Chapitre 13. Convergence(s) et approximations

Dans tout le chapitre on ne s'intéresse qu'à des variables aléatoires réelles discrètes ou bien à densité.

1 Inégalités probabilistes

On donne dans ce cours - et on démontre - deux inégalités classiques, bien que grossières, qui permettent d'obtenir des renseignements sur la concentration des valeurs prises par une variable aléatoire, en particulier autour de sa moyenne. Ces inégalités nous serviront principalement à démontrer les résultats de convergences en probabilités qui suivront.

1.1 Inégalité de Markov

Propriété

Inégalité de Markov.

Soit X une v.a. **positive** admettant une espérance. Alors,

$$\forall t > 0, \qquad P(X \geq t) \leq \frac{\mathrm{E}(X)}{t}$$

☞ On sait déjà que $t \mapsto P(X \geq t)$ est décroissante. Cette inégalité permet de voir que la décroissance se fait à une vitesse d'au moins $1/t$.

☞ Cette inégalité, bien que souvent très utile, reste très grossière pour estimer les *queues de probabilités*. Par exemple, si $X \hookrightarrow \mathcal{E}(\lambda)$, on sait que $P(X \geq t) = e^{-\lambda t} = o(1/\lambda t)$.

Remarque

En combinant l'inégalité de Markov (appliquée à $|X|^r$) avec la croissance de la fonction $t \mapsto t^r$ sur $\mathbb{R}_+$, on voit que, si X admet un moment d'ordre r, alors

$$\forall t > 0 \qquad P(|X| \geq t) \leq \frac{\mathrm{E}(|X|^r)}{t^r}.$$

☞ Ainsi, plus la variable admet des moments d'ordres élevés, plus les queues de probabilités décroissent vite.

Preuve. On distingue deux cas: si X est une variable aléatoire discrète et si X est à densité. Dans les deux cas, $X(\Omega) \subset \mathbb{R}_+$ ce qui permet les minorations qui vont apparaître. Soit $t > 0$.

- Si X est une variable aléatoire discrète, alors

$$\begin{aligned} E(X) &= \sum_{k\in X(\Omega)} kP(X=k) = \sum_{k<t} kP(X=k) + \sum_{k\geq t} kP(X=k) \\ &\geq \sum_{k\geq t} kP(X=k) \qquad\qquad (\text{car } k\geq 0) \\ &\geq t\sum_{k\geq t} P(X=k) = tP(X\geq t) \end{aligned}$$

et on a bien

$$P(X\geq t)\leq \frac{E(X)}{t}.$$

- Si X est une variable à densité et que f est une densité de X alors, comme X est à valeurs positives, $f(x)=0$ si $x<0$. Il suit que

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} xf(x)\mathrm{d}x = \int_0^t xf(x)\mathrm{d}x \\ &= \int_0^t xf(x)\mathrm{d}x + \int_t^{+\infty} xf(x)\mathrm{d}x \\ &\geq \int_t^{+\infty} xf(x)\mathrm{d}x \\ &\geq t\int_t^{+\infty} f(x)\mathrm{d}x = tP(X\geq t) \end{aligned}$$

et on a bien

$$P(X\geq t)\leq \frac{E(X)}{t}.$$

□

1.2 Inégalité de Bienaymé-Tchebychev

Propriété

Inégalité de Bienaymé-Tchebychev.

Soit X une v.a. admettant une variance (finie). Alors,

$$\forall t>0,\quad P\left(|X-E(X)|\geq t\right)\leq \frac{V(X)}{t^2}.$$

☞ Dans cette inégalité, on considère la variable *centrée* $X-E(X)$. On s'intéresse donc à la probabilité de déviation de X vis-à-vis de son espérance. Le résultat n'a, naturellement, d'intérêt que pour t grand.

☞ Nous savons que la variance quantifie l'étalement d'une variable aléatoire autour de son espérance. Ce théorème donne une majoration précise de la probabilité de déviation de cette espérance; X dévie de son espérance de plus de t avec une probabilité décroissante à vitesse $1/t^2$ (avec un certain coefficient multiplicatif: la variance de X).

Preuve. Il suffit d'appliquer l'inégalité de Markov à la variable $(X-E(X))^2$ (qui admet bien une espérance car X admet une variance). On a alors, par croissance de $t\mapsto t^2$ sur $\mathbb{R}_+$,

$$P(|X-E(X)|\geq t) = P((X-E(X)^2)\geq t^2)\leq \frac{E\left((X-E(X))^2\right)}{t^2} = \frac{V(X)}{t^2}.$$

□

Exercice 1. (Un théorème puissant mais parfois grossier)

(1) Soit $X \hookrightarrow \mathcal{U}([0;1])$.
   (a) Rappeler la valeur de $E(X)$ et de $V(X)$. Que vaut $P(|X - E(X)| > t)$ pour $t \geq 1/2$?
   (b) Exprimer l'inégalité de Bienaymé-Tchebychev. Commenter.

(2) Soit $X \hookrightarrow \mathcal{E}(1)$.
   (a) Rappeler la valeur de $E(X)$ et de $V(X)$. Que vaut $P(|X - E(X)| > t)$ pour $t > 1$?
   (b) Exprimer l'inégalité de Bienaymé-Tchebychev. Commenter.

Exercice 2. Soient $X_1, X_2, ..., X_n$ n v.a. (mutuellement) indépendantes de même loi $\mathcal{B}(p)$. Montrer que

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - p\right| \geq t\right) \leq \frac{p(1-p)}{nt^2}.$$

Exercice 3. On lance plusieurs fois une pièce parfaitement équilibrée. Les lancers sont indépendants. Combien de lancers faut-il effectuer pour pouvoir affirmer, avec un risque d'erreur inférieur à 5%, que la fréquence d'apparition du *Pile* au cours de ces lancers sera comprise entre 49% et 51%?

2 La loi faible des grands nombres

Le théorème suivant permet d'observer *a posteriori* la cohérence du modèle probabiliste avec l'approche empirique et intuitive de la notion de probabilité.

