

الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

2/96

وزارة التربية الوطنية

MINISTERE DE L'EDUCATION NATIONALE

ECOLE NATIONALE POLYTECHNIQUE

المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

DEPARTEMENT GENIE INDUSTRIEL

PROJET DE FIN D'ETUDES

SUJET

APPROCHE STATISTIQUE
DU PROBLEME DU
VOYAGEUR DE COMMERCE

1 DISQUETTE

Proposé par :

M. A. OUABDESSELAM
M. T. LAMRAOUI

Etudié par :

A. AOUFI
A. SOUAYEB

Dirigé par :

M. A. OUABDESSELAM
M. T. LAMRAOUI

PROMOTION

Juin 1996

République Algérienne Démocratique et Populaire
Ecole Nationale Polytechnique

المدرسة الوطنية المتعددة التقنيات
BIBLIOTHEQUE — المكتبة
Ecole Nationale Polytechnique

Mémoire de fin d'études
pour l'obtention du diplôme d'ingénieur d'Etat
en Génie Industriel

Thème

Approche statistique du problème du voyageur de
commerce

Proposé et dirigé par :

M^r : A. OUABDESSELAM

M^r : T. LAMRAOUI

Etudié par :

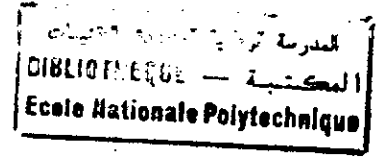
A .AOUFI

A .SOUAYEB

SESSION : Juin 1996

E.N.P.10, Avenue Hacén Badi El-harrach-Alger.

Dédicaces



*Je dédie ce modeste travail à
mes très chers parents,
mon frère et mes sœurs, à toute ma famille,
mes amis: Mohamed...
et tous ceux qui me sont chers.*

Ahmed

*A la mémoire de ma mère ,
mon très cher père,
mes frères et sœurs; spécialement Souâd et mes amis,
dont l'amour et le soutien , tout au long de ces années ,
ont permis de réaliser ce modeste travail.*

Abdelali



Remerciements

A vrai dire, les mots nous manquent pour exprimer nos profonds remerciements à M^r.OUABDESSELAM qui nous a honoré en nous dirigeant dans notre travail.

Nos remerciements vont également à notre co-promoteur M^r. LAMRAOUI pour ces précieuses et pertinentes suggestions.

Nous tenons à remercier aussi l'ensemble des enseignants qui ont contribué à notre formation , et aussi ceux d'entre eux qui ont eu l'aimabilité d'examiner ce modeste travail.

Enfin, nous ne saurions oublier de remercier tous ceux qui nous ont aidé de près ou de loin.

PLAN



Introduction.

1. Préliminaires.

- 1.1. Quelques généralités sur les graphes.
- 1.2. Rappel de quelques notions probabilistes et statistiques.

2. Présentation du problème du voyageur de commerce

- 2.1. Enoncé du problème.
- 2.2. Quelques variantes du TSP.
- 2.3. Approches de résolution.
- 2.4. Historique de l'approche statistique du TSP.

3. Comment générer des réseaux aléatoirement ?

- 3.1. Méthodes de génération des nombres aléatoires.
- 3.2. Création d'échantillons fictifs.
- 3.3. Test de l'hypothèse aléatoire
 - Procédure(1).
 - Résultats et interprétation.

4. Estimation statistique de la solution optimale par intervalle de confiance

- 4.1. Résultat de Fisher, Tippett et Fréchet.
- 4.2. Choix de l'heuristique.
- 4.3. Intervalle de confiance de la solution optimale
 - Procédure (2).
 - Résultats et interprétations.

5. Application de la méthode des moments

- 5.1. Calcul des moments.
- 5.2. Objectif.
- 5.3. Etude de la variation de « K » en fonction de la taille de l'échantillon.
 - Procédure(3).
 - Résultats et interprétations.

5.4. Etude de la variation de « W » d'un problème à un autre.

- Procédure (4).

- Résultats et interprétations.

5.5. Le modèle.

5.6. Vérification du modèle.

Conclusion.

Bibliographie.

Annexes.

1. Description de UNIFF.
2. Description de CURVE .
3. Test de l'hypothèse aléatoire .
4. La variation de K en fonction de la taille de l'échantillon .
5. La variation de W en fonction du nombre de villes .
6. L'erreur relative de la solution calculée .
7. Disquette d'accompagnement

INTRODUCTION

La pratique de la statistique a beaucoup évolué durant cette dernière décennie . Elle prend une place de plus en plus remarquable et importante dans presque tous les domaines de la science contemporaine.

