Chapitre 2

Chaînes de Markov à temps discret

2-1 Processus Stochastiques

Un processus aléatoire (ou processus stochastique) $(X_t)_{t\in T}$ est une famille de variables aléatoires (v.a) indexées par t et définies sur un même espace de probabilité (Ω, \mathcal{F}, P) . Les mots processus et stochastique signifient respectivement fonction et aléatoire.

Définition 1 Un processus stochastique $(X_t)_{t \in T}$ est une application X de Ω dans E:

$$X: \qquad \Omega \times T \to E$$

$$(w,t) \mapsto X(w,t) = X_t(w)$$

Un processus prend ses valeurs dans un espace des états noté E et évolue dans un espace des temps T. Un processus stochastique à temps discret est une famille $(X_n)_{n\in\mathbb{N}}$ de v.a .Dans ce cas, on note $T=\mathbb{N}$. Un processus stochastique à temps continu est une famille $(X_t)_{t\in\mathbb{R}}$ +de v.a .Dans ce cas, on note $T=\mathbb{R}^+$.

L'espace d'états E peut être discret (\mathbb{N} ou \mathbb{Z}) ou continu (\mathbb{R} ou \mathbb{R}^+)

Remarque 1

- Pour $t \in T$ fixé, $w \in \Omega \mapsto X_t(w)$ est une variable aléatoire réelle.
- pour $w \in \Omega$ fixé, $t \in T \mapsto X_t(w)$ est une fonction à valeurs réelles, appelée une trajectoire

Un processus stochastique est caractérisé par la nature de l'espace des états, de l'espace des temps et les relations de dépendance entre les variables aléatoires .

2-1-1 Processus à accroissements indépendants

Un processus $(X_t)_{t \in T}$ est dit à accroissements indépendants (noté P.A.I) si pour tout $n \geq 1$ et toute suite d'instants t_1, \ldots, t_n tels que $0 < t_1 < \cdots < t_n$, les variables aléatoires $X_{t_1}, (X_{t_2} - X_{t_1}), \ldots, (X_{t_n} - X_{t_{n-1}})$ sont indépendantes.

2-1-2 Processus à accroissements stationnaires

Un processus est à accroissements stationnaires (noté P.A.S) si pour tout n=1,2,... et toute suite d'instants $t_1,...,t_n$ et h>0, la distribution du vecteur $\left(X_{t_2+h}-X_{t_1+h},...,X_{t_n+h}-X_{t_{n-1}+h}\right)$ est la même. Cela signifie que la distribution du vecteur $\left(X_{t_2+h}-X_{t_1+h},...,X_{t_n+h}-X_{t_{n-1}+h}\right)$ dépend uniquement de la longueur de l'intervalle $]h,t_i[$ et non pas de h et t_i .

Définition 2 Un processus qui est à la fois un (P.A.I) et un (P.A.S) est appelé processus à accroissements indépendants et stationnaires (noté P.A.I.S)

2-1-3 Processus de Markov

Un processus $(X_t)_{t \in T}$ est un processus de Markov, s'il vérifie la propriété sans mémoire suivante :pour toute suite d'instants $t_1, t_2, \ldots, t_n, t_{n+1}$ tels que $t_1 < t_2 < \ldots < t_n < t_{n+1} \in T$, pour tout $i_1, \ldots, i_n, i_{n+1} \in E$ on a

$$P(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n, \dots, X_{t_1} = i_1) = P(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n)$$

2-2- Chaînes de Markov à temps discret

Définition3 On appelle chaîne de Markov à temps discret tout processus à temps discret $(X_n)_{n\in\mathbb{N}}$ à valeurs dans un espace d'étatsE discret tel que pour tout entier $n\geq 0$ et toute suite (i,j,i_{n-1},\ldots,i_0) dans E on a :

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0) = P(X_{n+1} = j | X_n = i) = p_{ij}(n)$$
 (1)

Avec

$$P(X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0) > 0$$

Cette propriété est connue sous le nom de **propriété de Markov**. Elle signifie que l'état futur de la chaine à l'instant n+1 ne dépend que de son état présent à l'instant n mais ne dépend pas de ses états antérieurs .

Lorsque les probabilités de transition (1) sont stationnaires (c'est-à-dire les mêmes pour tout entier $n \ge 0$), la chaine est dite **homogène**. i.e pour tout $n \ge 0$, pour $i, j \in E$,

$$P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i) = p_{ij}$$

Dans la suite, on étudiera que les chaînes de Markov homogènes à valeurs dans un espace d'états E fini.