Propriété

Loi faible des grands nombres.

Soit (X_n) une suite de v.a. (mutuellement) indépendantes, admettant une même espérance m et une même variance σ^2. Soit

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

Alors,

$$\forall \varepsilon > 0, \qquad \lim_{n\to+\infty} P\left(|\overline{X}_n - m| \geq \varepsilon\right) = 0 \quad \text{et} \quad \lim_{n\to+\infty} P\left(|\overline{X}_n - m| < \varepsilon\right) = 1.$$

☞ Ce théorème signifie donc moralement que la moyenne de n variables aléatoires indépendantes suivant une même loi *converge* vers leur espérance commune.

Preuve. Commençons par observer que (d'après la linéarité de l'espérance et les propriétés de la variance vues dans les chapitres précédents),

$$E(\overline{X}_n) = m, \qquad V(\overline{X}) = \frac{\sigma^2}{n}.$$

L'inégalité de Bienaymé-Tchebychev donne alors

$$P\left(|\overline{X}_n - m| \geq \varepsilon\right) = P\left(|\overline{X}_n - E(\overline{X}_n| \geq \varepsilon\right) \leq \frac{V(\overline{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \underset{n\to+\infty}{\longrightarrow} 0.$$

□

Remarque

On peut voir (avec l'exercice en annexe) que le résultat reste vrai en ne supposant que l'existence des moments d'ordre 1, mais la preuve est beaucoup plus difficile et ce cas particulier n'est pas au programme de ce cours.

3 Convergences

3.1 Convergence en loi

Définition

Soient (X_n) une suite de v.a. et X une autre v.a. On note respectivement F_{X_n} et F_X les fonctions de répartition de X_n et de X. On dit que X_n **converge en loi** vers X, si, pour tout réel t où la fonction F_x est continue, on a

$$\lim_{n\to+\infty} F_{X_n}(t) = F_X(t).$$

On note $X_n \xrightarrow{\mathcal{L}} X$.

✎ Pour montrer qu'une suite de v.a. converge en loi, on fixe d'abord un t (parmi les points où F_X est continue) et on fait ensuite tendre n vers l'infini.

☞ Si la variable X est à densité, sa fonction de répartition F_X est continue et la convergence ci-dessous doit donc avoir lieu en tout point $t \in \mathbb{R}$.

☞ Certaines suites de v.a. ne convergent pas. On gardera également à l'esprit qu'une suite de v.a. discrète peut converger vers une v.a. à densité et inversement.

Exercice 4. Soit (X_n) une suite de v.a. telle que, pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{E}(1+\frac{1}{n})$. Montrer que (X_n) converge en loi vers une v.a. Z à déterminer.

Exercice 5. Soient $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{U}([\![0;n]\!])$ et $Y_n = \frac{1}{n}X_n$.

(1) Montrer que

$$F_{X_n}(t) = \begin{cases} 0, & \text{si } t < 0 \\ \frac{\lfloor t\rfloor+1}{n+1}, & \text{si } t \in [0;n] \\ 1, & \text{si } t > 1 \end{cases}$$

(2) En déduire l'expression de $F_{Y_n}(t)$.

(3) Conclure que $Y_n \xrightarrow{\mathcal{L}} Z$, où $Z \hookrightarrow \mathcal{U}([0;1])$.

Exercice 6. Soit $X_n \hookrightarrow \mathcal{U}([n;n+1])$. Montrer que (X_n) ne converge pas en loi.

La proposition suivante est une conséquence immédiate des définitions de fonction de répartition et de convergence en loi.

Propriété

Soit (X_n) une suite de v.a. telle que $X_n \xrightarrow{\mathcal{L}} X$. Pour tous réels a, b points de continuité de F_X avec $a < b$, on a

$$\lim_{n\to+\infty} P(a < X_n \leq b) = P(a < X \leq b).$$

Une conséquence de cette proposition est le théorème suivant, permettant la caractérisation de la convergence en loi pour les v.a. discrètes à valeurs dans $\mathbb{N}$.

Propriété

Soient X une v.a. discrète telle que $X(\Omega) \subset \mathbb{N}$ et (X_n) une suite de v.a. discrètes avec la même propriété. Alors,

$$\left(X_n \xrightarrow{\mathcal{L}} X\right) \Longleftrightarrow \left(\forall k \in \mathbb{N},\ \lim_{n\to+\infty} P(X_n = k) = P(X = k)\right).$$

Exercice 7. Soit (X_n) une suite de v.a. telle que, pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{B}\left(\ln\left(2+\frac{1}{n}\right)\right)$. Montrer que $X_n \xrightarrow{\mathcal{L}} X$, où $X \hookrightarrow \mathcal{B}(\ln(2))$.

3.2 Application: convergence de binomiales vers Poisson

Le résultat suivant permet de comprendre la terminologie *loi des évènements rares* appliquée à la loi de Poisson. En effet, un processus de comptage de succès, lorsque la probabilité du succès est très petite devant le nombre d'observations, peut être approché par une loi de Poisson.

Propriété

Approximation d'une binomiale par une loi de Poisson.

