mathrm{Var}_p(X_1) = p(1-p)$. On en déduit que $\hat{p}$ est asymptotiquement normal.

5. D'après la question précédente, on a

$$\sqrt{n}\left(\frac{\hat{p}_n - p}{\sqrt{p(1-p)}}\right) \rightsquigarrow \mathcal{N}(0,1).$$

Pour en déduire un intervalle de confiance, on utilise la consistance de $\hat{p}_n$ et le Lemme de Slutsky, qui garantissent que

$$\sqrt{n}\left(\frac{\hat{p}_n - p}{\sqrt{\hat{p}_n(1-\hat{p}_n)}}\right) \rightsquigarrow \mathcal{N}(0,1).$$

Soit q le quantile d'ordre 99% d'une loi $\mathcal{N}(0,1)$. On a alors

$$P_p\left(\left\{\left|\sqrt{n}\left(\frac{\hat{p}_n - p}{\sqrt{\hat{p}_n(1-\hat{p}_n)}}\right)\right|\right\}\right) \xrightarrow[n\to+\infty]{} \mathbb{P}\left(|\mathcal{N}(0,1)| \leq q\right) = 98\%.$$

Or,

$$\left|\sqrt{n}\left(\frac{\hat{p}_n - p}{\sqrt{\hat{p}_n(1-\hat{p}_n)}}\right)\right| \leq q \quad \Leftrightarrow \quad \hat{p}_n - \frac{\sqrt{\hat{p}_n(1-\hat{p}_n)}}{\sqrt{n}}q \leq p \leq \hat{p}_n + \frac{\sqrt{\hat{p}_n(1-\hat{p}_n)}}{\sqrt{n}}q.$$

On en déduit que

$$\left[\hat{p}_n - \frac{\sqrt{\hat{p}_n(1-\hat{p}_n)}}{\sqrt{n}}q ; \hat{p}_n + \frac{\sqrt{\hat{p}_n(1-\hat{p}_n)}}{\sqrt{n}}q\right]$$

est un intervalle de niveau de confiance asymptotique 98%.

6. Comme $E_p(\hat{p}_n) = p$, on a
$$E_p((\hat{p}_n - p)^2) = \mathrm{Var}_p(\hat{p}_n) = \frac{\mathrm{Var}_p(X_1)}{n} = \frac{p(1-p)}{n} \leq \frac{1}{4n}.$$
7. On utilise ici l'inégalité de Bienaymé-Cebicev. Si $t > 0$,
$$P_p\left(\{|\hat{p}_n - p| > t\} \leq \frac{E_p((\hat{p}_n - p)^2)}{t^2}\right) \leq \frac{1}{4nt^2}.$$
On pose alors $t_{2\%} = 1/(2\sqrt{n2\%})$, de sorte que
$$P_p\left(\{|\hat{p}_n - p| > t_{2\%}\}\right) \leq 2\%.$$
Or,
$$|\hat{p}_n - p| \leq t_{2\%} \quad \Leftrightarrow \quad \hat{p}_n - 1/(2\sqrt{n2\%}) \leq p \leq \hat{p}_n + 1/(2\sqrt{n2\%}).$$
On en déduit que
$$\left[\hat{p}_n - 1/(2\sqrt{n2\%}); \hat{p}_n + 1/(2\sqrt{n2\%})\right]$$
est un intervalle de niveau de confiance 98%.
8. La précision de l'intervalle étant donné par sa longueur. Dans le deuxième cas, cette longueur vaut $1/\sqrt{n2\%}$, on a alors l'équation
$$\begin{aligned} 1/\sqrt{n2\%} \leq 10^{-2} \quad &\Leftrightarrow 2\%n \geq 10^4 \\ &\Leftrightarrow n \geq 500000. \end{aligned}$$
Dans le premier cas, la précision de l'intervalle vaut
$$2q\sqrt{\hat{p}_n(1-\hat{p}_n)}/\sqrt{n} \leq 2.33/\sqrt{n},$$
en utilisant $\hat{p}_n(1-\hat{p}_n) \leq 1/4$ et $q \approx 2.33$. L'équation devient alors
$$2.33/\sqrt{n} \leq 10^{-2} \quad \Leftrightarrow \quad n \geq 10^4 \times (2.33)^2.$$
Prendre $n \geq 57600$ suffit, ce qui fait quand même de l'ordre de 10 fois moins que pour l'intervalle précédent. Le premier intervalle (asymptotique) est donc plus précis.
9. On cherche à montrer que $p > 1/2$. On choisit donc les hypothèses
$$\begin{cases} H_0 & : \quad p = 1/2, \\ H_1 & : \quad p > 1/2. \end{cases}$$
10. Sous H_1 on s'attend à ce que $\hat{p}_n$ soit grand. On choisit donc une région de rejet de la forme $[t_{5\%}, +\infty[$, où $t_{5\%}$ doit vérifier.
$$P_{1/2}(\hat{p}_n \geq t_{5\%}) \leq 5\%.$$
Or, pour $p = 1/2$, $n\hat{p}_n \sim \mathcal{B}(n, 1/2)$. On a alors, pour $t \in \mathbb{N}$,
$$\begin{aligned} P_{1/2}(\hat{p}_n \geq t) \quad &\Leftrightarrow \quad P_{1/2}(n\hat{p}_n \geq nt) \\ &\Leftrightarrow \mathbb{P}\left(\mathcal{B}(n, 1/2) \geq nt\right). \end{aligned}$$
En posant $nt_{5\%} - 1 = q$, où q est le quantile d'ordre 95% d'une loi $\mathcal{B}(n, 1/2)$, on a bien
$$P_{1/2}(\hat{p}_n \geq t_{5\%}) = 1 - \mathbb{P}\left(\mathcal{B}(n, 1/2) \leq q\right) \leq 5\%.$$
Donc $T = \mathbb{1}_{\hat{p}_n \geq (q+1)/n}$ est un test de niveau 5%.

11. Comme $\mathbb{P}(\mathcal{B}(50, 1/2) \leq 30) \approx 0.9405$ et $\mathbb{P}(\mathcal{B}(50, 1/2) \leq 31) \approx 0.9675$, on en déduit que $q = 31$. On a alors $(q+1)/50 = 32/50 \leq 3/4 = \hat{p}_{50}$. On rejette donc l'hypothèse nulle, et on est sûrs à 95% que la pièce est truquée.

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.
6. On cherche encore à prouver que la pièce est truquée et tombe sur pile avec probabilité strictement plus petite que $1/2$. Donner les hypothèses de ce test.
7. Construire un test de niveau asymptotique 3%, basé sur $\hat{p}_n$, pour ces hypothèses.
8. **Application numérique**: On répète l'expérience $n = 200$ fois, et on observe un nombre moyen d'essais pour avoir un pile de $2,5$. Que peut-on en conclure (on pourra utiliser la table de quantiles donnée en annexe)?

SOLUTION 2. —

1. En notant Z_i le résultat du i-ème tirage comme auparavant, on observe cette fois ci $X = \min\{i \in \mathbb{N} \mid Z_i = 1\}$. Un rapide calcul (ou la connaissance du cours) donne $X \sim \mathcal{G}(p)$ (loi géométrique de paramètre p). On répète cette expérience n fois de manière indépendante (et on note X_i le résultat de la i-ème répétition). Le modèle est alors
$$\left((\mathbb{N}^*)^n, \left(\mathcal{G}(p)^{\otimes n}\right)_{p\in]0,1[}\right).$$
La densité s'écrit, pour $x_1, \ldots, x_n \geq 1$, $f_p(x_{1:n}) = \prod_{i=1}^n p(1-p)^{x_i - 1}$.
2. On utilise la méthode des moments. Comme $E_p(X) = \frac{1}{p}$, un estimateur par moment $\hat{p}_n$ vérifie
$$\bar{X}_n = \frac{1}{\hat{p}_n},$$
ou encore $\hat{p}_n = 1/\bar{X}_n$.
3. Comme $E_p(|X_1|) = 1/p < +\infty$, la loi des grands nombres donne
$$\bar{X}_n \xrightarrow[n\to+\infty]{\mathbb{P}} \frac{1}{p}.$$
Comme $x \mapsto 1/x$ est continue, $\hat{p}_n \xrightarrow[n\to+\infty]{\mathbb{P}} p$, et $\hat{p}_n$ est donc consistant.

4. Comme $E_p(X_1^2) < +\infty$, le théorème central limite donne
$$\sqrt{n}(\bar{X}_n - 1/p) \rightsquigarrow \mathcal{N}(0, \mathrm{Var}_p(X_1)),$$
avec $\mathrm{Var}_p(X_1) = (1-p)/p^2$. Comme $g : x \mapsto 1/x$ est dérivable en p, la méthode Δ donne alors
$$\sqrt{n}(\hat{p}_n - p) \rightsquigarrow g'(1/p)\mathcal{N}(0, \mathrm{Var}_p(X_1)) \sim \mathcal{N}(0, p^2(1-p)).$$
$\hat{p}_n$ est donc asymptotiquement normal.

5. D'après la question précédente,
$$\sqrt{n}\left(\frac{\hat{p}_n - p}{p\sqrt{1-p}}\right) \rightsquigarrow \mathcal{N}(0,1).$$
Soit q le quantile d'ordre 97% d'une loi $\mathcal{N}(0,1)$. On déduit alors que
$$P_p\left(\left\{\sqrt{n}\left(\frac{p - \hat{p}_n}{p\sqrt{1-p}}\right) \leq q\right\}\right) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}\left(\mathcal{N}(0,1) \leq q\right) = 97\%.$$
Or, pour $p \geq 0$ (p est un paramètre positif rappelons-le),
$$\sqrt{n}\left(\frac{p - \hat{p}_n}{p\sqrt{1-p}}\right) \leq q \quad \Leftrightarrow \quad 0 \leq p \leq \hat{p}_n + qp\sqrt{1-p}/\sqrt{n}.$$
Le terme de droite dépendant encore de p, ce n'est pas encore un intervalle de confiance. Il suffit de remarquer que, comme $p \in [0,1]$, $p\sqrt{1-p} \leq \sqrt{p}\sqrt{p(1-p)} \leq 1/2$, et donc que
$$0 \leq p \leq \hat{p}_n + qp\sqrt{1-p}/\sqrt{n} \quad \Rightarrow 0 \leq p \leq \hat{p}_n + q/(2\sqrt{n}).$$
On en déduit alors que
$$\left[0; \hat{p}_n + q/(2\sqrt{n})\right]$$
est un intervalle de niveau de confiance asymptotique 97%.