2-2-1 Matrice de transition – Graphe de transition

La matrice $P=\left(p_{ij}\right)_{i,j\in E}$ dont les éléments sont les probabilités p_{ij} , $i,j\in E=\{0,1..k\}$ est appelée matrice de transition

$$P = \begin{pmatrix} p_{00} & p_{01} & \dots & p_{0k} \\ p_{10} & p_{11} & \dots & p_{1k} \\ p_{20} & p_{21} & \dots & p_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ p_{k0} & p_{k1} & \dots & p_{kk} \end{pmatrix}$$

Les probabilités p_{ij} satisfont :

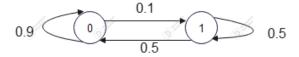
$$p_{ij} \geq 0$$
 , pour tout $i,j \in E$ et $\sum_{j \in E} p_{ij} = 1$

Une matrice vérifiant ces deux conditions est appelée matrice stochastique.

Le **graphe de transition** d'une chaîne de Markov estformé des sommets qui sont les états de E et les arcs correspondant aux transitions possibles pour lesquelles les probabilités p_{ij} sont positives.

Exemple 1 : Représenter le graphe de la matrice de transition P d'une chaine de Markov homogène suivante . On prendra comme espace des états $\{0,1\}$

$$P = \begin{pmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{pmatrix}$$



Définition 4 Pour une chaîne de Markov $(X_n)_{n\geq 0}$, la probabilité conditionnelle

$$p_{ii}^{(n)} = P(X_{n+1} = j | X_1 = i) = P(X_{n+k} = j | X_k = i)$$

est la probabilité de transition de l'état i vers l'état j en n étapes .

On note $P^{(n)}$ la matrice de transition en n étapes dont les éléments sont les probabilités $p_{ij}^{(n)}$.

Alors pour tout $n \ge 1$, on a $P^{(n)} = P^n$

Remarque 2: on a $P^{(1)} = P$ et $P^{(0)} = I$ ou I est la matrice identité.

Théorème 1 On considère une chaîne de Markov sur un espace d'états E, de matrice de transition P. Alors pour tout entiers positifs m et s on a

$$p_{ij}^{(m+s)} = \sum_{k \in E} p_{ik}^{(m)} p_{kj}^{(s)}$$

Cette équation est appelée équation de Chapman Kolmogorov.

Pour tout $n \ge 1$, cette dernière équation peut s'exprimer sous la forme suivante :

$$p_{ij}^{(n)} = \sum_{k \in E} p_{ik}^{(n-1)} p_{kj}$$

Exemple 2 : Considérons une chaîne de Markov à deux états $\{0,1\}$ de matrice de transition :

$$P = \begin{pmatrix} 4/10 & 6/10 \\ 3/10 & 7/10 \end{pmatrix}$$

Calculer la probabilité que la chaîne étant en 1 visite l'état 0 en deux transitions (ou 2 étapes)

$$p_{10}^2 = \sum_{k=0}^{1} p_{1k} p_{k0} = p_{10} p_{00} + p_{11} p_{10} = 3/10 \times 4/10 + 7/10 \times 3/10 = 33/100$$

2-2-2 Classification des états

Un état j est dit accessible depuis l'étati $(i \rightarrow j)$, s'il existe un entier n > 0 tel que $p_{ij}^{(n)} > 0$ Si l'état i est accessible depuisj et j est accessible depuisi on dit que ces états communiquent et l'on note $i \leftrightarrow j$. Cette relation est une relation d'équivalence :

- Réflexive : $i \leftrightarrow i$, puisque par convention $\,p_{ii}^{(0)} = 1 > 0\,$

- Symétrique : $i \leftrightarrow j \Leftrightarrow j \leftrightarrow i$

- Transitive : $i \leftrightarrow j$ et $j \leftrightarrow k$ alors $i \leftrightarrow k$

Le concept de communication permet de réaliser une partition de E en classes disjointes $E = C_1 \cup ... \cup C_r$ tel que tous les états d'une classe communiquent entre eux et deux états appartenant à deux classes différentes ne communiquent jamais.

1- Classe transitoire et classe récurrente

Une classe est **transitoire** (ou de transition) s'il est possible de sortir de cette classe mais dans ce cas, la chaine ne pourra plus jamais y retourner .

Une classe est récurrente (ou de récurrence) s'il est impossible de la quitter

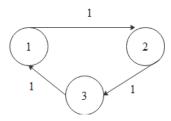
Une chaîne de Markov est **irréductible** si elle ne contient qu'une seule classe d'équivalence. Autrement dit tous les états de la chaine communiquent entre eux.