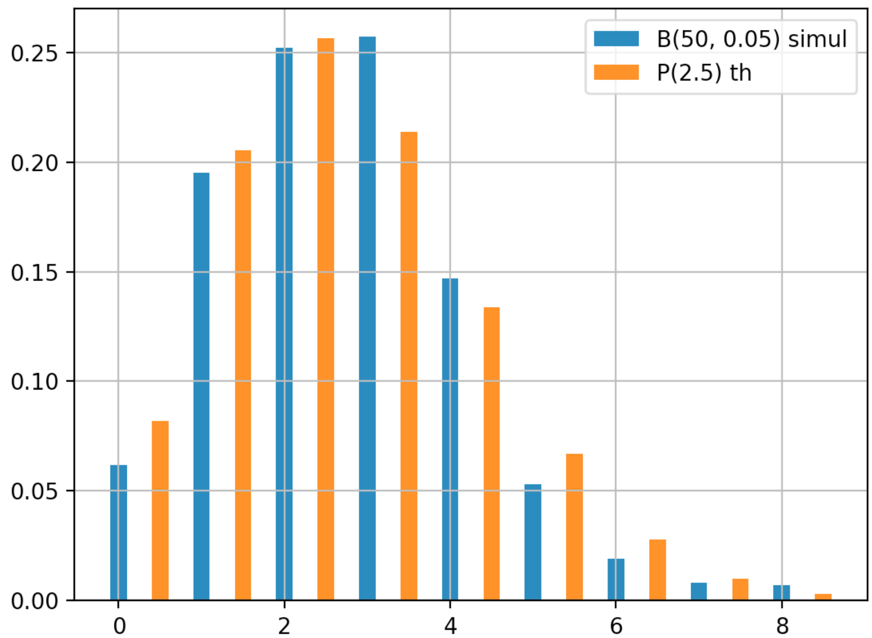
Soient λ un réel strictement positif et (X_n) une suite de v.a. telle que, pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{B}\left(n, \frac{\lambda}{n}\right)$. Alors, $X_n \xrightarrow{\mathcal{L}} X$, où $X \hookrightarrow \mathcal{P}(\lambda)$.

☞ En pratique, si p est petit et n grand, alors on approche une loi $\mathcal{B}(n,p)$ par la loi $\mathcal{P}(np)$. L'approximation sera toujours indiquée dans les sujets.

☞ L'exécution du programme Python suivant illustre la remarque précédente.

```python
import numpy as np
import numpy.random as rd
import matplotlib.pyplot as plt

X=[rd.binomial(50, 0.05) for k in range(1000)] #echantillon taille 1000
m=max(X)
freq=np.zeros(m+1)
for k in range(len(X)):
    freq[X[k]]+=1
freq=freq/1000

def distr_poisson_th(lam,n):
    y=np.zeros(n+1)
    y[0]=np.exp(-lam)
    for k in range(1, n+1):
        y[k]=y[k-1]*lam/k
    return y

N=[k for k in range(m+1)]
Z=distr_poisson_th(2.5, m)

plt.grid()
plt.bar(N, freq, width=0.2, label='B(50, 0.05) simul')
plt.bar([k+0.5 for k in range(m+1)], Z, width=0.2, label='P(2.5) th')
plt.legend()
plt.show()
```

Affichage Python



4 Théorème Central Limite

4.1 Le résultat

Le résultat suivant, qui peut être vu comme la finalité de ce cours de mathématiques appliquées, illustre le rôle crucial de la loi normale $\mathcal{N}(0,1)$ en statistique.

Théorème

Théorème Central Limite (TCL).

Soit (X_n) une suite de v.a. indépendantes de même loi, d'espérance (commune) finie m et de variance (commune) finie (et non nulle) σ^2. Pour $n \in \mathbb{N}^*$, on note

$$S_n = \sum_{k=1}^{n} X_k, \quad \overline{X}_n = \frac{S_n}{n}$$

et

$$S_n^* = \frac{S_n - E(S_n)}{\sigma(S_n)} = \frac{S_n - nm}{\sigma\sqrt{n}}, = \frac{\sqrt{n}(\overline{X}_n - m)}{\sigma} = \frac{\overline{X}_n - E(\overline{X}_n)}{\sigma(\overline{X}_n)} = \overline{X}_n^*.$$

Alors,

$$S_n^* \xrightarrow{\mathcal{L}} Z \qquad \text{et} \qquad \overline{X}_n^* \xrightarrow{\mathcal{L}} Z, \qquad \text{où} \qquad Z \hookrightarrow \mathcal{N}(0,1).$$

En particulier, notant Φ la fonction de répartition de la loi normale $\mathcal{N}(0,1)$, on a, pour tous réels a, b avec $a < b$,

$$\lim_{n \to +\infty} P\left(a < S_n^* \leq b\right) = \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} \mathrm{d}t.$$

☞ Pour reformuler "simplement" ce résultat, la somme (ou la moyenne) de n variables aléatoires indépendantes de même loi se comporte de manière approximativement gaussienne si n est suffisamment grand.

☞ On peut donc voir ce théorème comme une précision de la loi faible des grands nombres. Celle-ci stipule que la moyenne empirique converge vers la moyenne théorique. Le théorème central limite précise moralement que la convergence a lieu à vitesse de $1/\sqrt{n}$.

Exercice 8. Soit (U_n) une suite de v.a. indépendantes telle que, pour tout $n \in \mathbb{N}^*$, $U_n \hookrightarrow \mathcal{U}\left(\left[-\frac{1}{2}; \frac{1}{2}\right]\right)$. Montrer que

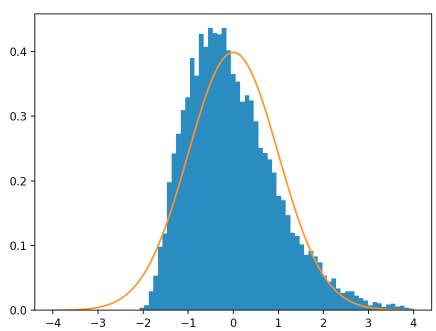
$$\sqrt{\frac{12}{n}} \sum_{k=1}^{n} U_k \xrightarrow{\mathcal{L}} U, \qquad \text{où} \quad U \hookrightarrow \mathcal{N}(0,1).$$

4.2 Une Illustration avec Python

On simule S_n^*, où $S_n = \sum_{i=1}^{n} X_i$ avec $X_i \hookrightarrow \mathcal{E}(1)$ (v.a.i.i.d). On observe $S_n^* \xrightarrow{\mathcal{L}} Z, \qquad Z \hookrightarrow \mathcal{N}(0,1).$