6. On cherche à montrer que $p < 1/2$. On choisit donc les hypothèses
$$\begin{cases} H_0 & : \quad p = 1/2, \\ H_1 & : \quad p < 1/2. \end{cases}$$

7. Le test demandé est asymptotique, on peut utiliser l'intervalle de confiance asymptotique de la question précédente, ou construire le test pas à pas (ce qu'on va faire ici). Sous H_1, on s'attend à ce que $\hat{p}_n$ soit petit. On choisit alors une région de rejet de la forme $]0, t_{3\%}]$, où $t_{3\%}$ doit vérifier
$$\lim \sup_{n\to+\infty} P_{1/2}(\hat{p}_n \leq t_{3\%}) \leq 3\%.$$
Or, si $p = 1/2$, la question 4 donne
$$2\sqrt{2}\sqrt{n}\left(\hat{p}_n - (1/2)\right) \rightsquigarrow \mathcal{N}(0,1).$$
Soit q le quantile d'ordre 3% d'une $\mathcal{N}(0,1)$, on a alors
$$P_{1/2}\left(2\sqrt{2}\sqrt{n}\left(\hat{p}_n - (1/2)\right) \leq q\right) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}\left(\mathcal{N}(0,1) \leq q\right) = 3\%.$$
Or
$$2\sqrt{2}\sqrt{n}\left(\hat{p}_n - (1/2)\right) \leq q \Leftrightarrow \hat{p}_n \leq (1/2) + \frac{q}{2\sqrt{2n}}.$$
On peut alors prendre $t_{3\%} = (1/2) + \frac{q}{2\sqrt{2n}}$, et le test devient alors
$$T(X_{1:n}) = \mathbb{1}_{\hat{p}_n \leq (1/2) + \frac{q}{2\sqrt{2n}}}.$$

8. On observe $\bar{X}_n = 2,5 = 18/8$, et donc $\hat{p}_n = 8/18$. Par ailleurs, la table de quantiles donne
$$q_{3\%} = -q_{97\%} \geq -1.89 > 2,$$
où $q_{3\%}$ et $q_{97\%}$ désignent les quantiles de la loi $\mathcal{N}(0,1)$ (d'ordre 3% et 97%). On a alors
$$1/2 + q/(2\sqrt{2n}) > 1/2 - 2/(2 \times 20) = 9/20.$$
Comme $\frac{8}{18} < \frac{9}{20}$, on rejette l'hypothèse nulle, et on est sûrs à 97% (asymptotiquement) que la pièce est truquée.

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.

Solution 3. —

1. En notant X le nombre de carpes pêchées, on sait que X suit une loi hypergéométrique de paramètres n, N_0, N, $\mathcal{H}(n, N_0, N)$, donnée par
$$P_p(X = k) = \frac{\binom{N_0}{k}\binom{N-N_0}{n-k}}{\binom{N}{n}}..$$
Le modèle est alors
$$\left([\![0, n]\!], (\mathcal{H}(n, N_0, N))_{N-n \geq N_0 \geq n}\right).$$
La densité est donnée par la formule $P_p(X = k)$ ci-dessus.
2. Il s'agit de calculer $E_{N_0}(X)$. On peut le faire par calcul direct, ou raisonner en introduisant $(Y_i)_{i=1,\ldots,n}$ la suite de variables aléatoires qui valent 1 si le i-ème poisson pêché est une carpe (0 sinon). Chaque Y_i suit une loi $\mathcal{B}(p_{N_0})$, avec $p_{N_0} = N_0/N$, mais elles ne sont pas indépendantes (tirages sans remise). On a alors $X = \sum_{i=1}^n Y_i$, et donc $E_{N_0}(X) = \sum_{i=1}^n E_{N_0}(X_i) = nN_0/N$. Un estimateur par moments $\hat{N}_0$ vérifie alors
$$X = \frac{n\hat{N}_0}{N},$$
et donc $\hat{N}_0 = (NX)/n$.

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 $[\hat{N}_-;+\infty[$.
6. On cherche à prouver qu'il y a strictement plus de 1000 poissons dans l'étang. Donner les hypothèses du test associé.
7. Construire un test de niveau 5% pour ces hypothèses, basé sur $\hat{N}_n$.
8. **Application numérique**: On suppose qu'on a marqué 100 poissons ($n_0 = 100$), et qu'on repêche 50 fois. On observe 4 poissons marqués repêchés. Que peut-on en conclure? On pourra utiliser la table donnée en annexe.

Solution 4. —

1. On note Y_i la variable qui vaut 1 si le i-ème poisson pêché fait partie des n_0 marqués. On a $Y_i \sim \mathcal{B}(p_N)$, avec $p_N = n_0/N$. Les poissons étant pêchés avec remise, on peut supposer les Y_i i.i.d., et donc le nombre total de poissons marqués pêchés $X = \sum_{i=1}^n Y_i \sim \mathcal{B}(n, p_N)$. Le modèle est alors
$$\left([\![0, n]\!], (\mathcal{B}(n, p_N))_{N \geq n_0} \right).$$
La densité s'écrit, pour $x \in [\![0, n]\!]$, $f_p(x) = \binom{n}{x} p_N^x (1 - p_N)^{n-x}$.
2. Comme $E_N(X) = nn_0/N$, un estimateur par moments vérifie
$$X = \frac{nn_0}{\hat{N}_n},$$
et donc $\hat{N}_n = \frac{n_0}{(X/n)}$.
3. Les Y_i étant i.i.d., et $E_N(|Y_1|) < +\infty$, la loi des grands nombres donne
$$\frac{X}{n} \xrightarrow[n\to+\infty]{\mathbb{P}} \frac{n_0}{N}.$$
Comme $g : x \mapsto n_0/x$ est continue, $\hat{N}_n \xrightarrow[n\to+\infty]{\mathbb{P}} N$, et donc $\hat{N}_n$ est consistant.
4. Comme $E_N(Y_1^2) < +\infty$, le théorème central limite donne
$$\sqrt{n} \left(\frac{X}{n} - \frac{n_0}{N} \right) \rightsquigarrow \mathcal{N}(0, \mathrm{Var}_N(Y_1)),$$
avec $\mathrm{Var}_N(Y_1) = p_N(1 - p_N)$. Comme g est dérivable en n_0/N, la méthode Δ donne
$$\sqrt{n} \left(\hat{N}_n - N \right) \rightsquigarrow g'(n_0/N)\mathcal{N}(0, p_N(1 - p_N)) \sim \frac{N^2}{n_0}\mathcal{N}(0, n_0(N - n_0)/N^2) \sim \mathcal{N}(0, N^2(N - n_0)/n_0).$$
$\hat{N}_n$ est donc asymptotiquement normal.
5. D'après la question précédente,
$$\sqrt{n} \left(\frac{\hat{N}_n - N}{N\sqrt{(N - n_0)/n_0}} \right) \rightsquigarrow \mathcal{N}(0, 1).$$

Comme $\hat{N}_n$ est consistant, le Lemme de Slutsky garantit alors que

$$\sqrt{n}\left(\frac{\hat{N}_n - N}{\hat{N}_n\sqrt{(\hat{N}_n - n_0)/n_0}}\right) \rightsquigarrow \mathcal{N}(0,1).$$

Soit q le quantile d'ordre 95% d'une loi $\mathcal{N}(0,1)$. On déduit alors que

$$P_N\left(\left\{\sqrt{n}\left(\frac{\hat{N}_n - N}{\hat{N}_n\sqrt{(\hat{N}_n - n_0)/n_0}}\right) \leq q\right\}\right) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}(\mathcal{N}(0,1) \leq q) = 95\%.$$

Or, pour tout $N > 0$,

$$\sqrt{n}\left(\frac{\hat{N}_n - N}{\hat{N}_n\sqrt{(\hat{N}_n - n_0)/n_0}}\right) \leq q \quad \Leftrightarrow \quad N \geq \hat{N}_n - q\frac{\hat{N}_n\sqrt{\hat{N}_n - n_0}}{\sqrt{nn_0}}.$$

On en déduit que

$$\left[\hat{N}_n - q\frac{\hat{N}_n\sqrt{\hat{N}_n - n_0}}{\sqrt{nn_0}}; +\infty\right]$$

est un intervalle de niveau de confiance asymptotique 95%.

6. On cherche à montrer que $N > 1000$. On choisit donc les hypothèses

$$\begin{cases} H_0 & : \quad N = 1000, \\ H_1 & : \quad N > 1000. \end{cases}$$

7. Le test demandé est non asymptotique, on ne peut donc pas utiliser l'intervalle de confiance asymptotique des questions précédentes. Sous H_1, on s'attend à ce que $\hat{N}_n$ soit grand, on choisit alors une région de rejet de la forme $[t_{5\%}, +\infty[$, où $t_5\%$ doit vérifier

$$P_{1000}(\hat{N}_n \geq t_{5\%}) \leq 5\%.$$

Or, pour un $t \geq 1$, on a

$$\hat{N}_n \geq t \quad \Leftrightarrow \frac{nn_0}{\hat{N}_n} \leq \frac{nn_0}{t}.$$

Or, pour $N = 1000$, $nn_0/\hat{N}_n \sim \mathcal{B}(n, n_0/1000)$. Soit alors q le quantile d'ordre 5% d'une loi $\mathcal{B}(n, n_0/1000)$. On a

$$P_{1000}\left(\hat{N}_n \geq \frac{nn_0}{q}\right) = P_{1000}\left(\frac{nn_0}{\hat{N}_n} \leq q\right) \leq 5\%.$$

On en déduit que

$$T(X) = \mathbb{1}_{\hat{N}_n \geq \frac{nn_0}{q}}$$

est un test de niveau 5%.