Une chaîne de Markov est **indécomposable** si elle est formée d'états de transition et une seule classe de récurrence.

Exemples 3 Considérons la chaîne de Markov dont l'espace des états est $E = \{1,2,3\}$, la matrice de transition

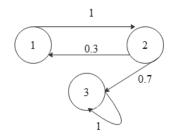
$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

son graphe de transition est



On constate qu'il y a une seule classe d'équivalence donc la chaîne est irréductible.

2- Soit la chaîne de Markov définie sur $E = \{1,2,3\}$ par le graphe de transition suivant :



Cette chaîne n'est pas irréductible car elle comporte 2 classes : $C_1 = \{1,2\}$ est une classe transitoire et $C_2 = \{3\}$ est une classe récurrente. La chaîne est indécomposable.

2- Période

La période d'un état i, notée d(i) estégale au plus grand diviseur commun de tous les n tel que $p_{ii}^n > 0$. On écrit :

$$d(i) = PGCD\{n \ge 1, p_{ii}^n > 0\}$$

- Si d(i) > 1alors l'état i est périodique
- Si d(i) = 1 alors l'état i est aperiodique

Propriétés:

i- si $p_{ii} > 0$ alors i est apériodique

ii- Si $i \leftrightarrow j$ alors ils ont la même période .

iii- Les états appartenant à la même classe ont la même période

Exemple 4 Considérons la chaîne de Markov dont la matrice de transition est donnée par

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

Trouver la période de la chaine

Solution : On a d'après le graphe $\;p_{11}>0\;$, $\;p_{11}^{(2)}>0\;$, $p_{11}^{(3)}>0$,.. donc

$$d(1) = PGCD\left\{n \ge 1; \ p_{11}^{(n)} > 0\right\} = PGCD\{1,2,3,\dots\} = 1$$

Puisque la chaîne est irréductible alors d(1) = d(2) = d(3) = 1

2-2-3 Loi de probabilité de $\boldsymbol{X_n}$

Pour tout $n \in \mathbb{N}$, on note $\pi(n)$ la loi de X_n , c'est-à-dire le vecteur ligne $\pi(n) = (\pi_1(n), \pi_2(n), \ldots)$ tel que

$$\pi_i(n) = P(X_n = i)$$
, $i \in E$ avec $\sum_{i \in E} \pi_i(n) = 1$

D'après le théorème des probabilités totales, on a

$$\pi_i(n) = \sum_{j \in E} \pi_j(0) p_{ji}^{(n)}$$
 (2)

Pour calculer la loi (ou la distribution) de X_n , il faut connaître la distribution initiale $\pi(0)=$ $(\pi_1(0),\pi_2(0),.)$

Où
$$\pi_i(0) = P(X_0 = i)$$
 pour $i \in E$

En notation matricielle la relation (2) s'écrit :

$$\pi(n) = \pi(0)P^n$$
, $n = 0.1,...$

 $\pi(n+1) = \pi(n)P$ De façon analogue, on a

La collection des $\pi(n)$, $n \ge 0$ représente le comportement transitoire.

Exemple 5 Considérons une chaîne de Markov à valeurs dans $E = \{0,1\}$ de matrice de transition :

$$P = \begin{pmatrix} 2/3 & 1/3 \\ 1/2 & 1/2 \end{pmatrix}$$

Supposons que $\pi(0) = (\pi_0(0), \pi_1(0)) = (0, 1)$. Calculer $\pi(1)$ et $\pi(2)$

Solution

Distribution de X_1 : $\pi(1) = \pi(0)P = \left(\frac{1}{2}, \frac{1}{2}\right)$

Distribution de X_2 :

$$\pi(2) = \pi(0)P^2 = \left(\frac{7}{12}, \frac{5}{12}\right)$$
 ou bien $\pi(2) = \pi(1)P = \left(\frac{7}{12}, \frac{5}{12}\right)$

Proposition1 Les probabilités de transition et la distribution de X_0 déterminent complètement la chaine .En effet, pour tout $i_0, ..., i_{n+1} \in E$, on a

$$P(X_{n+1} = i_{n+1}, X_n = i_n, ..., X_0 = i_0) = P(X_0 = i_0) p_{i_0 i_1} p_{i_1 i_2} ... p_{i_n i_{n+1}}$$

2-2-4 Distribution stationnaire

Une distribution de probabilité $\pi = (\pi_0, \pi_1, ...)$ est une distribution stationnaire (ou probabilité invariante) par rapport à une matrice P si

$$\pi P = \pi$$

Avec $\sum_{i \in E} \pi_i$, $\pi_i \geq 0$

Propriété: Toute chaîne de Markov finie possède toujours au moins une distribution stationnaire.