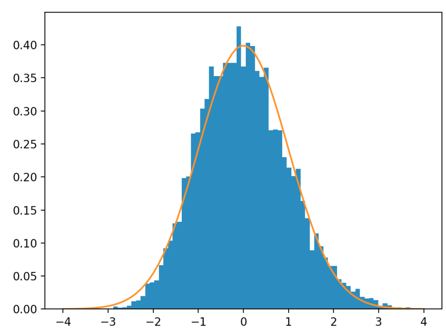
```
def S_star(n):
    S=[rd.exponential(1) for k in range(n)]
    return (np.sum(S)-n)/np.sqrt(n)

n=5 # puis n=25, n=50 et n=100
U=[S_star(n) for k in range(10000)]
X=np.linspace(-4, 4, 80)
Y=[np.exp(-(x**2/2))/np.sqrt(2*np.pi) for x in X]
plt.hist(U, X, density=True)
plt.plot(X,Y)
plt.show()
```

Affichage Python



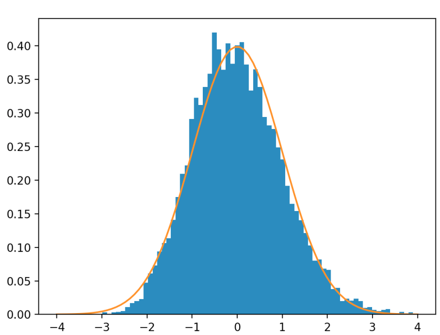
$n = 5$

Affichage Python



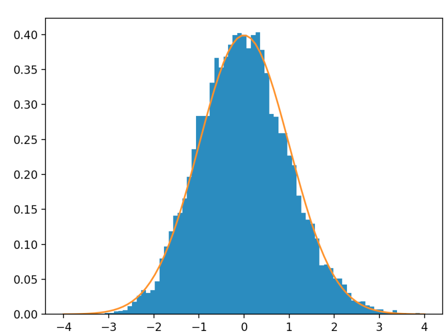
$n = 25$

Affichage Python



$n = 50$

Affichage Python



$n = 100$

5 Approximation

5.1 Approximation d'une binomiale par une loi normale

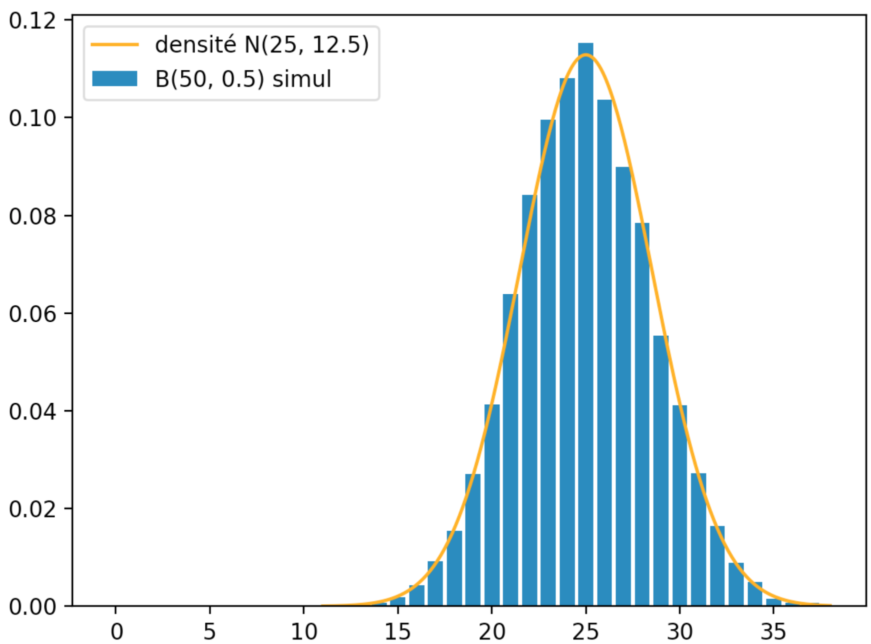
Propriété

Soit (S_n) une suite de v.a. indépendante de même loi binomiale $\mathcal{B}(n,p)$ (avec $p \in]0;1[$). Alors,

$$S_n^* = \frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathcal{L}} X, \qquad \text{où} \quad X \hookrightarrow \mathcal{N}(0,1).$$

☞ Une loi $\mathcal{B}(n,p)$ peut donc être approchée par une loi $\mathcal{N}(np, npq)$ lorsque n est assez grand. En pratique, cette approximation est d'autant plus valide que p est proche de 0.5.

Affichage Python



☞ On observera que l'on fait une approximation d'une loi discrète par une loi continue. Par conséquent, on approche en pratique $P(S_n = k)$ non pas par $P(X = k)$ (qui est nul comme tout le monde le sait bien), mais par $P(k - 1/2 \leq X \leq k + 1/2)$.

☞ Les instructions officielles précisent: *Toutes les indications devront êtres fournies aux candidats quant à la justification de l'utilisation des approximations*. Si le message est clair, on peut quand même mentionner que, souvent, on considère *raisonnable* d'approcher une loi binomiale $\mathcal{B}(n,p)$ par une loi normale $\mathcal{N}(np, npq)$ lorsque $n \geq 30$, $np \geq 5$ et $nq \geq 5$.

Exercice 9. On lance 900 fois une pièce de monnaie équilibrée et on note X le nombre de *Pile* obtenus. On veut estimer $\alpha = P(405 \leq X \leq 495)$.

(1) Quelle est la valeur exacte de α? (On exprimera le résultat sous forme d'une somme qu'on ne cherchera naturellement pas à calculer.)

(2) Utiliser la table de la loi normale centrée réduite pour donner une estimation de α.

