ECE 2 - Année 2018-2019
Lycée français de Vienne
Mathématiques - F. Gaunard
http://frederic.gaunard.com

Chapitre 12. Estimation

1 Introduction

Dans tous les chapitres précédents concernant les probabilités, les lois des processus aléatoires étaient connues, et c'est à partir de ces lois que nous avons fait les calculs.

Adoptons alors une approche différente d'ordre pratique; on **ne connait pas** la loi de X_i. Par exemple la durée de vie d'un composant électronique, le nombre de clients se présentant au guichet d'une banque un jour donné, la proportion d'un certain caractère dans une population, le résultat d'un scrutin à partir des premiers dépouillements...

Il peut être raisonnable de penser qu'elle appartient à une famille de lois usuelles (ce qui définit le **modèle** de loi) qui dépend d'un paramètre (inconnu) qu'on aimerait donc **estimer** à partir du résultat $(x_1, x_2, ..., x_n)$ de n expériences.

☞ Il s'agit bien d'une estimation car on ne peut pas faire une infinité d'expérience de même qu'on ne peut pas interroger l'ensemble de la population.

Ainsi, le $n-$uplet $(x_1, x_2, ..., x_n)$ correspond alors à l'**observation** d'un **échantillon**, c'est à dire la réalisation pour une certaine issue $\omega \in \Omega$. On peut alors, à partir des observations, vouloir produire une valeur approchée du paramètre (par exemple de l'espérance), ce qui s'appelle une *estimation ponctuelle* ou vouloir fournir un intervalle qui contient le paramètre avec une très grande probabilité, ce qu'on appelle *estimation par intervalle de confiance*.

2 Modèle statistique. Estimation ponctuelle

2.1 Modèle statistique

Définition 1. *Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. On appelle $n-$échantillon de X un $n-$uplet $(X_1, ..., X_n)$ de v.a.i.i.d, de même loi que X.*

Soit $(X_1, ..., X_n)$ un $n-$échantillon de X. Pour tout $\omega \in \Omega$, on appelle réalisation du $n-$échantillon $(X_1, ..., X_n)$ le $n-$uplet $(X_1(\omega), ..., X_n(\omega))$.

☞ Par convention, on note souvent une réalisation $(X_1(\omega), ..., X_n(\omega))$ par des minuscules $(x_1, ..., x_n)$.

☞ Il faut bien distinguer le $n-$échantillon $(X_1, ..., X_n)$, qui est une variable aléatoire (et donc une fonction), de la réalisation $(X_1(\omega), ..., X_n(\omega))$ qui est un élément de $\mathbb{R}^n$.

Définition 2. *Soit Θ un sous-ensemble de $\mathbb{R}$. Un **modèle statistique** est un ensemble de lois $\mathcal{M}_\Theta$. Étant donnée une observation $(x_1, x_2, ..., x_n)$ de X, le but (du statisticien) est d'identifier la loi, parmi celles de $\mathcal{M}_\Theta$, ayant permis de la générer.*

Exemple. On dispose d'une pièce dont on ne connait pas la probabilité de tomber sur *Pile*, que l'on note p, et que l'on aimerait estimer. On considère alors une v.a. X (qui vaut 1 en cas de *Pile* et 0 sinon) et qu'on va bien entendu identifier parmi un ensemble de lois de Bernoulli, ici $\mathcal{M}_\Theta = \{\mathcal{B}(p); p \in [0; 1]\}$.

On lance cette pièce n fois et on note X_i la v.a. qui vaut 1 si la pièce tombe sur *Pile* au $i-$ème lancer et 0 sinon. le $n-$uplet $(X_1, X_2, ..., X_n)$ est un $n-$échantillon de X.

On décide de prendre $n = 10$. Le résultat des 10 lancers donne $(1, 1, 0, 0, 1, 1, 0, 0, 0, 1)$ qui représente donc une réalisation du $10-$échantillon $(X_1, X_2, ..., X_{10})$.

Il va donc falloir introduire un *estimateur*, c'est à dire une *fonction* de l'échantillon qui, appliquée à l'observation donnera une estimation du paramètre cherché.

2.2 Estimateurs. Biais. Risque quadratique

Définition 3. *Soit $(X_1, X_2, ..., X_n)$ un $n-$échantillon d'une v.a. X dont la loi dépend d'un paramètre θ, θ appartenant à une partie $\Theta \subset \mathbb{R}$. On appelle **estimateur** de θ toute variable aléatoire T_n de la forme*

$$T_n = \varphi(X_1, X_2, ..., X_n),$$

*où φ est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$,) valeurs dans Θ, éventuellement dépendante de n mais indépendante de θ. Toute réalisation d'un estimateur est appelée **estimation ponctuelle** de θ.*

Exemple. Reprenons l'exemple précédent. La variable aléatoire

$$\overline{X}_{10} = \frac{X_1 + X_2 + ... + X_{10}}{10}$$

est un estimateur de p. À partir de l'observation $(1, 1, 0, 0, 1, 1, 0, 0, 0, 1)$, $\overline{X}_{10}$ donne une estimation (ponctuelle) de $1/2$ pour p.

☞ Certains estimateurs visent à estimer non pas le paramètre mais une fonction $g(\theta)$ du paramètre.

Par exemple, en reprenant le contexte ci-dessus, on peut vouloir estimer la variance de X, à savoir la quantité $g(p) = p(1 - p)$, qui est bien une fonction du paramètre p. Les deux variables aléatoires ci-dessous alors sont des estimateurs de $g(p)$

$$\frac{1}{10}\sum_{i=1}^{10}(X_i - \overline{X}_{10})^2, \qquad \overline{X}_{10}(1 - \overline{X}_{10}).$$

☞ La notion d'estimateur parait alors très floue (ou très générale), dans le sens "destinée à approcher la valeur de θ" puisqu'en réalité, toute fonction des données est donc un estimateur. Cependant, tous ne possèdent pas la même performance, ce que nous verrons ensuite.