Théorème 2 Une chaîne de Markov finie possède une unique distribution stationnaire si et seulement si elle comprend une seule classe récurrente.

Exemple 6 Considérons la chaîne de Markov définie sur $E = \{0,1,2\}$ de matrice de transition suivante :

$$P = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

Calculer la distribution stationnaire.

Solution

En traçant le graphe de transition, on remarque que la chaîne est irréductible donc elle admet une unique distribution stationnaire π tels que

$$\begin{cases} \frac{1}{2}\pi_0 + \frac{1}{4}\pi_1 + \frac{1}{3}\pi_2 = \pi_0 \\ \frac{1}{2}\pi_1 + \frac{1}{3}\pi_2 = \pi_1 \\ \frac{1}{2}\pi_0 + \frac{1}{4}\pi_1 + \frac{1}{3}\pi_2 = \pi_2 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases}$$

De l'équation (2) on a $\pi_2=\frac{3}{2}\pi_1$, on remplace π_2 par $\frac{3}{2}\pi_1$ dans l'équation (1) on trouve $\pi_0=\frac{3}{2}\pi_1$.

En tenant compte de $\pi_0 + \pi_1 + \pi_2 = 1$, on obtient

$$\pi_0 = \frac{3}{8}$$
, $\pi_1 = \frac{2}{8}$ et $\pi_2 = \frac{3}{8}$, $\pi = \left(\frac{3}{8}, \frac{2}{8}, \frac{3}{8}\right)$

Exemple 7 Soit la matrice de transition d'une chaîne de Markov donnée par

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Cette chaîne comporte une classe transitoire $\mathcal{C}_1=\{1\}$ et deux classes récurrentes $\mathcal{C}_1=\{2\}$ et $\mathcal{C}_2=\{3,4\}$ donc il n'existe pas une unique distribution stationnaire. Cette chaîne admet une infinité de distributions stationnaires définies par $(0,1-2\alpha,\alpha,\alpha)$, $0\leq\alpha\leq1/2$.

Remarque 3: Une chaîne de Markov infini n'admet pas toujours de distribution stationnaire.

2-2-5 Distribution limite

On dit qu'une chaîne de Markov converge vers π^* ou possède une distribution limite π^* si

$$\lim_{n\to+\infty}\pi(n)=\pi^*$$

Si la distribution $\pi(n)$ converge, lorsque n tend vers l'infini, vers une distribution limite, cette dernière définit **le régime permanent** du processus stochastique.

Théorème 3 Si au moins une puissance de la matrice de transition P n'a que des termes strictement positifs , alors la chaîne de Markov $(X_n)_n$ possède une distribution limite π^* qui est un vecteur de probabilité strictement positif. On écrit $\lim_{n \to +\infty} P^n = P^*$.

où P^* est une matrice dont toutes les lignes sont identiques au vecteur π^*

Théorème 4 Si la valeur propre 1 de P est simple et les autres valeurs propres sont de module strictement inférieur à 1, alors $\lim_{n\to+\infty}P^n=P^*$

Exemple 8 Considérons une chaîne de Markov de matrice $P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix}$. On a

$$P^{2} = \begin{pmatrix} + & + & + \\ 0 & + & + \\ + & + & + \end{pmatrix} P^{3} = \begin{pmatrix} + & + & + \\ + & + & + \\ + & + & + \end{pmatrix}$$

Le symbole + indique une valeur strictement positive. D'après le théorème 3, la chaîne converge vers π^* D'autre part , les valeurs propres de P sont $\lambda_1=-1/2$, $\lambda_2=0$ et $\lambda_3=1$. Ce qui confirme les résultats du théorème 4 . Autrement dit, la chaîne converge vers π^* .

Remarque 4 Les hypothèses des deux théorèmes précédents ne sont pas équivalentes.

Exemple 9 Soit la matrice stochastique $P = \begin{pmatrix} 1/2 & 1/2 \\ 0 & 1 \end{pmatrix}$

Le théorème 3 ne s'applique pas . mais les valeurs propres sont $\lambda_1=1/2$ et $\lambda_1=1$, ce qui permet d'utiliser le théorème 4 et de montrer que la chaîne converge vers une distribution limite.

1- Calcul de la distribution limite

Pour calculer P^n , on diagonalise la matrice P en l'écrivant sous la forme $P = SDS^{-1}$ donc $P^n = SD^nS^{-1}$ Où,S est la matrice des vecteurs propres, D est la matrice diagonale des valeurs propres et S^{-1} est la matrice inverse de S.