5.2 Approximation d'une loi de Poisson par une loi normale

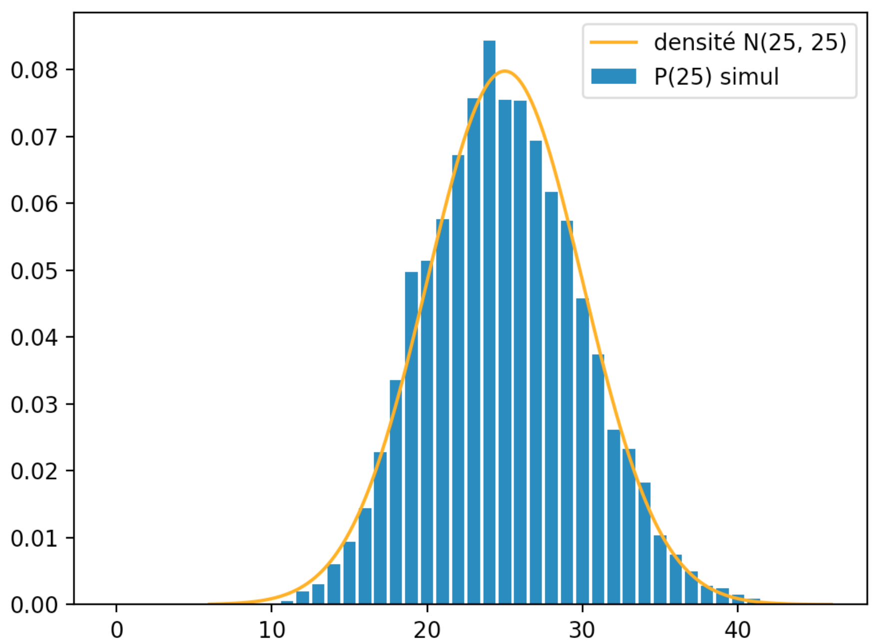
Propriété

Soit (S_n) une suite de v.a. indépendante de même loi de Poisson $\mathcal{P}(n\alpha)$ (avec $\alpha > 0$). Alors,

$$S_n^* = \frac{S_n - n\alpha}{\sqrt{n\alpha}} \xrightarrow{\mathcal{L}} X, \qquad \text{où} \quad X \hookrightarrow \mathcal{N}(0,1).$$

☞ Une loi $\mathcal{P}(\lambda)$ peut donc être approchée par une loi $\mathcal{N}(\lambda, \lambda)$ lorsque λ est assez grand.

Affichage Python



☞ Comme pour l'approximation précédente, on suivra les indications du texte mais on peut souvent considérer qu'il est raisonnable d'approcher une loi de Poisson $\mathcal{P}(\lambda)$ par une loi normale $\mathcal{N}(\lambda, \lambda)$ lorsque $\lambda \geq 18$.

6 Autres exercices

6.1 Manipulation des résultats et définitions

Exercice 1201. En utilisant l'inégalité de Bienaymé-Tchebychev, montrer que

$$\int_{-\infty}^{x} e^{-t^2/2} \mathrm{d}t \geq \sqrt{2\pi}\left(1 - \frac{1}{2x^2}\right).$$

Exercice 1202. Soient $n \in \mathbb{N}^*$ et $a > 0$. On considère une suite de v.a. (X_n) dont la fonction de répartition est donnée par

$$F_{X_n}(x) = \begin{cases} 0, & \text{si } x < 0 \\ 1 - \exp\left(-ax - \frac{x^2}{2n}\right), & \text{si } 0 \leq x < 2n \\ 1, & \text{si } x \geq 2n \end{cases}$$

Montrer que (X_n) converge en loi vers une v.a dont on précisera la loi.

Exercice 1203. Pour tout $n \in \mathbb{N}^*$, on considère la fonction f_n définie sur $\mathbb{R}$ par

$$f_n(x) = \begin{cases} (n+1)(1-x)^n, & \text{si } x \in [0;1] \\ 0, & \text{sinon} \end{cases}$$

(1) Montrer que f_n peut être considérée comme une densité de probabilité.
On note alors X_n une variable de densité f_n et on pose, pour tout $n \in \mathbb{N}^*$, $Y_n = nX_n$.
(2) Calculer la fonction de répartition F_n de la variable Y_n.
(3) Étudier la convergence en loi de la suite de variables aléatoires (Y_n).

Exercice 1204. Soit (X_n) une suite de v.a. (mutuellement) indépendantes avec $X_n \hookrightarrow \mathcal{B}(p_n)$. On note $\overline{X}_n = (X_1 + ... + X_n)/n$. On veut montrer que, pour tout $\varepsilon > 0$,

$$\lim_{n \to +\infty} P\left(\left|\overline{X}_n - \frac{1}{n}\sum_{i=1}^{n} p_i\right| \geq \varepsilon\right) = 0.$$

(1) Pourquoi ne peut-on pas appliquer la loi faible des grands nombres?
(2) Quelle est l'espérance de $\overline{X}_n$? Sa variance? Démontrer que $V(\overline{X}_n) \leq \frac{1}{n}$.
(3) En déduire le résultat.

Exercice 1205. Soit (X_n) une suite de v.a. indépendante de même loi $\mathcal{P}(1)$. On pose $S_n = X_1 + ... + X_n$.
(1) Quelle est la loi de S_n?
(2) Exprimer $P(S_n \leq n)$ en fonction de n, à l'aide d'une somme que l'on ne cherchera pas à calculer.
(3) En utilisant le théorème central limite, montrer que

$$\lim_{n \to +\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

Exercice 1206. Jean-Michel Jersey vend des maillots de foot dans sa boutique qui monte au Sacré-Coeur. Il a restreint des produits et ne propose que deux maillots (aux couleurs du PSG), celui de Messi (que l'on notera A) et celui de MBappé (noté B) et il dispose d'un stock de 40 exemplaires de A et 40 exemplaires de B. Chaque client demande soit A, soit B avec la probabilité $0,5$. Les demandes des clients sont indépendantes les unes des autres. Un samedi, 60 clients se présentent. On note x la probabilité de l'événement "Jean-Mi ne satisfait pas à toutes les demandes, ce jour là".
(1) Déterminer la loi de Y, nombre de clients demandant un maillot de Mbappé dans la matinée.
(2) Exprimer x à l'aide de Y .
(3) Déduire de l'inégalité de Bienaymé-Tchebychev un majorant de x.
(4) En approchant la loi de Y par une loi normale, exprimer x à l'aide de la fonction de répartition Φ puis utiliser la table de la loi $\mathcal{N}(0,1)$ pour donner une valeur approchée de x.