Définition 4. *Soit T_n un estimateur de $g(\theta)$ admettant une espérance $E_\theta(T_n)$. (pour tout $\theta \in \Theta$). On appelle **biais** de T_n le nombre*

$$b_\theta(T_n) = E_\theta(T_n) - g(\theta)$$

*c'est à dire l'écart "moyen" entre l'estimateur et le paramètre à estimer. Lorsque le biais est nul, on dit que T_n est **sans biais** ou **non biaisé**.*

Proposition 1. Soient X une v.a. d'espérance m et $(X_1, ..., X_n)$ un $n-$échantillon de X. Alors, la *moyenne empirique*

$$\overline{X}_n = \frac{1}{n}(X_1 + X_2 + ... + X_n)$$

est un estimateur sans biais de m.

Exercice 1. Soit $(X_1, X_2, ..., X_n)$ un $n-$échantillon d'une variable $X \in \mathcal{M}_\Theta = \{\mathcal{U}([0;\theta]); \theta \in \mathbb{R}_+\}$.

(1) Montrer que $\overline{X}_n$ est un estimateur biaisé et θ et préciser $b_\theta(\overline{X}_n)$.
(2) Proposer alors un estimateur V_n de θ sans biais, obtenu comme transformation simple de $\overline{X}_n$.
(3) On considère maintenant l'estimateur $M_n = \max(X_1, X_2, ..., X_n)$.
   (a) Déterminer la fonction de répartition F de M_n et en déduire une densité f de M_n.
   (b) Montrer alors que
$$E_\theta(M_n) = \frac{n\theta}{n+1}.$$
   (c) En déduire un estimateur sans biais Z_n à partir de M_n.

Exercice 2. (**Estimation de la variance**) Soit $(X_1, ..., X_n)$ un $n-$échantillon d'une v.a. X admettant pour espérance m et pour variance σ^2.

(1) On suppose que m est connu. Montrer que T_n est un estimateur non biaisé de σ^2, où
$$T_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2.$$
(2) On suppose que m n'est pas connu. On note $\overline{X}_n$ la moyenne empirique de l'échantillon et
$$U_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2.$$
   (a) Montrer que, pour tout $i \in [\![1;n]\!]$, $E((X_i - \overline{X}_n)^2) = V(X_i - \overline{X}_n)$.
   (b) Montrer que, pour tout $i \in [\![1;n]\!]$,
$$V(X_i - \overline{X}_n) = \left(1 - \frac{1}{n}\right)^2 V(X_i) + \frac{1}{n^2}\sum_{k\neq i} V(X_k).$$
   (c) En déduire que
$$V(X_i - \overline{X}_n) = \frac{n-1}{n}\sigma^2.$$
   (d) Montrer alors que U_n est un estimateur biaisé de σ^2. En déduire un estimateur sans biais de σ^2.

☞ Disposer d'un estimateur sans biais semble présenter un réel intérêt puisque son espérance est égale au paramètre cherché, mais ce seul fait ne garantit pas que l'estimateur fournit de bonnes approximations. Pour évaluer le défaut de moyenne, et juger de la qualité d'un estimateur, on calcule la moyenne des carrés des écarts au paramètre.

Définition 5. *Si T_n est un estimateur de $g(\theta)$ et admet un moment d'ordre 2 pour toute valeur de $\theta \in \Theta$, on appelle* ***risque quadratique*** *de T_n le réel*
$$r_\theta(T_n) = E_\theta((T_n - g(\theta))^2).$$
Si T_n et U_n sont deux estimateurs de $g(\theta)$, on dit que T_n est meilleur que U_n si et seulement si $r_\theta(T_n) \leq r_\theta(U_n)$ pour tout $\theta \in \Theta$.

☞ Le risque quadratique dépend (potentiellement) de θ et de n.

☞ $r_\theta(T_n)$ est toujours positif (ou nul). Dans le cas où $r_\theta(T_n) = 0$, alors $T_n = \theta$ presque sûrement. C'est donc l'estimateur parfait!

Proposition 2. Soit T_n un estimateur de $g(\theta)$ admettant espérance et variance (pour tout $\theta \in \Theta$). Alors,
$$r_\theta(T_n) = (b_\theta(T_n))^2 + V_\theta(T_n).$$

☞ Si T_n est un estimateur sans biais de $g(\theta)$, son risque quadratique est alors égal à sa variance et il suit que, parmi deux estimateurs sans biais, celui avec la plus petite variance est le meilleur.

Exercice 3. On reprend les notations de l'Exercice 1. C'est à dire que $X \in \mathcal{M}_\Theta = \{\mathcal{U}([0;\theta]); \theta \in \mathbb{R}_+\}$ et on a les deux estimateurs sans biais de θ

$$V_n = \frac{2}{n}\left(X_1 + X_2 + ... + X_n\right), \qquad Z_n = \frac{n+1}{n}\max(X_1, X_2, ..., X_n).$$

(1) Établir que

$$r_\theta(V_n) = V_\theta(V_n) = \frac{\theta^2}{3n}.$$

(2) Montrer que

$$E_\theta(Z_n^2) = \frac{(n+1)^2}{n^2}\int_0^\theta \frac{nt^{n+1}}{\theta^n}\mathrm{d}t.$$

En déduire que

$$r_\theta(Z_n) = V_\theta(Z_n) = \frac{\theta^2}{n(n+2)}$$

(3) Quel estimateur aura-t-on tendance à préférer en pratique?