Parmi les innombrables applications des méthodes statistiques , figurent les problèmes de l'optimisation combinatoire dont les solutions sont en nombre fini mais considérablement élevé.

Cette dernière catégorie de problèmes conduit à un problème majeur ; c'est celui du « voyageur de commerce ».

Dans notre mémoire , nous aborderons une approche statistique de celui - ci , à travers laquelle , nous espérons fournir une contribution à l'élargissement de l'interface entre la statistique et l'optimisation combinatoire .

Notre travail consiste à reprendre les travaux de GOLDEN, publiés dans son article (A Statistical approach to the TSP, 1977 b) et à essayer de développer une application de la méthode des moments.

Le cheminement adopté, dans ce mémoire s'articule autour de cinq parties :
Une première partie consistera à présenter les concepts de la théorie des graphes qui seront utilisés ainsi que quelques notions de probabilité et de statistique.

Une seconde partie sera consacrée à la présentation du problème du voyageur de commerce , à l'énumération de quelques unes de ses variantes, et à un aperçu de certaines approches de sa résolution.

La troisième partie sera axée sur la méthode de génération des réseaux (problèmes) aléatoires et test de l'hypothèse aléatoire.

Elle sera suivie d'une partie qui consistera à reprendre les travaux de GOLDEN, où une méthode d'estimation de la solution optimale par intervalle de confiance sera étudiée.

Dans une dernière partie nous développerons , une application de la méthode des moments en explicitant les différentes procédures élaborées et les résultats obtenus.

« Qui veut aller loin ménage sa monture »

CHAPITRE I

PRELIMINAIRES

- Quelques généralités sur les graphes.
- Rappel de quelques notions probabilistes et statistiques.

1 - Préliminaires

Le but de cette partie est de rappeler les principales notions qui seront utilisées dans le mémoire .

1.1 Généralités sur les graphes

Soit un graphe $G(x, u)$ caractérisé par :

- l'ensemble X ayant pour éléments les sommets (nœuds) .
- l'ensemble U ayant pour éléments des couples ordonnés de sommets appelés arcs.

Notons qu'un graphe peut être déterminé par l'ensemble des sommets et une matrice carrée C dont les éléments C_{ij} sont les coûts (distances) entre le sommet i et j telle que $i, j = 1, 2, \dots, n$, avec n : le nombre de sommets .

Définitions

♦ Chaîne de longueur N :

C'est une séquence de N arcs (u_1, u_2, \dots, u_N) de G telle que chaque arc ait une extrémité commune avec l'arc précédent , et l'autre extrémité en commun avec l'arc suivant.

♦ Cycle:

C'est une chaîne dont les extrémités coïncident.

♦ Chemin:

C'est une chaîne dont tous les arcs sont orientés dans le même sens.

♦ Circuit:

C'est un chemin dont les extrémités coïncident.

♦ Circuit hamiltonien (cycle hamiltonien) :

C'est un circuit (un cycle) qui passe par chacun des sommets de G une fois et une fois seulement .

1.2 Rappel de quelques concepts probabilistes et statistiques

a - La loi de Weibull

une variable aléatoire X est distribuée selon une loi de Weibull de paramètres: $a \geq 0$; $b > 0$; $c \geq 0$ si et seulement si:

- elle est continue ;
- elle prend des valeurs dans l'intervalles $[a , + \infty [$.
- sa densité de probabilité est donnée par :

$$f(x) = \frac{c}{b} \left(\frac{x-a}{b} \right)^{c-1} \exp \left[- \left(\frac{x-a}{b} \right)^c \right]$$

avec : a : paramètre de position.

b : paramètre d'échelle.

c : paramètre de forme.

La fonction de répartition est :

$$F(x) = Prob[X \leq x] = 1 - \exp \left[- \left(\frac{x-a}{b} \right)^c \right]$$

b -Théorème de Fisher , Tippett et Fréchet :

« La distribution de la plus petite (ou de la plus grande) valeur de n variables aléatoires indépendantes admettant même loi de probabilité tend vers la loi de Weibull lorsque n tend vers l'infini » [9].

c - Estimation ponctuelle des paramètres:

Soit le modèle $(X , P_{\theta} , \Theta \ni \theta)$ tel que :

X : est une variable aléatoire de loi P_θ , $\Theta \ni \theta$

Θ : est l'espace des paramètres.

Définition:

On appelle *estimateur* de θ l'application T :

$$(x_1, x_2, \dots, x_n) \xrightarrow{T} T(x_1, x_2, \dots, x_n).$$

La valeur de T en l'échantillon observé, (x_1, x_2, \dots, x_n) est l'*estimation* du paramètre θ

N.B : On utilisera la notation $\hat{\theta}$ pour un estimateur de θ .