8. Pour $n_0 = 100$ et $n = 50$, la table de la loi $\mathcal{B}(50, 0.1)$ fournie en annexe donne $q = 2$. On observe par ailleur $X = (nn_0/\hat{N}_n) = 4 > 2$. On ne peut donc pas rejeter l'hypothèse nulle (et on ne sait donc pas si le nombre de poissons dans l'étang dépasse 1000).

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 $r_n \to +\infty$, $r_n(N - \hat{N}_{n,2}) \rightsquigarrow 0$.
10. Construire un intervalle de niveau de confiance 98% sur N, de la forme $[\![\hat{N}_{n,2}, \hat{N}_+[$, basé sur $\hat{N}_{n,2}$.
11. Pour relancer la production de contre-mesures, on cherche à prouver que $N > 5000$. Donner les hypothèses du test associé.
12. Construire un test de niveau 99% pour ces hypothèses, basé sur $\hat{N}_{n,2}$.
13. En notant T_n ce test, montrer que $T_n - \mathbb{1}_{\hat{N}_{n,2} > 5000} \xrightarrow[n\to+\infty]{\mathbb{P}} 0$.

Solution 5. —

1. On note X_i le numéro du i-ème tank qui passe. On peut alors supposer que les X_i sont i.i.d. de loi $\mathcal{U}([\![1, N]\!])$. Le modèle est alors

$$\left((\mathbb{N}^*)^n, \left(\mathcal{U}([\![1, N]\!])^{\otimes n}\right)_{N\geq 1}\right).$$

La densité est donnée par, pour $x_1, \dots, x_n \geq 1$, $f_N(x_{1:n}) = \prod_{i=1}^n \mathbb{1}_{x_i \leq N}/N = N^{-n}\mathbb{1}_{m_n \leq N}$, avec $m_n = \max_{i=1,\dots,n} x_i$.

2. Comme $E_N(X_1) = (N+1)/2$, un estimateur par moments vérifie alors

$$\bar{X}_n = \frac{\hat{N}_{n,1} + 1}{2},$$

et donc $\hat{N}_{n,1} = 2\bar{X}_n - 1$.

3. Comme $E_N(X_1) < +\infty$, la loi des grands nombres donne $\bar{X}_n \xrightarrow[n\to+\infty]{\mathbb{P}} (N+1)/2$. Comme $x \mapsto 2x - 1$ est continue, $\hat{N}_{n,1} \xrightarrow[n\to+\infty]{\mathbb{P}} N$, et donc $\hat{N}_{n,1}$ est consistant.

4. Comme $E_N(X_1^2) < +\infty$, le théorème central limite donne

$$\sqrt{n}\left(\bar{X}_n - \frac{N+1}{2}\right) \rightsquigarrow \mathcal{N}(0, \mathrm{Var}_N(X_1)),$$

avec $\mathrm{Var}_N(X_1) = (N^2-1)/12$. On en déduit

$$\sqrt{n}\left(\hat{N}_{n,1} - N\right) = 2\sqrt{n}\left(\bar{X}_n - \frac{N+1}{2}\right) \rightsquigarrow \mathcal{N}(0, (N^2-1)/3),$$

et donc $\hat{N}_{n,1}$ est asymptotiquement normal.

5. La fonction $g : x \mapsto \log(x)$ étant dérivable en N (≥ 1), de dérivée $1/N$, la méthode Delta donne

$$\sqrt{n}\left(\log(\hat{N}_{n,1}) - \log(N)\right) \rightsquigarrow \frac{1}{N}\mathcal{N}(0, (N^2-1)/3) \sim \mathcal{N}(0, 1/3 - 1/(3N^2)),$$

et donc

$$\frac{N\sqrt{3n}}{\sqrt{N^2-1}}\left(\log(\hat{N}_{n,1}) - \log(N)\right) \rightsquigarrow \mathcal{N}(0,1).$$

Si q est le quantile d'ordre 98% d'une loi normale standard, on en déduit

$$P_N\left(\left\{\frac{N\sqrt{3n}}{\sqrt{N^2-1}}\left(\log(N) - \log(\hat{N}_{n,1})\right) \leq q\right\}\right) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}\left(\mathcal{N}(0,1) \leq q\right) = 98\%.$$

Or, pour $N \geq 1$ (et donc $\log(N) \geq 0$,

$$\begin{aligned}\frac{N\sqrt{3n}}{\sqrt{N^2-1}}\left(\log(N) - \log(\hat{N}_{n,1})\right) \leq q \quad &\Leftrightarrow \quad 0 \leq \log(N) \leq \log(\hat{N}_{n,1}) + q\frac{\sqrt{N^2-1}}{N\sqrt{3n}} \\ &\Rightarrow 0 \leq \log(N) \leq \log(\hat{N}_{n,1}) + q\frac{1}{\sqrt{3n}}.\end{aligned}$$

On en déduit que

$$\left[0; \log(\hat{N}_{n,1}) + q\frac{1}{\sqrt{3n}}\right]$$

est un intervalle de niveau de confiance asymptotique 98% sur $\log(N)$.

6. Comme

$$\lim\inf_{n\to+\infty} P_N\left(0 \leq \log(N) \leq \log(\hat{N}_{n,1}) + q\frac{1}{\sqrt{3n}}\right) \geq 98\%,$$

et $x \mapsto \exp(x)$ est croissante, on a

$$\lim\inf_{n\to+\infty} P_N\left(1 \leq N \leq \hat{N}_{n,1}\exp\left[q\frac{1}{\sqrt{3n}}\right]\right) \geq 98\%.$$

On en déduit que

$$\left[1; \hat{N}_{n,1}\exp\left(q\frac{1}{\sqrt{3n}}\right)\right]$$

est un intervalle de niveau de confiance asymptotique 98% sur N.

7. Pour $x_1, \dots, x_n \geq 1$, la vraisemblance s'écrit

$$V_{x_{1:n}}(N) = N^{-n} \mathbb{1}_{m_n \leq N},$$

où $m_n = \max_{i=1,\dots,n} x_i$. On en déduit que $V_{x_{1:n}}$ est maximale en m_n. L'estimateur du maximum de vraisemblance est $\hat{N}_{n,2} = \max_{i=1,\dots,n} X_i$.

8. Soit $\varepsilon > 0$. On a

$$P_N(|\hat{N}_{n,2} - N| \geq \varepsilon) \leq P_N(\hat{N}_{n,2} \neq N) = \left(1 - \frac{1}{N}\right)^n \underset{n \to +\infty}{\longrightarrow} 0,$$

$\hat{N}_{n,2}$ est donc consistant.

9. Soit r_n une suite positive tendant vers $+\infty$. On a, pour $t > 0$,

$$P_N(r_n|N - \hat{N}_{n,2}| \geq t) \leq P_N(\hat{N}_{n,2} \neq N) = \left(1 - \frac{1}{N}\right)^n \underset{n \to +\infty}{\longrightarrow} 0$$

On en déduit $r_n(N - \hat{N}_{n,2}) \xrightarrow[n \to +\infty]{\mathbb{P}} 0$, et donc $r_n(N - \hat{N}_{n,2}) \rightsquigarrow 0$.

10. Pour $t \in [1, N]$, on a

$$P_N\left(\hat{N}_{n,2} \leq t\right) = P_N\left(\max_{i=1,\dots n} X_i \leq \lfloor t \rfloor\right) = (\lfloor t \rfloor / N)^n.$$

Pour $t_{2\%} = N(2\%)^{1/n}$, on a

$$P_N\left(\hat{N}_{n,2} \leq t_{2\%}\right) \leq 2\%.$$

Donc,

$$P_N\left(t_{2\%} < \hat{N}_{n,2} \leq N\right) \geq 98\%.$$

Or,

$$\begin{aligned} t_{2\%} < \hat{N}_{n,2} &\Rightarrow N(2\%)^{1/n} < \hat{N}_{n,2} \\ &\Rightarrow N < \frac{\hat{N}_{n,2}}{(2\%)^{1/n}}. \end{aligned}$$

Comme par ailleur $\hat{N}_{n,2} \leq N$, on en déduit que

$$\left[\hat{N}_{n,2}; \hat{N}_{n,2}(2\%)^{-1/n}\right[$$

est un intervalle de niveau de confiance 98%.

11. On cherche à montrer que $N > 5000$. Les hypothèses de test sont donc

$$\begin{cases} H_0 & : \quad N = 5000, \\ H_1 & : \quad N > 5000. \end{cases}$$

12. Sous H_1 on s'attend à ce que $\hat{N}_{n,2}$ soit grand. On choisit alors une région de rejet de la forme $[\![t_{1\%};+\infty[$, où $t_{1\%}$ doit vérifier

$$P_{5000}(\hat{N}_{n,2} \geq t_{1\%}) \leq 1\%.$$

En reprenant la question 11−, pour $t \in [1;5000]$,

$$P_{5000}(\hat{N}_{n,2} \leq t) = (\lfloor t \rfloor/5000)^n.$$

Pour $t = \lceil 5000 \times (99\%)^{1/n} \rceil$, on a

$$P_{5000}(\hat{N}_{n,2} \leq t) \geq 99\%.$$

On en déduit que, pour $t_{1\%} = \lceil 5000 \times (99\%)^{1/n} \rceil + 1$,

$$P_{5000}(\hat{N}_{n,2} \geq t_{1\%}) \leq 1\%.$$

Le test

$$T(X_{1:n}) = \mathbb{1}_{\hat{N}_{n,2} \geq \lceil 5000 \times (99\%)^{1/n} \rceil + 1}$$

est donc de niveau 1%.

13. On a $(99\%)^{1/n} \underset{n\to+\infty}{\longrightarrow} 1$, donc, pour $n \geq n_0$, $\lceil 5000 \times (99\%)^{1/n} \rceil = 5000$. On en déduit alors que, pour $n \geq n_0$,

$$\mathbb{1}_{\hat{N}_{n,2} \geq \lceil 5000 \times (99\%)^{1/n} \rceil + 1} = \mathbb{1}_{\hat{N}_{n,2} > 5000},$$

et donc $\mathbb{1}_{\hat{N}_{n,2} \geq \lceil 5000 \times (99\%)^{1/n} \rceil + 1} - \mathbb{1}_{\hat{N}_{n,2} > 5000} \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0$.