2.3 Suite d'estimateurs

Définition 6. *Soit (T_n) une suite d'estimateurs. On dit que l'estimateur T_n est* ***asymptotiquement sans biais*** *si, pour tout $\theta \in \Theta$,*

$$b_\theta(T_n) \underset{n\to+\infty}{\longrightarrow} 0$$

ou bien, de manière équivalente,

$$\lim_{n\to+\infty} E_\theta(T_n) = g(\theta).$$

Exercice 4. Une fois de plus, on reprend le cadre de l'Exercice 1 où $X \hookrightarrow \mathcal{U}([0;\theta])$. Montrer que l'estimateur $M_n = \max(X_1, X_2, .., X_n)$ est asymptotiquement sans biais.

Définition 7. *Une suite d'estimateur (T_n) de $g(\theta)$ est dite convergente si, pour tout $\theta \in \Theta$,*

$$\forall \varepsilon > 0, \qquad \lim_{n\to+\infty} P_\theta(|T_n - g(\theta)| > \varepsilon) = 0.$$

☞ Par abus de langage on dit souvent simplement que l'estimateur T_n est convergent.

☞ Un estimateur T_n est convergent si, aussi petit qu'on choisisse $\varepsilon > 0$, la probabilité qu'une réalisation de T_n soit proche de $g(\theta)$ à ε près tend vers 1.

Théorème 1. *Soit T_n un estimateur tel quel, pour tout $\theta \in \Theta$, on ait*

$$\lim_{n\to+\infty} r_\theta(T_n) = 0.$$

Alors, T_n est un estimateur convergent.

Preuve. Soient $\theta \in \Theta$ et $n \in \mathbb{N}^*$. D'après l'inégalité de Markov, pour tout $\varepsilon > 0$, on a

$$P\left(|T_n - g(\theta)| > \varepsilon\right) = P\left(|\left(T_n - g(\theta)\right)^2| > \varepsilon^2\right) \leq \frac{r_\theta(T_n)}{\varepsilon^2} \underset{n\to+\infty}{\longrightarrow} 0.$$

□

Exercice 5. On considère la variable aléatoire X dont la loi est donnée par

$$P(X=-1) = p, \quad P(X=0) = 1-2p, \quad P(X=1) = p,$$

pour un certain paramètre $p \in \left]0; \frac{1}{2}\right[$. On dispose d'un $n-$échantillon $(X_1, ..., X_n)$ de X, et on cherche à déterminer le paramètre p.

(1) Calculer $E(\overline{X}_n)$. En déduire que l'estimateur $\overline{X}_n$ est biaisé.
(2) Peut-on trouver des réels a et b tels que $aX_n + b$ soit un estimateur sans biais de p?

(3) On note $T_n = \dfrac{1}{2n}\sum_{i=1}^{n} X_i^2$.
   (a) Montrer que T_n est un estimateur sans biais de p.
   (b) Calculer la variance $V_\theta(T_n)$.
   (c) Montrer que l'estimateur est convergent.

3 Estimation par intervalles de confiance

3.1 Motivation

À chaque estimation (observation d'un estimateur), correspond une valeur approchée, de précision non spécifiée, du paramètre θ. On peut vouloir préciser l'erreur commise (avec grande probabilité), c'est à dire déterminer un intervalle, basé sur l'observation, contenant θ avec une probabilité très élevée. C'est l'estimation par intervalle de confiance.

Par exemple, considérons un $n-$échantillon d'une loi de Bernoulli $X \hookrightarrow \mathcal{B}(p)$ dont on veut estimer p. On a pu voir précédemment que l'estimateur $T_n = \overline{X}_n$ est un estimateur non biaisé et convergent (son risque quadratique vaut $r_\theta(T_n) = p(1-p)/n)$ de p. À n fixé, l'inégalité de Bienaymé-Tchebychev donne alors

$$\begin{aligned} P\left(|T_n - p| > \varepsilon\right) \leq \frac{p(1-p)}{n\varepsilon^2} &\iff P\left(|T_n - p| \leq \varepsilon\right) \geq 1 - \frac{p(1-p)}{n\varepsilon^2} \\ &\iff P\left(T_n \in [p-\varepsilon; p+\varepsilon]\right) \geq 1 - \frac{p(1-p)}{n\varepsilon^2} \\ &\iff P\left(p \in [T_n-\varepsilon; T_n+\varepsilon]\right) \geq 1 - \frac{p(1-p)}{n\varepsilon^2} \end{aligned}$$

Considérons donné le problème suivant : étant donné un réel $\alpha \in]0;1[$ (appelé *niveau de confiance* ou seuil), déterminer un intervalle I_α (appelé *intervalle de confiance*) ne dépendant pas de p, contenant la vraie valeur de p avec probabilité supérieure à $1-\alpha$.

Par l'observation précédente, il suffit de déterminer ε tel que

$$1 - \frac{p(1-p)}{n\varepsilon^2} \geq 1-\alpha \iff \varepsilon \geq \sqrt{\frac{p(1-p)}{\alpha n}}.$$

On a alors

$$P\left(p \in \left[T_n - \sqrt{\frac{p(1-p)}{\alpha n}}; T_n + \sqrt{\frac{p(1-p)}{\alpha n}}\right]\right) \geq 1-\alpha.$$

Il apparaît alors un problème; l'intervalle censé encadrer p dépend lui-même de p. Mais qu'à cela ne tienne, on voit que $p \mapsto p(1-p)$ est majorée par $1/4$ sur $[0;1]$ ou encore que

$$\left[T_n - \sqrt{\frac{p(1-p)}{\alpha n}}; T_n + \sqrt{\frac{p(1-p)}{\alpha n}}\right] \subset \left[T_n - \sqrt{\frac{1}{4\alpha n}}; T_n + \sqrt{\frac{1}{4\alpha n}}\right]$$

et on peut donc proposer un encadrement

$$P\left(p \in \left[T_n - \sqrt{\frac{1}{4\alpha n}}; T_n + \sqrt{\frac{1}{4\alpha n}}\right]\right) \geq 1-\alpha.$$