Exercice 1207. Soit (X_n) une suite v.a.i.i.d de loi commune $\mathcal{U}(]0;1])$. Pour tout $n \in \mathbb{N}^*$, on note

$$Y_n = e^{\sqrt{n}} \left(\prod_{k=1}^{n} X_i\right)^{1/\sqrt{n}}, \qquad L_n = \ln(Y_n).$$

(1) Déterminer la loi de $-\ln(X_1)$. En déduire $E(\ln(X_1))$ et $V(\ln(X_1))$.
(2) À l'aide du TCL, montrer que (L_n) converge en loi vers une variable dont on donnera la loi.
(3) (a) Que vaut $Y_n(\Omega)$?
(b) Soit $t < 0$. Que vaut $\lim_{n \to +\infty} P(Y_n \leq t)$?

(4) On définit la fonction F sur $\mathbb{R}$ par

$$F(x) = \begin{cases} 0, & \text{si } x \leq 0 \\ \Phi(\ln(x)), & \text{si } x > 0 \end{cases}$$

(a) Montrer que F est la fonction de répartition d'une variable aléatoire Z.
(b) Montrer que (Y_n) converge en loi vers Z.

Exercice 1208. Soit (X_n) une suite de v.a. telle que, pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{U}([-1/n; 1/n])$. Montrer que (X_n) converge en loi vers une v.a. X certaine égale à 0.

Exercice 1209. Soient $a > 0$ un réel (X_n) une suite de v.a.i.i.d de loi $\mathcal{U}([0; a])$. On pose

$$T_n = \min(X_1, ..., X_n), \quad \text{et} \quad M_n = \max(X_1, ..., X_n)$$

(1) Montrer que (M_n) converge en loi vers une variable certaine dont on précisera la valeur.
(2) Justifier un résultat analogue pour (T_n).
(3) On prend $b \in \mathbb{N}^*$ et on remplace les loi uniformes *continues* précédentes par des loi uniformes *discrètes* $\mathcal{U}([\![0; b]\!])$. Les variables T_n et M_n restent définies de la même manière. Étudier la convergence en loi des suites (T_n) et (M_n) dans ce cas.

Exercice 1210. Afin de progresser en maths, Ismaël s'entraîne tous les soirs sur un exercice de l'un des trois grands thèmes du programme (Algèbre Linéaire (1), Probabilités (2) ou Analyse (3)). Il choisit le thème au hasard et ne travaille jamais deux soirs de suite sur le même thème. Le premier soir, il commence avec l'Algèbre. On note X_n la variable aléatoire correspondant au thème travaillé le soir n. On note

$$U_n = \begin{pmatrix} P(X_n = 1) \\ P(X_n = 2) \\ P(X_n = 3) \end{pmatrix}$$

(1) Que vaut U_1 ?
(2) Déterminer une matrice M telle que, pour tout $n \in \mathbb{N}^*$, on ait $U_{n+1} = MU_n$.
(3) Diagonaliser M. En déduire M^n puis la loi de X_n. *On donne* $\text{Sp}(M) = \{1, -1/2\}$.
(4) Montrer que X_n converge en loi vers une v.a. X à préciser.

6.2 Exercices d'annales

Exercice 1211. (D'après **EDHEC 2013**) On considère n variables aléatoires $X_1, X_2, \ldots X_n$, $n \geqslant 2$, toutes définies sur $(\Omega; \mathcal{A}, P)$ indépendantes et suivant la loi uniforme sur $[0; 1]$. On pose $I_n = \min(X_1, X_2, ..., X_n)$. Déterminer la fonction de répartition de I_n et montrer que la suite $(I_n)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire dont on précisera la loi.

Exercice 1212. (D'après **HEC 2003**) Soient n un entier naturel non nul, N un entier supérieur ou égal à 2, et p un réel strictement compris entre 0 et 1.

Une compagnie aérienne a vendu n billets à cent euros pour le vol 714 (pour Sydney) qui peut accueillir jusqu'à N passagers. La probabilité pour qu'un acheteur se présente à l'embarquement est p et les comportements des acheteurs sont supposés indépendants les uns des autres.
Un acheteur qui ne se présente pas à l'embarquement est remboursé à 80%, tandis qu'un acheteur qui se présente à l'embarquement mais n'obtient pas de place, le vol étant déjà complet, est remboursé à 200%.
Soit X la variable aléatoire désignant le nombre d'acheteurs d'un billet se présentant à l'embarquement, soit Y la variable aléatoire désignant le nombre d'acheteurs d'un billet se présentant à l'embarquement mais n'obtenant pas de place et soit G la variable aléatoire désignant le montant en centaines d'euros du chiffre d'affaire de la compagnie sur le vol considéré.

(1) Quelle est la loi de X? Donner son espérance et sa variance.
(2) Préciser, pour tout élément ω de Ω, la valeur de $Y(\omega)$ en fonction de N et de $X(\omega)$, en distinguant les cas $X(\omega) > N$ et $X(\omega) \leqslant N$.
(3) Écrire l'expression de G en fonction de n, X, Y.
(4) On suppose, dans cette question seulement, que n est inférieur ou égal à N.
Calculer alors l'espérance $E(G)$ de la variable aléatoire G.

La compagnie cherche alors à évaluer la probabilité $P([X \geqslant N])$ et à savoir si le nombre n aurait pu être choisi de façon à optimiser son chiffre d'affaire.

(5) On suppose, dans cette question, que p est égal à $0,5$.

(a) Soit X^* la variable aléatoire définie par:

$$X^* = \frac{2X - n}{\sqrt{n}}.$$

Donner l'espérance et la variance de la variable aléatoire X^*.