Exercice 6. — Stylométrie de discours

Un chercheur en sciences sociales s'intéresse aux occurences du mot 'sécurité' dans les discours de J. Chirac. Ce nombre d'occurences 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.

10. Un adversaire politique cherche à montrer que J. Chirac est obnubilé par la sécurité, c'est à dire que $\lambda > 50$. Quelles sont les hypothèses de test associées?
11. Construire un test de niveau **asymptotique** 6% pour ces hypothèses, basé sur $\hat{\lambda}_n$.
12. **Application**: Sur 200 discours, J. Chirac a employé le mot sécurité 10200 fois. Que peut-on en conclure? On pourra utiliser les tables de quantiles données en annexe.

SOLUTION 6. —

1. On note X_i le nombre d'occurences du mot 'sécurité' dans le i-ème discours de J. Chirac. On peut supposer que les X_i sont i.i.d. de loi $\mathcal{P}(\lambda)$, pour un paramètre $\lambda > 0$ inconnu. Le modèle s'écrit alors

$$\left(\mathbb{N}^n, \left(\mathcal{P}(\lambda)^{\otimes n}\right)_{\lambda>0}\right).$$

La densité est donnée par, pour $x_1, \dots, x_n \in \mathbb{N}$,

$$f_\lambda(x_{1:n}) = \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i} \prod_{i=1}^n \frac{1}{x_i!}.$$

2. Par la méthode du maximum de vraisemblance. Pour $x_{1:n} \in \mathbb{N}^n$, la vraisemblance s'écrit

$$V_{x_{1:n}}(\lambda) = e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i} \prod_{i=1}^n \frac{1}{x_i!}.$$

Comme pour tout $\lambda > 0$, $V_{x_{1:n}}(\lambda) > 0$, on peut maximiser la log-vraisemblance

$$\ell_{x_{1:n}}(\lambda) = \log(V_{x_{1:n}}(\lambda)) = -n\lambda + \sum_{i=1}^n x_i \log(\lambda) + \log\left(\prod_{i=1}^n \frac{1}{x_i!}\right).$$

Cette fonction est strictement concave, un extremum local sera nécessairement un maximum. On résout alors

$$\begin{aligned} \ell'_{x_{1:n}}(\lambda) = 0 \quad &\Leftrightarrow \quad -n + \left(\sum_{i=1}^n x_i\right)/\lambda = 0 \\ &\Leftrightarrow \quad \lambda = \bar{x}_n. \end{aligned}$$

Attention, si $x_{1:n} = 0_{1:n}$, la dérivée ne s'annule pas (elle vaut $-n$).

Traitons donc ce cas à part. Pour $x_{1:n} = 0_{1:n}$, $\ell_{x_{1:n}}(\lambda) = -n\lambda$ est maximale en $\lambda = 0$.

Dans les deux cas, on a un maximum en $\lambda = \bar{x}_n$. L'estimateur du maximum de vraisemblance est alors

$$\hat{\lambda}_n = \bar{X}_n.$$

3. Comme $E_\lambda(|X_1|) = \lambda < +\infty$, la loi des grands nombres donne

$$\bar{X}_n = \hat{\lambda}_n \xrightarrow[n\to+\infty]{\mathbb{P}} \lambda.$$

$\hat{\lambda}_n$ est donc consistant.

4. Comme $\mathrm{Var}_\lambda(X_1) = \lambda < +\infty$, le théorème central limite donne

$$\sqrt{n}(\hat{\lambda}_n - \lambda) \rightsquigarrow \mathcal{N}(0, \mathrm{Var}_\lambda(X_1)),$$

avec $\mathrm{Var}_\lambda(X_1) = \lambda$. $\hat{\lambda}_n$ est donc asymptotiquement normal.

5. Comme $g : x \mapsto \sqrt{x}$ est dérivable en λ, la méthode Delta donne
$$\sqrt{n}(\sqrt{\hat{\lambda}_n} - \sqrt{\lambda}) \rightsquigarrow g'(\lambda)\mathcal{N}(0,\lambda) \sim \frac{1}{2\sqrt{\lambda}}\mathcal{N}(0,\lambda) \sim \mathcal{N}(0,1/4).$$

6. D'après la question précédente on déduit
$$2\sqrt{n}(\sqrt{\hat{\lambda}_n} - \sqrt{\lambda}) \rightsquigarrow \mathcal{N}(0,1).$$
Soit q le quantile d'ordre 98% d'une loi $\mathcal{N}(0,1)$. On a alors
$$P_\lambda\left(\left\{\left|2\sqrt{n}(\sqrt{\hat{\lambda}_n} - \sqrt{\lambda})\right| \leq q\right\}\right) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}\left(|\mathcal{N}(0,1)| \leq q\right) = 96\%.$$
Or, pour $\lambda > 0$,
$$\left|2\sqrt{n}(\sqrt{\hat{\lambda}_n} - \sqrt{\lambda})\right| \leq q \quad \Leftrightarrow \quad \sqrt{\hat{\lambda}_n} - \frac{q}{2\sqrt{n}} \leq \sqrt{\lambda} \leq \sqrt{\hat{\lambda}_n} + \frac{q}{2\sqrt{n}}.$$
On en déduit que
$$\left[\sqrt{\hat{\lambda}_n} - \frac{q}{2\sqrt{n}}; \sqrt{\hat{\lambda}_n} + \frac{q}{2\sqrt{n}}\right]$$
est un intervalle de niveau de confiance asymptotique 96%.

7. Comme $x \mapsto x^2$ est croissante sur $\mathbb{R}^+$, on déduit de la question précédente que
$$I_1 = \left[\left(\sqrt{\hat{\lambda}_n} - \frac{q}{2\sqrt{n}}\right)^2; \left(\sqrt{\hat{\lambda}_n} + \frac{q}{2\sqrt{n}}\right)^2\right]$$
est un intervalle de niveau de confiance asymptotique 96% sur λ.

8. On a $E_\lambda(\hat{\lambda}_n) = \lambda$, et $\mathrm{Var}_\lambda(\hat{\lambda}_n) = \mathrm{Var}_\lambda(X_1)/n = \lambda/n$. Une décomposition biais-variance donne alors
$$E_\lambda((\hat{\lambda}_n - \lambda)^2) = \mathrm{Var}_\lambda(\hat{\lambda}_n) + (E_\lambda(\hat{\lambda}_n) - \lambda)^2 = \lambda/n.$$

9. On utilise l'inégalité de Markov. Pour $t > 0$,
$$P_\lambda\left(\left\{|\hat{\lambda}_n - \lambda| > t\right\}\right) = P_\lambda\left(\left\{(\hat{\lambda}_n - \lambda)^2 > t^2\right\}\right) \leq \frac{E_\lambda\left((\hat{\lambda}_n - \lambda)^2\right)}{t^2} = \frac{\lambda}{nt^2}.$$
Posons $t_{4\%} = \sqrt{\frac{\lambda}{n4\%}}$, de sorte que
$$P_\lambda\left(\left\{|\hat{\lambda}_n - \lambda| \leq t_{4\%}\right\}\right) \geq 1 - 4\% = 96\%,$$
d'après ce qui précède. On peut réécrire, pour $\lambda > 0$,
$$|\hat{\lambda}_n - \lambda| \leq t_{4\%} \quad \Leftrightarrow \quad \hat{\lambda}_n - \sqrt{\frac{\lambda}{n4\%}} \leq \lambda \leq \hat{\lambda}_n + \sqrt{\frac{\lambda}{n4\%}}.$$
La première inéquation peut se réécrire
$$\left(\sqrt{\lambda} + \frac{1}{2\sqrt{n4\%}}\right)^2 \geq \hat{\lambda}_n + \frac{1}{n16\%},$$

ce qui implique

$$\sqrt{\lambda} \geq \sqrt{\hat{\lambda}_n + \frac{1}{n16\%}} - \frac{1}{2\sqrt{n4\%}}.$$

Le terme de droite étant positif, cela implique

$$\lambda \geq \left(\sqrt{\hat{\lambda}_n + \frac{1}{n16\%}} - \frac{1}{2\sqrt{n4\%}} \right)^2.$$

La deuxième inéquation se réécrit comme

$$\left(\sqrt{\lambda} - \frac{1}{2\sqrt{n4\%}} \right)^2 \leq \hat{\lambda}_n + \frac{1}{n16\%},$$

ce qui mène à

$$\sqrt{\lambda} \leq \frac{1}{2\sqrt{n4\%}} + \sqrt{\hat{\lambda}_n + \frac{1}{n16\%}},$$

et donc à

$$\lambda \leq \left(\frac{1}{2\sqrt{n4\%}} + \sqrt{\hat{\lambda}_n + \frac{1}{n16\%}} \right)^2.$$

Un intervalle au niveau de confiance 96% pour λ est donc

$$\left[\left(\sqrt{\hat{\lambda}_n + \frac{1}{n16\%}} - \frac{1}{2\sqrt{n4\%}} \right)^2 ; \left(\frac{1}{2\sqrt{n4\%}} + \sqrt{\hat{\lambda}_n + \frac{1}{n16\%}} \right)^2 \right].$$

10. On cherche à montrer que $\lambda > 50$. Les hypothèses de test sont donc

$$\begin{cases} H_0 & : \quad \lambda = 50, \\ H_1 & : \quad \lambda > 50. \end{cases}$$

11. Sous H_1, on s'attend à ce que $\hat{\lambda}_n$ soit grand. On choisit donc pour région de rejet $[t_{6\%}; +\infty[$, où $t_{6\%}$ doit vérifier

$$P_{50}\left(\hat{\lambda}_n \geq t_{6\%} \right) \leq 6\%.$$

Or, pour $\lambda = 50$, la normalité asymptotique obtenue en question 4 donne

$$\sqrt{n/50}(50 - \hat{\lambda}_n) \rightsquigarrow \mathcal{N}(0,1).$$

Si q est le quantile d'ordre 6% d'une loi normale standard, on en déduit que

$$P_{50}\left((\sqrt{n/50}(50 - \hat{\lambda}_n) \leq q \right) \underset{n \to +\infty}{\longrightarrow} 6\%,$$

et donc que

$$T(X_{1:n}) = \mathbb{1}_{\hat{\lambda}_n \geq 50 - q\sqrt{50/n}}$$

est un test de niveau asymptotique 6%.