3.2 Intervalles de confiance

Définition 8. *Soient $(X_1, ..., X_n)$ un $n-$échantillon de la loi X dépendant d'un paramètre θ, U_n et V_n deux estimateurs de $g(\theta)$ tels que, pour tous $n \in \mathbb{N}$ et $\theta \in \Theta$, $P(U_n \leq V_n) = 1$. Soit $\alpha \in [0;1]$.*

On dit que l'intervalle $[U_n, V_n]$ est un ***intervalle de confiance*** *au niveau $1-\alpha$ (ou au risque α) pour $g(\theta)$ si, pour tout $\theta \in \Theta$,*

$$P\left(g(\theta) \in [U_n, V_n]\right) \geq 1 - \alpha$$

ou de manière équivalente

$$P\left(g(\theta) \notin [U_n; V_n]\right) \leq \alpha.$$

☞ En pratique, le risque α étant donné, on utilise un estimateur non biaisé T_n de $g(\theta)$ et on cherche $\varepsilon > 0$ tel que $P(|T_n - g(\theta)| \leq \varepsilon) \geq 1 - \alpha$ à l'aide de l'inégalité de Bienaymé-Tchébychev ou du théorème central limite (approximation par la loi normale).

☞ Les bornes de l'intervalle de confiance **ne doivent jamais dépendre** du paramètre à estimer.

Exercice 6. Dernier retour sur l'Exercice 1. On considère à nouveau l'estimateur V_n du paramètre θ de la loi $x \in \mathcal{M}_\Theta = \{\mathcal{U}([0;\theta]),\ \theta \geq 0\}$.

(1) Montrer, à l'aide de l'inégalité de Bienaymé-Tchébychev, que

$$P(|V_n - \theta| > \varepsilon) \leq \frac{\theta^2}{3n\varepsilon^2}.$$

(2) Montrer que

$$\theta \in \left[V_n - \sqrt{\frac{\theta^2}{3n\alpha}}; V_n + \sqrt{\frac{\theta^2}{3n\alpha}}\right] \Longleftrightarrow \frac{V_n}{1 + \frac{1}{\sqrt{3n\alpha}}} \leq \theta \leq \frac{V_n}{1 - \frac{1}{\sqrt{3n\alpha}}}.$$

(3) En déduire un intervalle de confiance au risque α pour θ.

3.3 Intervalles de confiance asymptotiques

Définition 9. *Soient $(X_1, ..., X_n)$ un $n-$échantillon de la loi X dépendant d'un paramètre θ, U_n et V_n deux estimateurs de $g(\theta)$ tels que, pour tous $n \in \mathbb{N}$ et $\theta \in \Theta$, $P(U_n \leq V_n) = 1$. Soit $\alpha \in [0;1]$.*

On dit que l'intervalle $[U_n, V_n]$ est un ***intervalle de confiance asymptotique*** *pour $g(\theta)$ au niveau $1-\alpha$ (ou au risque α) si il existe une suite (α_n) de réels de $[0;1]$ vérifiant $\alpha_n \longrightarrow \alpha$, $n \to +\infty$, et telle que, pour tout $\theta \in \Theta$,*

$$P\left(g(\theta) \in [U_n, V_n]\right) \geq 1 - \alpha_n$$

ou de manière équivalente

$$\lim_{n \to +\infty} P\left(U_n \leq g(\theta) \leq V_n\right) \geq 1 - \alpha.$$

☞ Les intervalles de confiance asymptotiques sont toujours déduits d'une convergence en loi. En particulier, le théorème central limite est très utile lorsque l'estimateur T_n est la moyenne empirique.

Exercice 7. Afin d'avoir une idée du nombre de poissons N présents dans un étang, on en pêche une certaine quantité K, que l'on marque, puis que l'on remet à l'eau. On revient un jour plus tard, et l'on pêche n poissons (avec remise pour simplifier). Pour tout $i \in\in [\![1; n]\!]$, on note X_i la variable qui vaut 1 si le $i-$ième poisson pêché est marqué ou non. On suppose que ces variables sont indépendantes, et on introduit la moyenne empirique $\overline{X}_n$.

(1) Déterminer la loi de X_i puis calculer l'espérance et la variance de $\overline{X}_n$.

(2) Montrer que, si on note $p = K/N$, alors

$$\sqrt{\frac{n}{p(1-p)}}\left(\overline{X}_n - p\right) \xrightarrow{\mathcal{L}} X, \qquad X \hookrightarrow \mathcal{N}(0;1).$$

(3) Déterminer alors un intervalle de confiance asymptotique de risque α pour $p = K/N$ basé sur $\overline{X}_n$, puis en déduire un intervalle de confiance asymptotique de même risque pour N basé sur $\overline{X}_n$.

(4) *Application numérique.* $K = 50,\ n = 300,\ X_n = 0.6, \alpha = 0.05$.

3.4 Paramètre d'une Bernoulli. Comparaison des intervalles de confiance

On considère un $n-$échantillon $(X_1, ..., X_n)$ d'une loi de Bernoulli $X \hookrightarrow \mathcal{B}(p)$. On cherche un intervalle de confiance pour p au risque α. On sait que $T_n = \overline{X}_n$ est un estimateur non biaisé de p.