(b) Par quelle loi approcher la loi de X^* si n est assez grand ? Montrer qu'alors une valeur approchée de la probabilité $P([X \geqslant N])$ est

$$\Phi\left(\frac{n+1-2N}{\sqrt{n}}\right),$$

où Φ désigne la fonction de répartition de la loi normale centrée réduite.

(c) Pour tout réel x supérieur ou égal à 1, on pose

$$f(x) = \frac{x+1-2N}{\sqrt{x}}.$$

Montrer que la fonction f est croissante.

(d) On suppose que N est égal à 320 et on donne

$$\Phi\left(\frac{7}{\sqrt{646}}\right) \approx 0,609\ ; \quad \Phi\left(\frac{6}{\sqrt{645}}\right) \approx 0,592.$$

Que peut-on en déduire pour $P([X \geqslant N])$ si n est inférieur ou égal à 645, puis si n est supérieur ou égal à 646?

(6) Pour tout entier naturel non nul m, on considère la fonction g_m définie sur $\mathbb{R}_+^*$ par

$$g_m(x) = e^{-x} \sum_{k=0}^{m} \frac{x^k}{k!}$$

(a) Montrer que la fonction dérivée de g_m est définie sur $\mathbb{R}_+^*$ par

$$g_m'(x) = -e^{-x} \frac{x^m}{m!}.$$

Montrer qu'elle vérifie la double inégalité

$$\forall x \in \mathbb{R}_+^*, \quad -e^{-m} \frac{m^m}{m!} \leqslant g_m'(x) \leqslant 0.$$

(b) En déduire que, si a et b sont deux réels vérifiant $0 < a < b$, on a

$$0 \leqslant g_m(a) - g_m(b) \leqslant (b-a) e^{-m} \frac{m^m}{m!}$$

(7) On suppose, dans cette question, que p est égal à $0,99$ et que n est strictement supérieur à N.

(a) Préciser la loi de la variable aléatoire $n - X$.

(b) On supposera, dans les prochains calculs, que la loi de la variable aléatoire $n - X$ peut être remplacée par la loi de Poisson de paramètre $0,01\,n$ dont on note F la fonction de répartition.

Que vaut alors $P([X \geqslant N])$?

(c) Exprimer le nombre $F(n - N)$ à l'aide d'une fonction g_m particulière de la Question **??**?

(d) On suppose que N est égal à 300.

Pour tout réel strictement positif α, on note F_α la fonction de répartition de la loi de Poisson de paramètre α et on donne:

$$F_3(2) \approx 0,423\ ; \quad F_3(3) \approx 0,647\ ; \quad e^{-3}\frac{3^3}{3!} \approx 0,224$$

Montrer que, si n est égal à 302, $P([X \geqslant N])$ est au plus égal à $0,5$ et que, si n est égal à 303, $P([X \geqslant N])$ est strictement supérieur à $0,6$.

Exercice 1213. (D'après **ESSEC II 2013**) Pour tout entier n strictement positif, on se donne un réel p_n strictement positif et n variables aléatoires (X_n) indépendantes et suivant une loi de Bernoulli de paramètre p_n. On suppose que np_n a une limite finie strictement positive et on pose

$$\lim_{n\to+\infty} np_n = \lambda.$$

(1) Quelle est la loi de $S_n = \sum_{k=1}^{n} X_k$?

(2) Soit k un entier naturel.
  (a) Donner l'expression de $P(S_n = k)$ pour n supérieur ou égal à k.
  (b) Que peut-on dire de la limite de p_n quand n tend vers l'infini? Étudier la limite de $(1-p_n)^n$ quand n tend vers l'infini.
  (c) Montrer que

$$\lim_{n\to+\infty} P(S_n = k) = e^{\lambda}\frac{\lambda^k}{k!}.$$

(3) On pose $N_n = \max(X_1, \ldots, X_n)$.
  (a) Quelles sont les valeurs que peut prendre la variable aléatoire N_n?
  (b) Calculer

$$\lim_{n\to+\infty} P(N_n = 0).$$

  (c) En déduire la limite en loi de la suite $(N_n)_{n\in\mathbb{N}^*}$.

Exercice 1214. (D'après **EML 2014**)

Pour tout entier n supérieur ou égal à 2, on considère une urne contenant n boules numérotées de 1 à n, dans laquelle on effectue une succession de $(n+1)$ tirages d'une boule avec remise et l'on note X_n la variable aléatoire égale au numéro du tirage où, pour la première fois, on a obtenu un numéro supérieur ou égal au numéro précédent.

Ainsi, pour tout entier n supérieur ou égal à 2, la variables X_n prend ses valeurs dans $[\![2; n+1]\!]$. Par exemple, si $n = 5$ et si les tirages amènent successivement les numéros 5,3,2,2,6,3, alors $X_5 = 4$. Pour tout k de $[\![1; n+1]\!]$, on note N_k la variable aléatoire égale au numéro obtenu au k-ième tirage.

Partie I : Étude du cas $n = 3$

(1) (a) Exprimer l'évènement $(X_3 = 4)$ à l'aide des variables N_1, N_2, N_3. En déduire $P(X_3 = 4)$.
  (b) Montrer que $P(X_3 = 2) = 2/3$, et en déduire $P(X_3 = 3)$.

(2) Calculer l'espérance de X_3.

Partie II : Cas général $n \geq 2$

(3) Pour tout k de $[\![1; n+1]\!]$, reconnaître la loi de N_k et rappeler son espérance et sa variance.
(4) Calculer $P(X_n = n+1)$.
(5) Montrer, pour tout i de $[\![1; n]\!]$:

$$P_{(N_1=i)}(X_n = 2) = \frac{n-i+1}{n}.$$

(6) En déduire une expression simple de $P(X_n = 2)$.
(7) Soit $k \in [\![2; n]\!]$. Justifier l'égalité d'événements suivante: $(X_n > k) = (N_1 > N_2 > \ldots > N_k)$. En déduire que

$$P(X_n > k) = \frac{1}{n^k}\binom{n}{k}.$$

Vérifier que cette dernière égalité reste valable pour $k = 0$ et pour $k = 1$.
(8) Exprimer, pour tout $k \in [\![2; n+1]\!]$, $P(X_n = k)$ à l'aide de $P(X_n > k-1)$ et de $P(X_n > k)$.