Exemple 10 Reprenons l'exemple 8 et calculons la matrice limite P^* ainisi que la distribution limite π^*

les vecteurs propres correspondant aux valeurs propres $\lambda_1 = -1/2$, $\lambda_2 = 0$, $\lambda_3 = 1$ sont respectivement

$$V_1 = (-1,2,-1)$$
 , $V_2 = (0,1,-1)$ et $V_3 = (1,1,1)$

ďoù,

$$D = \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad S = \begin{pmatrix} -1 & 0 & 1 \\ 2 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix} \quad \text{et} \quad S^{-1} = \begin{pmatrix} -2/3 & 1/3 & 1/3 \\ 1 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

On écrit
$$P^n = \begin{pmatrix} -1 & 0 & 1 \\ 2 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} (-1/2)^n & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2/3 & 1/3 & 1/3 \\ 1 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

$$\lim_{n \to +\infty} P^n = S \lim_{n \to +\infty} D^n S^{-1} = \begin{pmatrix} -1 & 0 & 1 \\ 2 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2/3 & 1/3 & 1/3 \\ 1 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

et la distribution limite $\pi^* = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$

2- Distribution limite et distribution stationnaire

Si π^* est la distribution limite d'une chaîne de Markov, alors π^* est l'unique distribution stationnaire de cette chaîne. c'est à dire $\pi^* = \pi$. La réciproque est fausse.

2-2-6 Chaines de Markov absorbantes

Définition 5 : Un état *i* est absorbant si la chaine ne peut plus le quitter une fois qu'elle y est entré. et $p_{ii}^{(n)} = 1, \ \forall n \ge 1.$

Une chaîne de Markov est absorbante si elle comprend au moins un état absorbant et si l'on peut passer de n'importe quel état à un état absorbant.

1- Délais et probabilités d'absorption

Soit N_i : nombre de transitions jusqu'à l'absorption en partant de i

 $t_i = E(N_i)$: temps moyen jusqu' à l'absorption en partant de i

 b_{ij} : Probabilité que le processus soit absorbé dans j si son état initial est i

On a $t_i = 0$ si *i* est absorbant

 $b_{ii} = 1$ si i est absorbant

 $b_{ij} = 0$ si i est absorbant et $i \neq j$

Théorème 5 Les quantités t_i sont solutions du système :

$$t_i = 1 + \sum_{k \in F'} p_{ik} t_k$$
, $i \in E'$

Où , E' est l'ensemble de tous les états non absorbants .

Théorème 6 Soit j un état absorbant et E' l'ensemble des états non absorbants .Les probabilités $b_{i,i}$, $i \in$ E' sont solutions du système d'équations suivant :

$$b_{ij} = p_{ij} + \sum_{k \in E'} p_{ik} \, b_{kj}$$

Exemple 11 Considérons une chaîne de Markov à espace d'états $E = \{1,2,3\}$ et de matrice de transition

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \\ 0 & 0 & 1 \end{pmatrix}$$

- 1- Déterminer la nature des classes
- 2- Calculer la probabilité est $P(N_1 = k)$, k = 1,2,3
- 3- Les temps moyens jusqu'à l'absorption

Solution

La classe $C_1=\{1,2\}$ est transitoire ,la classe $C_2=\{3\}$ est récurrente et absorbante . l'état 3 est absorbant La probabilité du nombre de transitions jusqu'à l'état absorbant 3 en partant de 1 est $P(N_1=k), k=1,2,..$

Pour , k = 1,2,3 , on a

$$P(N_1 = 1) = p_{13} = 1/2$$

$$P(N_1 = 2) = p_{12}p_{23} = 1/3$$

$$P(N_1 = 3) = p_{12}p_{21}p_{13} = 1/12$$

Les temps moyens jusqu'à l'absorption $\ t_i = 1 + \sum_{k \in E'} p_{ik} t_k$, $\ i = 1,2$.

$$t_1 = 1 + p_{11}t_1 + p_{12}t_2 = 1 + \frac{1}{2}t_2$$

$$t_2 = 1 + p_{21}t_1 + p_{22}t_2 = 1 + \frac{1}{3}t_1$$

En remplaçant t_1 par $1+\frac{1}{2}t_2$ dans la seconde équation , on obtient $t_2=1+\frac{1}{3}\left(1+\frac{1}{2}t_2\right)$ donc $\frac{5}{6}t_2=\frac{4}{3}$

D'où
$$t_2 = \frac{8}{5}$$
 et $t_1 = \frac{9}{5}$.