- Comme vu précédemment, l'inégalité de Bienaymé-Tchébychev fournit l'intervalle de confiance cherché
$$P\left(p \in \left[T_n - \sqrt{\frac{1}{4\alpha n}}; T_n + \sqrt{\frac{1}{4\alpha n}}\right]\right) \geq 1 - \alpha.$$
- On peut aussi procéder avec le théorème central limite, ce qu'on a davantage tendance à faire. En effet, ce dernier permet d'affirmer que
$$T_n^* = \sqrt{\frac{n}{p(1-p)}}\,(T_n - p) \xrightarrow{\mathcal{L}} Z, \quad Z \hookrightarrow \mathcal{N}(0;1).$$
Utilisant encore que $p(1-p) \leq 1/4$ (et donc que $\sqrt{n/p(1-p)} \geq 2\sqrt{n}$), et notant Φ la fonction de répartition de la loi normale centrée réduite, on a
$$\begin{array}{rcl} P\left(|T_n - p| \leq \varepsilon\right) & = & P\left(|T_n^*| \leq \varepsilon\sqrt{\dfrac{n}{p(1-p)}}\right) \\ & \leq & P\left(|T_n^*| \leq 2\varepsilon\sqrt{n}\right) \\ & \underset{n\to+\infty}{\longrightarrow} & 2\Phi\left(2\varepsilon\sqrt{n}\right) - 1. \end{array}$$
Il suffit donc de choisir ε de sorte que
$$2\Phi\left(2\varepsilon\sqrt{n}\right) - 1 \geq 1 - \alpha \Longleftrightarrow \Phi\left(2\varepsilon\sqrt{n}\right) - 1 \geq 1 - \frac{\alpha}{2} \Longleftrightarrow \varepsilon \geq \frac{1}{2\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right).$$
On obtient alors un autre intervalle de confiance asymptotique au risque α
$$\lim_{n\to+\infty} P\left(p \in \left[T_n - \frac{1}{2\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right); T_n + \frac{1}{2\sqrt{n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right]\right) \geq 1 - \alpha.$$

Examinons en détails les avantages et inconvénients de ces deux intervalle: ☞ Clairement, l'avantage du premier intervalle par rapport au deuxième est son caractère exact, c'est-à-dire non asymptotique.

☞ Cependant, l'intervalle asymptotique possède une étendue plus restreinte, et donne donc une estimation plus précise de p.

Lequel préférer? On voit que, pour de petites valeurs de n, l'intervalle asymptotique fournira des bornes trop restreintes, et que l'inégalité
$$P(g(\theta) \in [U_n; V_n]) \geq 1 - \alpha$$
ne sera pas du tout respectée. La question se reformule donc mieux ainsi : à partir de quelle valeur de n est-il raisonnable d'échanger le caractère exact de l'intervalle de confiance contre un intervalle plus réduit? On considère qu'en pratique, on doit au moins avoir $n > 20$.

✎ La majoration de la quantité $p(1-p)$ par $1/4$ (sur l'intervalle $[0;1]$) est très pratique et très souvent utilisée, parfois sans rappel ni indication, dans les problèmes de concours.

Exercice 8. Dans une population, des électeurs doivent choisir parmi 2 candidats, Emmanuel Maquereau et François Filou, le futur président. On note p la proportion d'électeurs désirant voter pour François Filou. On choisit un échantillon $(X_1, ..., X_{100})$ où l'on a noté $X_i = 1$ si la personne a voté pour Filou, et 0 sinon. Parmi ces personnes, 55% déclarent vouloir voter pour François. Peut-on déclarer au risque $\alpha = 5\%$ que François Filou sera élu président? On utilisera le théorème central limite.

4 Autres exercices

Exercice 1201. Soit a un réel strictement positif. On note f la fonction définie sur $\mathbb{R}$ par

$$f(x) = \begin{cases} 0, & \text{si } x < a \\ \dfrac{3a^3}{x^4}, & \text{si } x \geq a \end{cases}$$

(1) Montrer que f est une densité de probabilité.

Un capteur mesure **en permanence** le taux de gaz carbonique émis par un moteur. On suppose que le temps écoulé entre le démarrage du moteur et l'instant précis (en heures) où son taux de gaz carbonique devient non réglementaire est une variable aléatoire T de densité f.

(2) Montrer que T admet une espérance et une variance de valeurs:

$$E(T) = \frac{3a}{2}, \quad V(t) = \frac{3a^2}{4}.$$

(3) (a) Déterminer la fonction de répartition de T.
(b) Calculer les probabilités $P(T > 2a)$ et $P_{(T>2a)}(T > 6a)$.

(4) On met en route n moteurs de modèle identique au précédent, et indépendants. On note $T_1, T_2, ..., T_n$ les temps respectifs pendant lesquels ces moteurs ont un taux de gaz carbonique réglementaire ($T_1, T_2, ..., T_n$ suivent donc la même loi que T et sont indépendantes).

(a) Montrer que la variable

$$Z_n = \frac{2}{3n} \sum_{k=1}^{n} T_k$$

est un estimateur sans biais du paramètre a.
(b) Calculer son risque quadratique $r(Z_n)$. En déduire que T_n est un estimateur convergent.

Exercice 1202. Soit (X_k) une suite de v.a.i.i.d. suivant une loi de Poisson $\mathcal{P}(\lambda)$, où $\lambda > 0$. On pose, pour tout $n \in \mathbb{N}^*$, $S_n = X_1 + X_2 + ... + X_n$.

(1) (a) Montrer que $S_2 \hookrightarrow \mathcal{P}(2\lambda)$.
(b) En déduire, par récurrence sur $n \in \mathbb{N}^*$, que $S_n \hookrightarrow \mathcal{P}(n\lambda)$.

(2) Pour $n \in \mathbb{N}$, $n \geq 2$, on pose, $Y_n = \left(1 - \dfrac{1}{n}\right)^{S_n}$.