12. On a $n = 200$, et $\hat{\lambda}_n = \bar{X}_n = 10200/200 = 51$. Les tables en annexe donnent $1.75 \leq -q \leq 1.76$, et donc
$$50 - q\sqrt{50/n} \leq 50 + \frac{1.76}{2} < \hat{\lambda}_n.$$
On rejette donc H_0, et on est sûrs à 94% que J. Chirac est obnubilé par la sécurité (dans ses discours).

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 $[0, \hat{\lambda}_+]$).
6. L'usager cherche à prouver que la qualité de service s'est dégradée. Auparavant il avait un temps d'attente moyen de 5. Donner les hypothèses du test correspondant.
7. Construire un test de niveau asymptotique 4% basé sur $\hat{\lambda}_n$ pour ces deux hypothèses.
8. **Application numérique**: L'usager répète l'expérience 25 fois, et observe un temps d'attente moyen de 6. Que peut-il en conclure (on pourra utiliser les tables de la loi normale données en annexe)?

Solution 7. —

1. Pour $j \in [\![1,3]\!]$ et $i \in [\![1,n]\!]$, on note $Y_{i,j}$ le temps que met le j-ème ascenseur à arriver au jour i. On observe $(X_i)_{i=1,\ldots,n}$ i.i.d., avec $X_i = \min_{j=1,2,3} Y_{i,j}$. Calculons la loi de X_1. Comme $(Y_{1,1}, Y_{1,2}, Y_{1,3})$ sont indépendantes de loi $\mathcal{E}(\lambda)$, pour $t > 0$ on a
$$\begin{aligned} P_\lambda(X_1 > t) &= P_\lambda(\min_{j=1,2,3} Y_{1,j} > t) = P_\lambda(Y_{1,1} > t)^3 \qquad \text{(par indépendance des } Y_{1,j}) \\ &= e^{-3\lambda t}. \end{aligned}$$
On en déduit que $X_1 \sim \mathcal{E}(3\lambda)$. Le modèle est alors
$$\left( ]0, +\infty[^n, \left(\mathcal{E}(3\lambda)^{\otimes n}\right)_{\lambda>0} \right).$$
La densité est donnée par, pour $x_1, \ldots, x_n > 0$,
$$f_\lambda(x_{1:n}) = \prod_{i=1}^n (3\lambda e^{-3\lambda x_i}) = (3\lambda)^n e^{-3\lambda \sum_{i=1}^n x_i}.$$

2. Par la méthode des moments. On a $E_\lambda X_1 = 1/(3\lambda)$. L'estimateur par moments $\hat{\lambda}_n$ vérifie alors l'équation
$$\bar{X}_n = \frac{1}{3\hat{\lambda}_n},$$
ce dont on déduit $\hat{\lambda}_n = \frac{1}{3\bar{X}_n}$.

3. Comme $E_\lambda|X_1| = 1/\lambda < +\infty$, la loi des grands nombres donne

$$\bar{X}_n \xrightarrow[n\to+\infty]{\mathbb{P}} 1/(3\lambda).$$

Comme $f : x \mapsto 1/3x$ est continue en $1/(3\lambda)$, on en déduit

$$\hat{\lambda}_n = f(\bar{X}_n) \xrightarrow[n\to+\infty]{\mathbb{P}} f(1/(3\lambda)) = \lambda.$$

$\hat{\lambda}_n$ est donc consistant.

4. Comme $\mathrm{Var}_\lambda(X_1) = 1/(9\lambda^2) < +\infty$, le théorème central limite donne

$$\sqrt{n}(\bar{X}_n - 1/(3\lambda)) \rightsquigarrow \mathcal{N}(0, \mathrm{Var}_\lambda(X_1)),$$

avec $\mathrm{Var}_\lambda(X_1) = 1/(9\lambda^2)$. Comme f est différentiable en $1/(3\lambda)$, on en déduit par la méthode Δ

$$\sqrt{n}(\hat{\lambda}_n - \lambda) = \sqrt{n}(f(\bar{X}_n) - f(1/(3\lambda))) \rightsquigarrow f'(1/(3\lambda))\mathcal{N}(0, (9\lambda^2)^{-1}) \sim \mathcal{N}(0, \lambda^2).$$

5. De la question précédente on déduit que, pour tout $\lambda > 0$,

$$\sqrt{n}\left(\frac{\hat{\lambda}_n - \lambda}{\lambda}\right) \rightsquigarrow \mathcal{N}(0,1).$$

Soit q le quantile d'ordre 90% d'une loi $\mathcal{N}(0,1)$. On déduit alors que

$$P_\lambda\left[\left\{\sqrt{n}\left(\frac{\hat{\lambda}_n}{\lambda} - 1\right) < -q\right\}\right] \xrightarrow[n\to+\infty]{} 10\%.$$

Or, pour $\lambda > 0$,

$$\sqrt{n}\left(\frac{\hat{\lambda}_n}{\lambda} - 1\right) < -q \quad \Leftrightarrow \quad \lambda > \frac{\hat{\lambda}_n}{1 - q/\sqrt{n}}.$$

On en déduit alors que

$$\left]0, \frac{\hat{\lambda}_n}{1 - q/\sqrt{n}}\right]$$

est un intervalle de niveau de confiance asymptotique 90% (on rappelle que $\lambda > 0$).

6. Le temps moyen d'attente étant $E_\lambda(X_1) = 1/(3\lambda)$. On cherche donc à prouver que $1/3\lambda > 5$, ou encore $\lambda < 1/15$. Les hypothèses de test sont alors

$$\begin{cases} H_0 & : \quad \lambda = 1/15, \\ H_1 & : \quad \lambda < 1/15. \end{cases}$$

7. Remarque: comme $\bar{X}_n \sim \Gamma(n, 3n\lambda)$, on pourrait construire un test non asymptotique. Encore une remarque: on peut se servir de l'intervalle de confiance asymptotique précédent. Ici on construit pas à pas. Sous H_1, on s'attend à ce que $\hat{\lambda}_n$ soit petit, on choisit alors une région de rejet de la forme $]0; t_{4\%}]$, où $t_{4\%}$ doit vérifier

$$\lim \sup_{n\to+\infty} P_{1/15}(\hat{\lambda}_n \leq t_{4\%}) \leq 4\%.$$

De la normalité asymptotique donnée en question 4 on déduit, pour $\lambda = 1/15$,

$$15\sqrt{n}(\hat{\lambda}_n - 1/15) \rightsquigarrow \mathcal{N}(0,1).$$

Soit q le quantile d'ordre 4% d'une loi $\mathcal{N}(0,1)$. On en déduit alors que

$$P_{1/2}\left(15\sqrt{n}(\hat{\lambda}_n - 1/15) \leq q\right) \underset{n\to+\infty}{\longrightarrow} 4\%.$$

Or,

$$15\sqrt{n}(\hat{\lambda}_n - 1/15) \leq q \quad \Leftrightarrow \quad \hat{\lambda}_n \leq (1/15) + q/(15\sqrt{n}).$$

On en déduit alors que

$$T(X_{1:n}) = \mathbb{1}_{\hat{\lambda}_n \leq (1/15)+q/(15\sqrt{n})}$$

est un test de niveau asymptotique 4%.

8. On a $n = 25$ et $\bar{X}_n = 6$, donc $\hat{\lambda}_n = 1/18$. La table des quantiles de la loi normale donne

$$-1.76 \leq q \leq -1.74,$$

en remarquant que pour une telle loi le quantile d'ordre 4% vaut moins le quantile d'ordre 96%. On en déduit alors que

$$\frac{1}{15} + \frac{q}{15\sqrt{n}} \leq \frac{1}{15} - \frac{1.75}{75} = \frac{13}{300}.$$

Comme $\frac{1}{18} > \frac{13}{300}$, l'usager accepte H_0 et ne peut rien en conclure.

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 $\sqrt{n}(\bar{X}_n - \mu)^2 \underset{n\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} 0$ (on pourra utiliser l'inégalité de Markov).
8. Montrer que $\hat{v}$ asymptotiquement normal. On admettra que si $X \sim \mathcal{N}(0,1)$, $\mathrm{Var}(X^2) = 2$.
9. On suppose que v est connue. Donner un intervalle bilatère de niveau de confiance 96% sur μ.
10. On suppose que v est inconnue. Donner un intervalle de niveau de confiance asymptotique 96% sur μ.

11. L'étudiant veut prouver que son niveau moyen est supérieur à la note seuil (10). Quelles sont les hypothèses du test correspondant?

12. Dans le cas où v est connue, construire un test de niveau 5% pour ces hypothèses, basé sur $\hat{\mu}_n$.

13. Dans le cas où v est inconnue, construire un test de niveau asymptotique 5% pour ces hypothèses, basé sur $\hat{\mu}_n$.

14. **Application numérique**: L'étudiant a observé une moyenne de 11 sur les 16 épreuves du premier semestre, avec une variance empirique de 4. Que peut-il en conclure? On pourra utiliser la table des quantiles de la loi normale fournie en annexe.

15. On suppose que μ et v sont inconnus. Donner un intervalle de niveau de confiance asymptotique 80% pour v de type $[\hat{v}_-, +\infty[$.