L'estimation ponctuelle des paramètres d'une loi de probabilité, peut se faire de deux manières:

- Méthode des moments .
- Méthode de maximum de vraisemblance.

Dans le cas où les paramètres à estimer sont les paramètres de loi de Weibull , les moments sont donnés par des formules très compliquées. c'est pour cela que la méthode de maximum de vraisemblance est préférée.

d - Estimation par intervalle de confiance:

Soit $\alpha \in [0, 1]$ donné, on appelle intervalle (ou région) de confiance pour le paramètre θ de niveau de confiance $1-\alpha$, la famille non vide des ensembles $C_x(\theta)$ tel que:

$$\text{Prob}[\theta \in C_x(\theta)] = 1 - \alpha$$

Les bornes de $C_x(\theta)$ dépendent de l'échantillon (x_1, x_2, \dots, x_n) considéré.

e - Test d'hypothèse:

C'est une règle de décision qui permet d'accepter ou de refuser une hypothèse en vu d'un échantillon et avec un risque de se tromper.

f - Test de Kolmogorov - Smirnov (K - S) :

Ce test est basé sur la comparaison de la fonction de répartition empirique ou observée de l'échantillon , notée $F_n(x)$ et de la fonction de répartition théorique notée $F(x)$ de la population.

Principe du test:

On détermine l'écart maximal (en valeur absolue) existant entre la fonction de répartition empirique et la fonction de répartition théorique ,i.e on constitue:

$$D = \max_x |F_n(x) - F(x)|$$

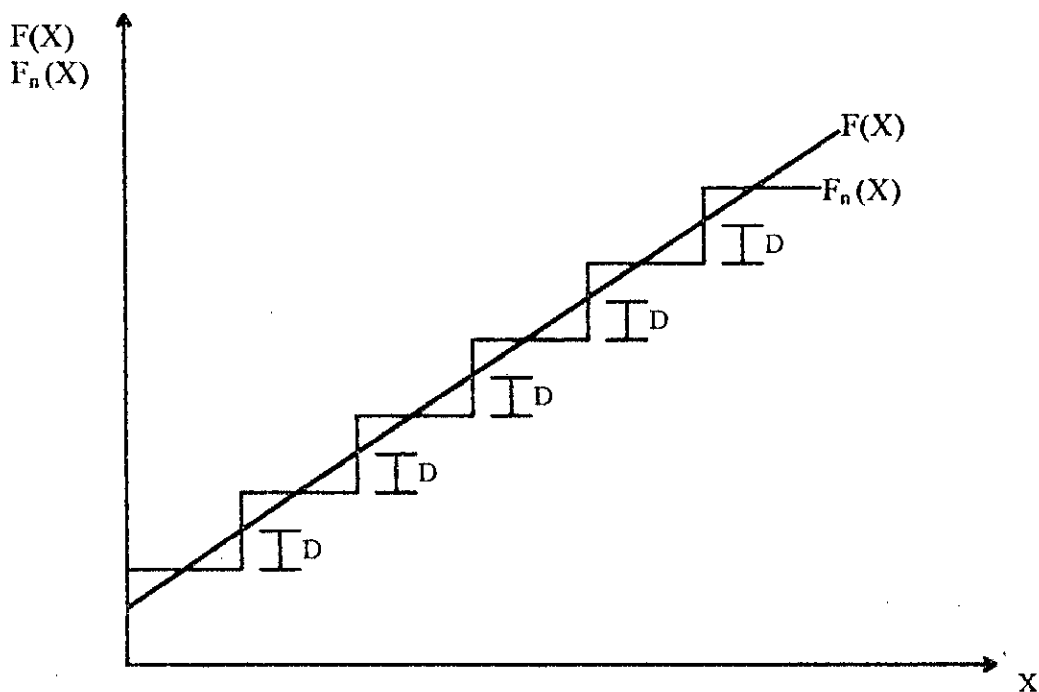


Fig 1.1 - Illustration du principe du test de Kolmogorov - Smirnov

Exemple

Pour un échantillon d'effectif $n > 35$ et en prenant des seuils égaux respectivement à 0.05 et 0.01, on compare la valeur de D aux valeurs $1.36 \sqrt{n}$ et $1.63 \sqrt{n}$.

Si $D > 1.36 \sqrt{n}$ (respectivement $1.63 \sqrt{n}$), on rejette hypothèse d'ajustement au seuil 0.05 (respectivement 0.01).