(a) Montrer que Y_n est une v.a. discrète à valeurs dans $]0; 1]$.
(b) Déterminer la loi de Y_n, son espérance et sa variance.
(c) En déduire un estimateur sans biais de $e^{-\lambda}$.

Exercice 1203. La sécurité routière fait une enquête sur le nombre d'accidents survenus par semaine sur un tronçon d'autoroute. Soit X la v.a. égale au nombre d'accidents en une semaine. On suppose que $X \in \mathcal{M}_\Theta = \{\mathcal{P}(\lambda),\ \lambda > 0\}$. On se propose d'estimer le paramètre $e^{-\lambda} = P(X = 0)$. On note $(X_1, ..., X_n)$ un $n-$échantillon de X.

(1) Soit Y_n le nombre de fois où l'on a pas observé d'accident pendant la semaine, *i.e.*

$$Y_n = \#\{i \in [\![1; n]\!];\ X_i = 0\}.$$

(a) Montrer que Y_n/n est un estimateur non biaisé de $e^{-\lambda}$.
(b) Déterminer son risque quadratique, noté ici $r(Y_n/n)$.

(2) Vérifier que $\overline{X}_n$ est un estimateur sans biais de λ.

(3) On pose $S_n = X_1 + ... + X_n$. Quelle est la loi de S_n? À l'aide du théorème de transfert, déterminer l'espérance de $e^{-\overline{X}_n}$. En déduire que $e^{-\overline{X}_n}$ est un estimateur biaisé de $e^{-\lambda}$. Est-il asymptotiquement sans biais?

Exercice 1204. (D'après **EDHEC 2014**)

Dans cet exercice, θ désigne un réel strictement positif et n un entier naturel supérieur ou égal à 2. Pour tout k de $\mathbb{N}$, on pose

$$u_k = \frac{1}{1+\theta}\left(\frac{\theta}{1+\theta}\right)^k.$$

(1) Montrer que la suite (u_k) définit bien une loi de probabilité.

On considère maintenant une variable aléatoire X prenant ses valeurs dans $\mathbb{N}$ et dont la loi est donnée par

$$\forall k \in \mathbb{N}, \qquad P(X=k) = u_k.$$

(2) (a) On pose $Y = X + 1$. Reconnaître la loi de Y et en déduire l'espérance et la variance de X.

(b) Compléter la fonction `SciLab` suivante pour qu'elle simule la loi d'une variable aléatoire X:

```
function x=X(theta)
    Y=1;
    while.....................
        Y=Y+1;
    end
    x=.....................
endfunction
```

(3) Dans cette question, on souhaite estimer le paramètre θ par la méthode du *maximum de vraisemblance*. Pour ce faire, on considère un échantillon $(X_1, X_2, \ldots, X_n)$ composé de variables aléatoires indépendantes ayant toutes la même loi que X et on introduit la fonction L, de $\mathbb{R}_+^*$ dans $\mathbb{R}$, définie par

$$\forall \theta \in \mathbb{R}_+^*, \quad L(\theta) = \prod_{k=1}^{n} P(X_k = x_k),$$

où $x_1, x_2, \ldots, x_n$ désignent des entiers naturels éléments de $X(\Omega)$. L'objectif est de choisir la valeur de θ qui rend $L(\theta)$ maximale.

(a) Écrire $\ln(L(\theta))$ en fonction de θ et de $S_n = \sum_{k=1}^{n} x_k$.

(b) On considère la fonction φ, définie par

$$\forall \theta \in]0;+\infty[, \qquad \varphi(\theta) = S_n \ln\theta - (S_n + n)\ln(1+\theta).$$

Montrer que la fonction φ admet un maximum, atteint en un seul réel que l'on notera $\hat{\theta}_n$ et que l'on exprimera en fonction de S_n. Que représente $\hat{\theta}_n$ pour la fonction L?

On pose dorénavant

$$T_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

La variable T_n est appelée *estimateur du maximum de vraisemblance* pour θ.

(c) Vérifier que T_n est un estimateur sans biais de θ.

(d) Calculer le risque quadratique $r_{T_n}(\theta)$ de T_n et vérifier que $\lim\limits_{n\to+\infty} r_{T_n}(\theta) = 0$.

Exercice 1205. (D'après **ESC 2009**)

Soit θ un réel strictement positif. On considère la fonction f définie sur $\mathbb{R}$ par

$$f(x) = \begin{cases} e^{\theta - x}, & \text{si } x \geq \theta \\ 0, & \text{si } x < \theta \end{cases}$$

(1) (a) Vérifier que, pour tout réel $A \geq \theta$,

$$\int_{\theta}^{A} f(x)\mathrm{d}x = 1 - e^{\theta - A}.$$

(b) Montrer que f est une densité.

On note X une variable aléatoire réelle de densité f.

(2) Déterminer la fonction de répartition de X.

(3) On considère la variable aléatoire $Y = X - \theta$.

(a) Montrer que la fonction de répartition F_Y de Y est définie par

$$F_Y(x) = \begin{cases} 1 - e^{-x}, & \text{si } x \geq 0 \\ 0, & \text{si } x < 0 \end{cases}$$

(b) En déduire que Y est une variable à densité qui suit une loi classique dont on précisera le paramètre. Préciser son espérance et sa variance.

(c) En déduire l'espérance et la variance de X.

(4) Dans toute la suite, n désigne un entier naturel non nul et $X_1, X_2, \ldots, X_n$ des variables aléatoires mutuellement indépendantes de même loi que X.

On cherche à estimer le réel θ à l'aide de la variable aléatoire $S_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n}(X_k - 1)$.

(a) Montrer que S_n est un estimateur sans biais de θ.