SOLUTION 8. —

1. On note X_i la note de l'étudiant à la i-ème matière. On peut supposer que les X_i sont i.i.d., de loi $\mathcal{N}(\mu, v)$. Le modèle s'écrit alors

$$\left(\mathbb{R}^n, \left(\mathcal{N}(\mu, v)^{\otimes n}\right)_{\mu \geq 0, v > 0}\right).$$

La densité s'écrit, pour $x_1, \ldots, x_n \in \mathbb{R}$,

$$f_{\mu,v}(x_{1:n}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi v}} e^{-(x_i-\mu)^2/(2v)} = (2\pi v)^{-n/2} e^{-\sum_{i=1}^{n}(x_i-\bar{x}_n)^2/(2v) - n(\bar{x}_n-\mu)^2/2v}.$$

2. Par la méthode du maximum de vraisemblance. Pour $x_1, \ldots, x_n \in \mathbb{R}$, la vraisemblance s'écrit

$$V_{x_{1:n}}(\mu, v) = (2\pi v)^{-n/2} e^{-\sum_{i=1}^{n}(x_i-\bar{x}_n)^2/(2v) - n(\bar{x}_n-\mu)^2/2v}.$$

Comme, pour tout $\mu \in \mathbb{R}$ et $v > 0$, $V_{x_{1:n}}(\mu, v) > 0$, on peut se contenter de maximiser la log-vraisemblance

$$\ell_{x_{1:n}}(\mu, v) = \log\left(V_{x_{1:n}}(\mu, v)\right) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(v) - \frac{1}{2v}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 - \frac{n}{2v}(\bar{x}_n - \mu)^2.$$

Il est immédiat que, pour tout $\mu \in \mathbb{R}$, $\ell_{x_{1:n}}(\mu, v) \leq \ell(\bar{x}_n, v)$ (avec égalité ssi $\mu = \bar{x}_n$). Par ailleurs, $g : v \mapsto \ell(\bar{x}_n, v)$ est concave, on peut donc se contenter d'annuler sa dérivée. Pour $t > 0$, on a

$$\begin{aligned} g'(t) = 0 \quad &\Leftrightarrow \quad -\frac{n}{2t} + \frac{1}{2t^2}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 = 0 \\ &\Leftrightarrow \quad t = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2. \end{aligned}$$

On en déduit que les estimateurs du maximum de vraisemblance sont

$$\begin{aligned} \hat{\mu}_n &= \bar{X}_n, \\ \hat{v}_n &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2. \end{aligned}$$

3. Comme $E_{\mu,v}|X_1| < +\infty$, la loi des grands nombres donne
$$\hat{\mu}_n = \bar{X}_n \xrightarrow[n\to+\infty]{\mathbb{P}} E_{\mu,v}X_1 = \mu,$$
et donc $\hat{\mu}_n$ est consistant.
4. Pour $\hat{v}_n$, on peut écrire
$$\hat{v}_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 - (\bar{X}_n - \mu)^2.$$
5. Comme $E_{\mu,v}(X_1-\mu)^2 = v < +\infty$, la loi des grands nombres donne
$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 \xrightarrow[n\to+\infty]{\mathbb{P}} E_{\mu,v}(X_1-\mu)^2 = v.$$
Comme $\bar{X}_n \xrightarrow[n\to+\infty]{\mathbb{P}} \mu$, et que $x \mapsto x^2$ est continue, on déduit que $(\bar{X}_n-\mu)^2 \xrightarrow[n\to+\infty]{\mathbb{P}} 0$, et donc, en utilisant la question précédente, que
$$\hat{v}_n = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 - (\bar{X}_n-\mu)^2 \xrightarrow[n\to+\infty]{\mathbb{P}} v - 0,$$
$\hat{v}_n$ est donc consistant.
6. On remarque que $\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}(\mu, v/n)$ (les X_i sont i.i.d. $\mathcal{N}(\mu,v)$). Donc
$$(\hat{\mu}_n - \mu) \sim \mathcal{N}(0, v/n),$$
et $\hat{\mu}_n$ est **non-asymptotiquement** normal (c'est une égalité en loi).
7. Comme $(\bar{X}_n-\mu) \sim \mathcal{N}(0,v/n)$, $E_{\mu,v}(\sqrt{n}(\bar{X}_n-\mu)^2) = v/\sqrt{n} \xrightarrow[n\to+\infty]{} 0$. Soit $\varepsilon > 0$, on a alors, en utilisant l'inégalité de Markov,
$$P_{\mu,v}\left(\left|\sqrt{n}(\bar{X}_n-\mu)^2\right| > \varepsilon\right) \leq \frac{v}{\sqrt{n}\varepsilon} \xrightarrow[n\to+\infty]{} 0.$$
Donc $\sqrt{n}(\bar{X}_n-\mu)^2) \xrightarrow[n\to+\infty]{\mathbb{P}} 0$.
8. Pour la normalité asymptotique de $\hat{v}_n$, on commence par remarquer que $E_{\mu,v}(X_1-\mu)^4 < +\infty$. Le théorème central limite donne alors
$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 - v\right) \rightsquigarrow \mathcal{N}(0, \mathrm{Var}_{\mu,v}(X_1-\mu)^2).$$
Comme $(X_1-\mu)^2) \sim vZ^2$, où $Z \sim \mathcal{N}(0,1)$, on a $\mathrm{Var}_{\mu,v}((X_1-\mu)^2)) = 2v^2$ (en utilisant $\mathrm{Var}(Z^2) = 2$), et donc
$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 - v\right) \rightsquigarrow \mathcal{N}(0, 2v^2).$$
Par ailleurs, on a
$$\sqrt{n}(\hat{v}_n - v) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 - v\right) - \sqrt{n}(\bar{X}_n-\mu)^2.$$

Comme $\sqrt{n}(\bar{X}_n - \mu)^2) \xrightarrow[n\to+\infty]{\mathbb{P}} 0$ (cf. question précédente), le lemme de Slutsky donne alors

$$\sqrt{n}(\hat{v}_n - v) \rightsquigarrow \mathcal{N}(0, 2v^2),$$

donc $\hat{v}_n$ est asymptotiquement normal.

9. D'après la question précédente, on a

$$\sqrt{n}\left(\frac{\hat{\mu}_n - \mu}{\sqrt{v}}\right) \sim \mathcal{N}(0,1).$$

Soit q le quantile d'ordre 98% d'une loi $\mathcal{N}(0,1)$. On a alors

$$P_\mu\left(\left\{\left|\frac{\hat{\mu}_n - \mu}{\sqrt{v}}\right| \leq q\right\}\right) = 96\%.$$

On en déduit immédiatement que

$$\left[\hat{\mu}_n - q\frac{\sqrt{v}}{\sqrt{n}}; \hat{\mu}_n + q\frac{\sqrt{v}}{\sqrt{n}}\right]$$

est un intervalle de niveau de confiance 96%.

10. Comme $\hat{v}_n$ est consistant, la question précédente et le Lemme de Slutsky assurent que

$$\sqrt{n}\left(\frac{\hat{\mu}_n - \mu}{\sqrt{\hat{v}_n}}\right) \rightsquigarrow \mathcal{N}(0,1).$$

Comme à la question précédente on en déduit que

$$\left[\hat{\mu}_n - q\frac{\sqrt{\hat{v}_n}}{\sqrt{n}}; \hat{\mu}_n + q\frac{\sqrt{\hat{v}_n}}{\sqrt{n}}\right]$$

est un intervalle de niveau de confiance asymptotique 96%.

11. L'étudiant veut prouver que $\mu > 10$. Les hypothèses de test sont alors

$$\begin{cases} H_0 & : \quad \mu = 10, \\ H_1 & : \quad \mu > 10. \end{cases}$$

12. Sous l'hypothèse H_1, on s'attend à ce que $\hat{\mu}_n$ soit grand. On choisit donc une région de rejet de la forme $[t_{5\%}; +\infty[$, où $t_{5\%}$ doit vérifier

$$P_{10}(\hat{\mu}_n \geq t_{5\%}) \leq 5\%.$$

D'après la question 6$-$, si $\mu = 10$, on a

$$\sqrt{n/v}(10 - \hat{\mu}_n) \sim \mathcal{N}(0,1),$$

c'est une égalité en loi (pas une convergence). On en déduit que, si q est le quantile d'ordre 5% d'une loi $\mathcal{N}(0,1)$,

$$P_{10}\left(\sqrt{n/v}(10 - \hat{\mu}_n) \leq q\right) = 5\%.$$

Or

$$\sqrt{n/v}(10 - \hat{\mu}_n) \leq q \quad \Leftrightarrow \quad \hat{\mu}_n \geq 10 - \frac{q\sqrt{v}}{\sqrt{n}}.$$

Comme v est connue, on en déduit que

$$T(X_{1:n}) = \mathbb{1}_{\hat{\mu}_n \geq 10 - \frac{q\sqrt{v}}{\sqrt{n}}}$$

est un test de niveau 5% pour ces deux hypothèses.

13. On a la même forme de région de rejet qu'à la question précédente, mais cette fois-ci on utilise la consistance de $\hat{v}_n$ et le Lemme de Slutsky pour déduire que, si $\mu = 10$,

$$\sqrt{n/\hat{v}_n}(10 - \hat{\mu}_n) \rightsquigarrow \mathcal{N}(0,1),$$

cette fois-ci on a une convergence. On en déduit alors, avec q le quantile d'ordre 5% d'une loi $\mathcal{N}(0,1)$, que

$$P_{10}\left(\sqrt{n/\hat{v}_n}(10 - \hat{\mu}_n) \leq q\right) \underset{n\to+\infty}{\longrightarrow} 5\%,$$

et donc que

$$T(X_{1:n}) = \mathbb{1}_{\hat{\mu}_n \geq 10 - \frac{q\sqrt{\hat{v}_n}}{\sqrt{n}}}$$

est un test de niveau **asymptotique** 5% pour ces deux hypothèses. Remarque: en fait la loi de $\sqrt{n}(\hat{\mu}_n - \mu)/\sqrt{\hat{v}_n/(n-1)v)}$ est connue (loi de Student), on aurait pu en faire un test non asymptotique.

14. On a $\hat{\mu}_n = 11$, $n = 16$ et $\hat{v}_n = 4$. Par ailleurs, la table fournie en annexe donne $-q \leq 1.645$ (en rappelant que par symétrie $-q$ est le quantile d'ordre 95% d'une loi normale standard. On a alors

$$10 - \frac{q\sqrt{\hat{v}_n}}{\sqrt{n}} \leq 10 - 1.645\sqrt{1/4} < 11.$$

On rejette donc H_0, et l'étudiant est sûr à 95% que son niveau moyen est supérieur à la note seuil.