(9) En déduire que

$$\mathrm{E}(X_n) = \sum_{k=0}^{n} P(X_n > k).$$

Calculer ensuite $\mathrm{E}(X_n)$.

(10) Montrer que, pour tout $k \in [\![2; n+1]\!]$, $P(X_n = k) = \dfrac{k-1}{n^k}\dbinom{n+1}{k}$.

Partie III : Une convergence en loi

(11) Soit k un entier fixé supérieur ou égal à 2. Montrer

$$\lim_{n\to+\infty} P(X_n = k) = \frac{k-1}{k!}.$$

(12) Montrer que la série $\sum_{k\geqslant 2} \frac{k-1}{k!}$ converge et calculer sa somme.

On admet qu'il existe une variable aléatoire Z à valeurs dans $\mathbb{N} \setminus \{0; 1\}$ telle que

$$\forall k \geq 2, \qquad P(Z = k) = \frac{k-1}{k!}.$$

(13) Montrer que Z admet une espérance et la calculer. Comparer $\mathrm{E}(Z)$ et

$$\lim_{n\to+\infty} \mathrm{E}(X_n).$$

7 Annexe : allègement des hypothèses de la loi faible des grands nombres

On rappelle que si X est une variable aléatoire **admettant une variance**, alors la loi faible des grands nombres permet d'affirmer que, notant $(X_1, ..., X_n)$ un $n-$échantillon de X, on a

$$\forall t > 0, \qquad P\left(\left|\frac{X_1 + ... + X_n}{n} - E(X)\right| > t\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

La preuve de ce résultat repose sur l'inégalité de Bienaymé-Tchebychev (on invite d'ailleurs à la reprendre avant d'attaquer ce problème).
L'objectif est de montrer que ce résultat reste vrai lorsque X admet une espérance mais pas de variance.

Pour simplifier les calculs, on se place dans la situation d'une variable X centrée.

On considère donc une variable aléatoire X de densité[1] f, d'espérance nulle dont on a un $n-$échantillon $(X_1, X_2, ..., X_n)$.

Partie 1 - Des variables tronquées

Soit $A > 0$ fixé. Pour tout $k \in [\![1; n]\!]$, on introduit les variables

$$Y_k = X_k \cdot \mathbb{1}_{(|X_k| \leq A)}, \qquad \text{et} \qquad Z_k = X_k \cdot \mathbb{1}_{(|X_k| > A)}.$$

(1) Quelle relation a-t-on entre X_k, Y_k et Z_k?
(2) Montrer que Y_k admet un moment d'ordre 2 et que $E(Y_k^2) \leq A^2$.
(3) Montrer que

$$\lim_{A\to+\infty} E(Y_k) = E(X_k) = 0 \quad \text{et que} \quad \lim_{A\to+\infty} E(Z_k) = 0.$$

[1]Tout ceci fonctionne de manière analogue dans le cas d'une variable discrète

Partie 2 - Inégalités

Dans toute la suite, on considère $\varepsilon > 0$ fixé.

(4) Soient $x, y \in \mathbb{R}$. Montrer que

$$|x+y| > t \Longrightarrow \left(\left[|x| > \frac{t}{2}\right] \text{ ou } \left[|y| > \frac{t}{2}\right]\right)$$

(5) On note alors

$$\overline{X}_n = \frac{X_1 + ... + X_n}{n}, \quad \overline{Y}_n = \frac{Y_1 + ... + Y_n}{n} \quad \text{et} \quad \overline{Z}_n = \frac{Z_1 + ... + Z_n}{n}.$$

Déduire de la question précédente que

$$P\left(|\overline{X}_n| > t\right) \leq P\left(|\overline{Y}_n| > \frac{t}{2}\right) + P\left(|\overline{Z}_n| > \frac{t}{2}\right).$$

(6) (a) En réutilisant la Question (4), montrer que

$$P\left(|\overline{Z}_n| > \frac{t}{2}\right) \leq nP\left(|Z_1| > \frac{nt}{2}\right).$$

(b) En déduire, à l'aide de l'inégalité de Markov, qu'il existe $A_1 > 0$, tel que, si $A \geq A_1$,

$$P\left(|\overline{Z}_n| > \frac{t}{2}\right) \leq \frac{\varepsilon}{2}.$$

(7) (a) Montrer que

$$E(\overline{Y}_n^2) = \frac{1}{n^2}\left(\sum_{k=1}^{n} E(Y_k^2) + 2\sum_{1\leq i<j\leq n} E(Y_iY_j)\right).$$

(b) Montrer que, pour $i \neq j$,

$$E(Y_iY_j) = E(Y_i)E(Y_j) = E(Y_1)^2$$

En déduire que

$$2\sum_{1\leq i<j\leq n} E(Y_iY_j) \leq n(n-1)E(Y_1)^2.$$

(c) Obtenir ensuite que

$$E(\overline{Y}_n^2) \leq \frac{A^2}{n} + E(Y_1)^2.$$

(d) Justifier de l'existence d'un $A_2 > 0$ tel que pour $A \geq A_2$,

$$E(Y_1)^2 \leq \frac{16\varepsilon}{t^2}.$$

(e) Utiliser une inégalité de Markov pour obtenir que

$$P\left(|\overline{Y}_n| > \frac{t}{2}\right) \leq \frac{t^2A^2}{4n} + \frac{\varepsilon}{4}.$$

(8) Montrer que si $A \geq \max(A_1, A_2)$,

$$P\left(|\overline{X}_n| > t\right) \leq \frac{t^2A^2}{4n} + \frac{3\varepsilon}{2}.$$

(9) Conclure.

8 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