(b) Calculer son risque quadratique noté $r(S_n)$.

Exercice 1206. (D'après **ESC 2006**)

Dans cet exercice R désigne un réel fixé strictement positif et on considère la fonction f définie sur $\mathbb{R}$ par

$$f(t) = \begin{cases} 0, & \text{si } t \notin [0\,;\,R] \\ \frac{2t}{R^2}, & \text{si } t \in [0\,;\,R] \end{cases}$$

(1) (a) Étudier la continuité de f.

(b) Montrer que f est une densité de probabilité.

On note dans toute la suite X une variable aléatoire réelle de densité f et F_X désigne sa fonction de répartition.

(2) (a) Déterminer la valeur $F_X(x)$ lorsque $x < 0$, puis lorsque $x > R$.

(b) Montrer que pour tout réel x de $[0\,;\,R]$,

$$F_X(x) = \frac{x^2}{R^2}.$$

(3) Montrer que X admet une espérance et une variance et que

$$E(X) = \frac{2R}{3}, \qquad V(X) = \frac{R^2}{18}.$$

Dans toute la suite n désigne un entier naturel non nul et $(X_1, ..., X_n)$ un $n-$échantillon de X visant à estimer le réel R. On note

$$T_n = \frac{3}{2n}\sum_{k=1}^{n} X_k.$$

(3) Montrer que T_n est un estimateur sans biais de R et calculer son risque quadratique noté $r(T_n)$.

(4) On note $M_n = \max(X_1, ..., X_n)$.
   (a) Montrer que pour tout réel x , $\mathrm{P}(M_n \leq x) = (F_X(x))^n$. En déduire la fonction de répartition de M_n, puis montrer que M_n est une variable aléatoire à densité.
   (b) Montrer qu'une densité possible de M_n est la fonction g_n définie sur $\mathbb{R}$ par :

$$g_n(t) = \begin{cases} \dfrac{2nt^{2n-1}}{R^{2n}}, & \text{si } t \in [0\,;\, R] \\ \\ 0, & \text{si } t \notin [0\,;\, R] \end{cases}$$

   (c) Montrer que M_n admet une espérance et une variance, et que:

$$E(M_n) = \frac{2n}{2n+1}R \qquad \text{et} \qquad V(M_n) = \frac{n}{(n+1)(2n+1)^2}R^2.$$

   (d) Calculer le biais de M_n, noté $b(M_n)$, et son risque quadratique noté $r(M_n)$.

(5) (a) Déterminer un équivalent simple lorsque n tend vers $+\infty$ de $b(M_n)$ et $r(M_n)$.
   (b) Quels sont les avantages et les inconvénients réciproques des estimateurs T_n et M_n ?

Exercice 1207. Soit $(X_1, \ldots, X_n)$ un n-échantillon i.i.d de loi parente la loi normale $\mathcal{N}(m, \sigma^2)$ de variance σ^2 connue. *i.i.d signifie indépendantes et identiquement distribuées.*
On cherche un intervalle de confiance pour l'espérance m et on considère pour cela l'estimateur $T_n = \overline{X_n}$.

(1) Montrer que T_n est un estimateur sans biais et convergent de m.

(2) **Utilisation de l'inégalité de Bienaymé-Tchebychev**.
   (a) Montrer, à l'aide de l'inégalité de BT, que pour tout $t > 0$, on a

$$P(T_n - t \leq m \leq T_n + t) \geq 1 - \frac{\sigma^2}{nt^2}.$$

   (b) En déduire que l'intervalle

$$I_n(\alpha) = \left[T_n - \frac{\sigma}{\sqrt{n\alpha}}; T_n + \frac{\sigma}{\sqrt{n\alpha}}\right]$$

   est un intervalle de confiance de m au niveau de confiance $1-\alpha$.

(3) **Utilisation de la loi normale centrée réduite**.
   (a) Montrer que T_n suit la loi normale $\mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$.
   (b) En déduire que pour tout $t > 0$

$$P\left(\left|\sqrt{\frac{n}{\sigma^2}}(T_n - m)\right| \leq t\right) = 2\Phi(t) - 1.$$

   (c) Justifier que pour tout $\alpha \in]0; 1[$, il existe un réel t_α tel que : $2\Phi(t) - 1 = 1 - \alpha$.
   (d) En déduire que l'intervalle

$$J_n(\alpha) = \left[T_n - \frac{t_\alpha \sigma}{\sqrt{n}}; T_n + \frac{t_\alpha \sigma}{\sqrt{n}}\right]$$

   est un intervalle de confiance de m au niveau de confiance $1-\alpha$.

(4) (a) Déterminer l'étendue des intervalles $I_n(\alpha)$ et $J_n(\alpha)$.

(b) La commande SciLab `cdfnor("X",0,1,1-a,a)` renvoie la valeur de x vérifiant $\Phi(x) = 1-a$. On considère alors le programme suivant :

```
R = zeros(1,10)
T = zeros(1,10)
i=1
for a = [0.01 : 0.01 : 0.1] do
    R(i) = 1/sqrt(a)
    T(i) = cdfnor("X",0,1,1-a/2,a/2)
    i=i+1
end
```

qui renvoie les données suivantes :

```
R  =  10.  7.071  5.773  5.  4.472  4.082  3.779  3.535  3.333  3.162
T  =  2.575  2.326  2.170  2.053  1.959  1.880  1.811  1.750  1.695  1.644
```

Commenter ce programme (notamment la commande `cdfnor("X",0,1,1-a/2,a/2)`) puis, au vu de ces résultats, déterminer quel est l'intervalle le plus précis entre $I_n(\alpha)$ et $J_n(\alpha)$?