15. D'après la question 4-, on a

$$\sqrt{n}\left(\frac{\hat{v}_n - v}{\sqrt{2}v}\right) \rightsquigarrow \mathcal{N}(0,1).$$

Soit q le quantile d'ordre 80% d'une loi $\mathcal{N}(0,1)$. On a alors

$$P_{\mu,v}\left(\left\{\sqrt{n}\left(\frac{\hat{v}_n}{\sqrt{2}v} - \frac{1}{\sqrt{2}}\right) > q\right\}\right) \underset{n\to+\infty}{\longrightarrow} 20\%.$$

Or,

$$\sqrt{n}\left(\frac{\hat{v}_n}{\sqrt{2}v} - \frac{1}{\sqrt{2}}\right) > q \quad \Leftrightarrow \quad v < \frac{\hat{v}_n}{1 + \sqrt{2}q/\sqrt{n}}.$$

On en déduit que

$$\left[\frac{\hat{v}_n}{1 + \sqrt{2}q/\sqrt{n}}; +\infty\right[$$

est un intervalle de niveau de confiance asymptotique (exact) 80%.

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. On veut montrer que l'âge de l'univers est plus petit que 10^{10}. Donner les hypothèses du test associé.
7. Construire un test de niveau 3% basé sur $\hat{A}_{n,1}$ pour ces hypothèses.
8. Donner un estimateur de A par la méthode du maximum de vraisemblance. On le notera $\hat{A}_{n,2}$.
9. Montrer que $\hat{A}_{n,2}$ est consistant.
10. Montrer que $n(A - \hat{A}_{n,2}) \rightsquigarrow \mathcal{E}(a)$, où a est un paramètre à déterminer.
11. Construire un test de niveau asymptotique 3% basé sur $\hat{A}_{n,2}$ pour les hypothèses de la question 6.
12. **Application numérique**: En admettant que $\log(1/0.03) \approx 3.506$, on observe sur 40 corps célestes un âge maximal de 9×10^9. Que peut-on en conclure?
13. Montrer que $E_A((\hat{A}_{n,2} - A)^2) = 2A^2/((n+1)(n+2))$. Comparer avec le risque quadratique de $\hat{A}_{n,1}$.
14. Donner un intervalle de niveau de confiance 95% sur A basé sur $\hat{A}_{n,1}$. On le notera I_1.
15. Donner un intervalle de niveau de confiance 95% sur A basé sur $\hat{A}_{n,2}$. On le notera I_2.
16. Comparer I_1 et I_2 sur la base de leurs longueurs lorsque $n \to +\infty$.

Solution 9. —

1. On note X_i l'âge du i-ème corps céleste arrivé sur Terre. On peut supposer que les X_i sont i.i.d., de loi $\mathcal{U}(]0, A])$, où A est inconnu. Le modèle est donc

$$\left( ]0, +\infty[^n, \left(\mathcal{U}(]0, A])^{\otimes n}\right)_{A>0} \right).$$

La densité est donnée par, pour $x_1, \ldots, x_n > 0$,

$$f_A(x_{1:n}) = \prod_{i=1}^{n} \mathbb{1}_{]0,A]}(x_i) \frac{1}{A} = A^{-n} \mathbb{1}_{\max_{i=1,\ldots,n} x_i \leq A}.$$

2. Comme $X_1 \sim \mathcal{U}([0, A[)$, $E_A(X_1) = A/2$. Un estimateur par moments $\hat{A}_{n,1}$ vérifie donc

$$\bar{X}_n = \frac{\hat{A}_{n,1}}{2},$$

et donc $\hat{A}_{n,1} = 2\bar{X}_n$.

3. Comme $E_A(|X_1|) = A/2 < +\infty$, la loi des grands nombres donne

$$\bar{X}_n \xrightarrow[n \to +\infty]{\mathbb{P}} E_A(X_1) = A/2,$$

et donc $\hat{A}_{n,1} \xrightarrow[n \to +\infty]{\mathbb{P}} A$. $\hat{A}_{n,1}$ est donc consistant.

4. Comme $E_A(X_1^2) \leq A^2 < +\infty$, le théorème central limite donne

$$\sqrt{n}(\bar{X}_n - A/2) \rightsquigarrow \mathcal{N}(0, \mathrm{Var}_A(X_1)),$$

avec $\mathrm{Var}_A(X_1) = A^2/12$. On en déduit que

$$\sqrt{n}(\hat{A}_{n,1} - A) \rightsquigarrow 2\mathcal{N}(0, A^2/12) \sim \mathcal{N}(0, A^2/3),$$

et donc que $\hat{A}_{n,1}$ est asymptotiquement normal.

5. Soit $A > 0$. On a

$$\begin{aligned} E_A((\hat{A}_{n,1} - A)^2) &= \mathrm{Var}_A(\hat{A}_{n,1}) + \Big(E_A(\hat{A}_{n,1}) - A\Big)^2 \\ &= \mathrm{Var}_A(\hat{A}_{n,1}) = \mathrm{Var}_A(2\bar{X}_n) = \frac{4}{n}\mathrm{Var}_A(X_1) = \frac{A^2}{3n}. \end{aligned}$$

6. On veut montrer que $A < 10^{10}$. Les hypothèses de test sont donc

$$\begin{cases} H_0 & : \quad A = 10^{10}, \\ H_1 & : \quad A < 10^{10}. \end{cases}$$

7. La loi de $\hat{A}_{n,1}$ étant peu accessible sous H_0, on va se baser sur l'inégalité de Markov. Sous H_1, on s'attend à ce que $\hat{A}_{n,1}$ soit petit. On choisit donc une zone de rejet de la forme $[0, t_{3\%}]$, où $t_{3\%}$ doit vérifier

$$P_{10^{10}}\left(\hat{A}_{n,1} \leq t_{3\%}\right) \leq 3\%.$$

En utilisant la question 5 et l'inégalité de Bienaymé Cebicev on obtient, pour $t > 0$,

$$P_{10^{10}}(\hat{A}_{n,1} \leq 10^{10} - t) \leq P_{10^{10}}(|\hat{A}_{n,1} - 10^{10}| \geq t) \leq \frac{10^{20}}{3nt^2}.$$

En posant $t_0 = \frac{10^{10}}{0.3\sqrt{n}}$, on obtient alors

$$P_{10^{10}}(\hat{A}_{n,1} \leq 10^{10} - t_0) \leq 3\%.$$

On en déduit que

$$T_1(X_{1:n}) = \mathbb{1}_{\hat{A}_{n,1} \leq 10^{10}(1 - 1/(0.3\sqrt{n}))}$$

est un test de niveau 3%.

8. Soit $x_1, \ldots, x_n > 0$. La vraisemblance s'écrit, pour $A > 0$,

$$V_{x_{1:n}}(A) = A^{-n}\mathbb{1}_{\max_{i=1,\ldots,n} x_i \leq A}.$$

Comme $V_{x_{1:n}}(A) = 0$ pour $A < \max_{i=1,\ldots,n} x_i$, on ne peut pas passer au logarithme. Cela dit, on a $V_{x_{1:n}}(A) = 0$ pour $A < \max_{i=1,\ldots,n} x_i$ et $V_{x_{1:n}}(A) = A^{-n}$ pour $A \geq \max_{i=1,\ldots,n} x_i$, il est donc immédiat que

$$\arg\max_{A>0} V_{x_{1:n}}(A) = \max_{i=1,\ldots,n} x_i.$$

On en déduit que $\hat{A}_{n,2} = \max_{i=1,\ldots,n} X_i$.

9. Soit $A \geq \varepsilon > 0$ (pour $\varepsilon > A$, $P_A(|\hat{A}_{n,2} - A| > \varepsilon) = 0$), et $A > 0$.

$$\begin{aligned} P_A(|\hat{A}_{n,2} - A| > \varepsilon) &= P_A(\hat{A}_{n,2} < A - \varepsilon) \\ &= P_A(\max_{i=1,\dots,n} X_i < A - \varepsilon) \\ &= \left(1 - \frac{\varepsilon}{A}\right)^n \underset{n \to +\infty}{\longrightarrow} 0. \end{aligned}$$

On en déduit que $\hat{A}_{n,2}$ est consistant.

10. Soit $t > 0$, et $A > 0$. D'après ce qui précède

$$\begin{aligned} P_A(n(A - \hat{A}_{n,2}) > t) &= P_A\left(\hat{A}_{n,2} < A - \frac{t}{n}\right) \\ &= \left(1 - \frac{t}{An}\right)^n \mathbb{1}_{t \leq nA} \\ &\underset{n \to +\infty}{\longrightarrow} \exp(-t/A). \end{aligned}$$

On en déduit que $n(A - \hat{A}_{n,2}) \rightsquigarrow \mathcal{E}(1/A)$.

11. Comme auparavant on s'attend à ce que $\hat{A}_{n,2}$ soit petit sous H_1. La forme de la zone de rejet est donc $[0, t_{3\%}]$, où $t_{3\%}$ doit vérifier

$$\limsup_{n \to +\infty} P_{10^{10}}(\hat{A}_{n,2} \leq t_{3\%}) \leq 3\%.$$

D'après la question précédente, pour $A = 10^{10}$ on a

$$n\left(1 - \frac{\hat{A}_{n,2}}{10^{10}}\right) \rightsquigarrow \mathcal{E}(1).$$

Si q est le quantile d'ordre 97% d'une loi $\mathcal{E}(1)$ (par le calcul $q = \log(1/3\%)$), on a alors

$$P_{10^{10}}\left(n\left(1 - \frac{\hat{A}_{n,2}}{10^{10}}\right) \leq q\right) \underset{n \to +\infty}{\longrightarrow} 97\%,$$

et donc

$$P_{10^{10}}\left(n\left(1 - \frac{\hat{A}_{n,2}}{10^{10}}\right) \geq q\right) cv 3\%.$$

Or,

$$n\left(1 - \frac{\hat{A}_{n,2}}{10^{10}}\right) \geq q \quad \Leftrightarrow \quad \hat{A}_{n,2} \leq 10^{10}(1 - q/n).$$

On en déduit que

$$T(X_{1:n}) = \mathbb{1}_{\hat{A}_{n,2} \leq 10^{10}(1-q/n)}$$

est un test de niveau asymptotique 3%.

12. On a $n = 40$ et $\hat{A}_{n,2} = 9 \times 10^9$. Par ailleurs $q \leq 4$, et donc

$$10^{10}(1 - q/n) > 10^{10}(1 - 4/40) = 9 \times 10^9.$$

On rejette donc H_0, et est sûr à 97% que l'univers est d'âge inférieur à 10^{10}.

13. Commençons par calculer la loi de $\hat{A}_{n,2}$. Pour $t \in [0, A[$,

$$P_A(\hat{A}_{n,2} \leq t) = (t/A)^n.$$

On en déduit que $\hat{A}_{n,2}$ a pour densité

$$\frac{nt^{n-1}}{A^n} \mathbb{1}_{]0,A]}(t).$$

On calcule alors

$$\begin{aligned} E_A(\hat{A}_{n,2}) &= \int_0^A \frac{nt^n}{A^n} = \frac{nA}{n+1}. \\ E_A(\hat{A}_{n,2}^2) &= \int_0^A \frac{nt^{n+1}}{A^n} = \frac{nA^2}{n+2}. \end{aligned}$$

On en déduit alors que

$$\begin{aligned} E_A((\hat{A}_{n,2} - A)^2) &= E_A(\hat{A}_{n,2}^2) - 2AE_A(\hat{A}_{n,2}) + A^2 \\ &= A^2 \left(\frac{n}{n+2} - \frac{2n}{n+1} + 1 \right) \\ &= \frac{2A^2}{(n+1)(n+2)}. \end{aligned}$$

Pour $A > 0$, on a $E_A((\hat{A}_{n,2} - A)^2) \leq E_A((\hat{A}_{n,1} - A)^2)$ équivaut à

$$(n+1)(n+2) \geq 6n \quad \Leftrightarrow n^2 - 3n + 2 \geq 0.$$

Les racines de ce polynôme étant 1 et 2, on en déduit que $E_A((\hat{A}_{n,2} - A)^2) \leq E_A((\hat{A}_{n,1} - A)^2)$ dès lors que $n \geq 2$ (c'est une inégalité stricte pour $n \geq 3$).

14. La loi de $\hat{A}_{n,1}$ étant peu simple à déterminer, on utilise l'inégalité de Bienaymé-Cevicev. Pour $t > 0$, on a

$$P_A\left(\left|\hat{A}_{n,1} - A\right| > t\right) \leq \frac{E_A((\hat{A}_{n,1} - A)^2)}{t^2} = \frac{A^2}{3nt^2}.$$

On pose alors $t_{5\%} = A/\sqrt{3n5\%}$, de telle sorte que

$$P_A\left(\left|\hat{A}_{n,1} - A\right| > t_{5\%}\right) \leq 5\%.$$

Or,

$$\begin{aligned} \left|\hat{A}_{n,1} - A\right| \leq t_{5\%} \quad &\Leftrightarrow \quad \hat{A}_{n,1} - A/\sqrt{3n5\%} \leq A \leq \hat{A}_{n,1} + A/\sqrt{3n5\%} \\ &\Leftrightarrow \quad \frac{\hat{A}_{n,1}}{1 + 1/\sqrt{3n5\%}} \leq A \leq \frac{\hat{A}_{n,1}}{1 - 1/\sqrt{3n5\%}}, \end{aligned}$$

pour $3n5\% < 1$. On en déduit que

$$I_1 = \left[\frac{\hat{A}_{n,1}}{1 + 1/\sqrt{3n5\%}} ; \frac{\hat{A}_{n,1}}{1 - 1/\sqrt{3n5\%}} \right]$$

est un intervalle de niveau de confiance 95%.

15. La loi de $\hat{A}_{n,2}$ étant connue (cf Question 9), on peut se baser dessus. On repart de

$$P_A(\hat{A}_{n,2} \leq t) = (t/A)^n.$$

Comme $\hat{A}_{n,2} \leq A$, cela donne, pour $t > 0$,

$$\begin{aligned} P_A(\hat{A}_{n,2} \notin [A-t, A]) &= P_A(\hat{A}_{n,2} \leq A - t) \\ &= \left(1 - \frac{t}{A}\right)^n \mathbb{1}_{t \leq A}. \end{aligned}$$

On pose $t_{5\%} = A - A(5\%)^{1/n}$, de sorte que

$$P_A(\hat{A}_{n,2} \notin [A - t_{5\%}, A]) = 5\%.$$

Par ailleurs,

$$A - t_{5\%} \leq \hat{A}_{n,2} \leq A \quad \Leftrightarrow \quad \hat{A}_{n,2} \leq A \leq \hat{A}_{n,2}(5\%)^{-1/n}.$$

On en déduit que

$$I_2 = \left[\hat{A}_{n,2}; \hat{A}_{n,2}(5\%)^{-1/n}\right]$$

est un intervalle de niveau de confiance 95%.

16. En notant L la fonction longueur, on a

$$\begin{aligned} L(I_1) &= \hat{A}_{n,1}\left((1 - 1/\sqrt{3n5\%})^{-1} - (1 - 1/\sqrt{3n5\%})^{-1}\right) \\ &= \hat{A}_{n,1}\left(1 + 1/\sqrt{3n5\%} - (1 - 1/\sqrt{3n5\%}) + O(1/n)\right) \\ &\sim_{n\to+\infty} 2A/\sqrt{3n5\%}. \\ L(I_2) &= \hat{A}_{n,2}((5\%)^{-1/n} - 1) \\ &= \hat{A}_{n,2}\,(1 + \log(1/5\%)/n + o(1/n) - 1) \\ &\sim_{n\to+\infty} A\log(1/5\%)/n. \end{aligned}$$

On en déduit que $L(I_2) =_{n\to+\infty} o(L(I_1))$, et donc que I_2 est beaucoup plus précis.

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^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.

Annexe: tables de quantiles

Fonction de répartition de la loi normale standard

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-x^2/2}\, dx$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

z	0.841	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
$\Phi(z)$	0.8000	0.9000	0.9500	0.9750	0.9800	0.9900	0.9950	0.9975	0.9990	0.9995

Loi Binomiale (suite)

$$P(X = k) = C_n^k p^k (1-p)^{n-k}$$

(k le nombre d'occurrences parmi n)

n = 50

k	p 0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
0	0,0769	0,0052	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,2794	0,0338	0,0029	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
2	0,5405	0,1117	0,0142	0,0013	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000
3	0,7604	0,2503	0,0460	0,0057	0,0005	0,0000	0,0000	0,0000	0,0000	0,0000
4	0,8964	0,4312	0,1121	0,0185	0,0021	0,0002	0,0000	0,0000	0,0000	0,0000
5	0,9622	0,6161	0,2194	0,0480	0,0070	0,0007	0,0001	0,0000	0,0000	0,0000
6	0,9882	0,7702	0,3613	0,1034	0,0194	0,0025	0,0002	0,0000	0,0000	0,0000
7	0,9968	0,8779	0,5188	0,1904	0,0453	0,0073	0,0008	0,0001	0,0000	0,0000
8	0,9992	0,9421	0,6681	0,3073	0,0916	0,0183	0,0025	0,0002	0,0000	0,0000
9	0,9998	0,9755	0,7911	0,4437	0,1637	0,0402	0,0067	0,0008	0,0001	0,0000
10	1,0000	0,9906	0,8801	0,5836	0,2622	0,0789	0,0160	0,0022	0,0002	0,0000
11	1,0000	0,9968	0,9372	0,7107	0,3816	0,1390	0,0342	0,0057	0,0006	0,0000
12	1,0000	0,9990	0,9699	0,8139	0,5110	0,2229	0,0661	0,0133	0,0018	0,0002
13	1,0000	0,9997	0,9868	0,8894	0,6370	0,3279	0,1163	0,0280	0,0045	0,0005
14	1,0000	0,9999	0,9947	0,9393	0,7481	0,4468	0,1878	0,0540	0,0104	0,0013
15	1,0000	1,0000	0,9981	0,9692	0,8369	0,5692	0,2801	0,0955	0,0220	0,0033
16	1,0000	1,0000	0,9993	0,9856	0,9017	0,6839	0,3889	0,1561	0,0427	0,0077
17	1,0000	1,0000	0,9998	0,9937	0,9449	0,7822	0,5060	0,2369	0,0765	0,0164
18	1,0000	1,0000	0,9999	0,9975	0,9713	0,8594	0,6216	0,3356	0,1273	0,0325
19	1,0000	1,0000	1,0000	0,9991	0,9861	0,9152	0,7264	0,4465	0,1974	0,0595
20	1,0000	1,0000	1,0000	0,9997	0,9937	0,9522	0,8139	0,5610	0,2862	0,1013
21	1,0000	1,0000	1,0000	0,9999	0,9974	0,9749	0,8813	0,6701	0,3900	0,1611
22	1,0000	1,0000	1,0000	1,0000	0,9990	0,9877	0,9290	0,7660	0,5019	0,2399
23	1,0000	1,0000	1,0000	1,0000	0,9996	0,9944	0,9604	0,8438	0,6134	0,3359
24	1,0000	1,0000	1,0000	1,0000	0,9999	0,9976	0,9793	0,9022	0,7160	0,4439
25	1,0000	1,0000	1,0000	1,0000	1,0000	0,9991	0,9900	0,9427	0,8034	0,5561
26	1,0000	1,0000	1,0000	1,0000	1,0000	0,9997	0,9955	0,9686	0,8721	0,6641
27	1,0000	1,0000	1,0000	1,0000	1,0000	0,9999	0,9981	0,9840	0,9220	0,7601
28	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9993	0,9924	0,9556	0,8389
29	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9997	0,9966	0,9765	0,8987
30	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9999	0,9986	0,9884	0,9405
31	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9995	0,9947	0,9675
32	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9998	0,9978	0,9836
33	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9999	0,9991	0,9923
34	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9997	0,9967
35	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9999	0,9987
36	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9995
37	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	0,9998
38	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000