(5) On suppose que le temps de travail hebdomadaire d'un étudiant en classe prépa est une variable aléatoire suivant la loi normale $\mathcal{N}(m, 100)$ de moyenne m inconnue que l'on cherche à estimer. Une enquête réalisée auprès de 100 étudiants donne un temps de travail hebdomadaire moyen de 29.32h.

(a) Déterminer l'intervalle de confiance $J_n(0.05)$ au niveau de confiance $\alpha = 0.05$ pour la moyenne m.

(b) Quelle valeur de n faut-il choisir pour que la longueur de l'intervalle de confiance au niveau de confiance $\alpha = 0.05$ soit inférieure à 2 ?

Exercice 1208. Soit $n \in \mathbb{N}^*$. On considère un échantillon $(X_1, \ldots, X_n)$ de la loi normale $\mathcal{N}(m, m/5)$, avec $m > 0$, le paramètre m étant inconnu.

Soit $\alpha \in]0;1[$ et t_α le réel positif tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$. On suppose que $n > \frac{t_\alpha^2}{25}$.

On note

$$\overline{X_n} = \frac{X_1 + \ldots + X_n}{n}, \qquad \text{et} \qquad Y_n = 5\sqrt{n}\frac{\overline{X_n} - m}{m}.$$

(1) Justifier que la variable aléatoire Y_n converge en loi vers une loi normale centrée réduite.

(2) Justifier alors que pour n assez grand, on peut écrire

$$P\left(|Y_n| \leq t_\alpha\right) = 1 - \alpha.$$

(3) Montrer que l'intervalle

$$\left[5\sqrt{n}\frac{\overline{X_n}}{5\sqrt{n} + t_\alpha}; 5\sqrt{n}\frac{\overline{X_n}}{5\sqrt{n} - t_\alpha}\right]$$

est un intervalle de confiance de m au niveau de confiance $1 - \alpha$.

(4) Pour $n = 100$, une réalisation de ce n-échantillon nous donne une moyenne empirique de 12. Déterminer une estimation d'un intervalle de confiance de m à 95%.
On donne $\Phi(1,96) = 0,975$.

Exercice 1209. (D'après **HEC 2015**)

Soit λ un paramètre réel strictement positif, **inconnu**. Pour $n \in \mathbb{N}^*$, on considère un n-échantillon $(X_1, X_2, ..., X_n)$ de variables aléatoires à valeurs strictement positives, indépendantes, de même loi exponentielle de paramètre λ.

On pose, pour tout $n \in \mathbb{N}^*$,

$$S_n = \sum_{k=1}^{n} X_k, \quad \text{et} \quad J_n = \lambda S_n.$$

(1) Calculer, pour tout $n \in \mathbb{N}^*$, $E(S_n)$, $V(S_n)$, $E(J_n)$ et $V(J_n)$.

(2) On admet qu'une densité f_{J_n} de J_n est donnée par la formule

$$f_{J_n}(x) = \begin{cases} \frac{e^{-x}x^{n-1}}{(n-1)!}, \text{ si } x > 0 \\ 0, \text{ si } x \leq 0 \end{cases}.$$

(a) À l'aide du théorème de transfert, établir pour tout entier $n \geq 3$, l'existence de $E\left(\frac{1}{J_n}\right)$ et de $E\left(\frac{1}{J_n^2}\right)$, et donner leur valeurs respectives.

(b) On pose, pour tout entier $n \geq 3$,

$$\widehat{\lambda_n} = \frac{n}{S_n}.$$

Justifier que $\widehat{\lambda_n}$ est un estimateur de λ. Est-il sans biais? Calculer la limite, lorsque n tend vers $+\infty$, du risque quadratique associé à $\widehat{\lambda_n}$ en λ.

(3) Dans cette question, on veut déterminer un intervalle de confiance du paramètre λ au risque α. On note Φ la fonction de répartition de la loi normale centrée réduite, et u_α le réel strictement positif tel que

$$\Phi(u_\alpha) = 1 - \frac{\alpha}{2}.$$

(a) Enoncer le théorème de la limite centrée. En déduire que la variable aléatoire N_n définie par $N_n = \lambda \frac{S_n}{\sqrt{n}} - \sqrt{n}$ converge en loi vers la loi normale centrée réduite.

(b) En déduire que pour n assez grand, on a approximativement

$$P([-u_\alpha \leq N_n \leq u_\alpha]) = 1 - \alpha.$$

(c) Montrer que pour n assez grand, l'intervalle

$$\left[\left(1 - \frac{u_\alpha}{\sqrt{n}}\right)\widehat{\lambda_n}, \left(1 + \frac{u_\alpha}{\sqrt{n}}\right)\widehat{\lambda_n}\right]$$

est un intervalle de confiance de λ au risque α. On note λ_0 la réalisation de $\widehat{\lambda_n}$ sur le n-échantillon.

(4) Avec le n-échantillon $(X_1, X_2, ..., X_n)$, on construit un nouvel intervalle de confiance de λ au risque β ($\beta \neq \alpha$), tel que la longueur de cet intervalle soit k ($k > 1$) fois plus petite que celle obtenue avec le risque α.

(a) Justifier l'existence de la fonction réciproque Φ^{-1} de Φ. Quel est le domaine de définition de Φ^{-1} ?

(b) Établir l'égalité

$$\beta = 2\Phi\left(\frac{1}{k}\Phi^{-1}(\alpha/2)\right).$$

(c) En déduire que $\beta > \alpha$. Ce dernier résultat était-il prévisible?

5 Appendice: Table de la loi normale centrée réduite $\mathcal{N}(0,1)$

$$\phi(t)(x) = P(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \mathrm{d}t \quad \text{et} \quad \phi(-x) = 1 - \phi(x).$$

x	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
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000