Les valeurs 1.36 et 1.63 sont des valeurs théoriques tirées de la table de la loi de K-S.

g - Ajustement de courbes et corrélation

En pratique, il arrive souvent que l'on mette en évidence une relation entre deux variables x , y (ou plus) et que l'on souhaite exprimer cette relation sous forme mathématique en déterminant une relation entre ces variables.

L'étude d'un échantillon nous donne les valeurs x_1, x_2, \dots, x_n de la variable dite indépendante, et les valeurs correspondantes y_1, y_2, \dots, y_n de la variable dite dépendante.

On émet l'hypothèse que les données suivent approximativement une courbe continue; le coefficient de corrélation noté r nous permet de mesurer la qualité d'ajustement par une courbe de régression non linéaire (il existe un coefficient de corrélation pour les régressions linéaires).

La formule de r est donnée par:

$$r^2 = \frac{\text{Ecart expliqué}}{\text{Ecart total}} = \frac{\sum_{i=1}^n (y_{\text{est}} - \bar{y})^2}{\sum_{i=1}^n (y - \bar{y})^2}, \quad -1 \leq r \leq 1$$

avec: \bar{y} : la moyenne des observations y_i , $i = 1, 2, \dots, n$.

y_{est} : la valeur estimée de y pour une valeur donnée de x , obtenue à partir de la courbe de régression de y sur x .

La meilleure courbe d'ajustement est celle qui donne un coefficient de corrélation r proche de 1.

2- Présentation du problème du voyageur de commerce

2-1 Enoncé du problème

Le problème classique du voyageur de commerce, (*the traveling salesman problem (TSP) en anglais*); s'énonce comme suit :

Un voyageur de commerce doit visiter N villes A_1, A_2, \dots, A_N ; chacune une fois et une fois seulement ; et revenir à son point de départ . Connaissant la matrice des coûts de transport $C=(C_{ij})$, où C_{ij} est le coût de transport de la ville A_i à la ville A_j . Dans quel ordre doit - il visiter ces villes pour que le coût total soit minimal ?

En termes de la théorie des graphes, ce problème revient à rechercher un circuit hamiltonien de longueur totale minimale dans le graphe complet G construit sur l'ensemble des sommets (les villes), les arcs étant munis des coûts C_{ij} .

Propriétés

La matrice des coûts $C = (C_{ij})$ nous permet de distinguer les cas suivants:

- Si $C_{ij} = C_{ji}$ pour tout A_i et A_j appartenant à l'ensemble des villes ; alors le TSP est symétrique ; si non, il est asymétrique.
- Si $C_{ij} < C_{ih} + C_{hj}$ pour tout A_i, A_j et A_h appartenant à l'ensemble des villes ; alors le TSP satisfait à l'inégalité triangulaire .

Remarque

Les problèmes du voyageur de commerce euclidiens sont symétriques et satisfont à l'inégalité triangulaire.

2.2 Quelques variantes du TSP

Le TSP apparaît dans plusieurs cas. Considérons, par exemple, les situations suivantes:

Exemple 1:

Dans un système d'exploitation automatisé, une machine doit régulièrement effectuer N types de tâches différentes x_1, x_2, \dots, x_N . On connaît de plus les différentes durées de commutations entre deux tâches x_i et x_j quelconques.

Trouver une séquence de x_i qui minimise le temps total des commutations.

Exemple 2:

Soit un réseaux informatique par commutation de paquets comprenant N nœuds indicés de 1 à N . On désire installer au nœud 1, un logiciel permettant de distribuer des instructions de travail à chacun des autres nœuds. Mais cette distribution ne se fait pas en « étoile » c'est-à-dire directement du nœud 1 à chacun des nœuds, au contraire le mode opératoire choisi est le suivant:

Le nœud 1 envoie un paquet dit « paquet de service » à un seul nœud, soit i . Ce paquet contient les instructions de travail et la « liste ordonnée » décrivant l'ordre suivant lequel les nœuds doivent être successivement informés.

Ainsi lorsque ce nœud i a reçu le paquet de service, il le transmet au nœud j qui est désigné comme le suivant de i dans la liste ordonnée et ainsi de suite...

Le dernier nœud atteint renvoie le paquet de service au nœud 1, l'informant ainsi que tous les nœuds ont reçu leurs instructions de travail.

En chaque nœud, bien évidemment, le paquet de service ne passe qu'une fois et une seule. Constituer la liste ordonnée des nœuds, de telle sorte que le temps total d'acheminement des messages soit minimal.

Exemple 3 :

Étant données N tâches de durée a_i , $i = 1, 2, \dots, N$. Soit d_i la date où la tâche i doit être terminée (date souhaitée). Déterminer la séquence des tâches sur une machine unique de façon à minimiser le retard maximum c'est-à-dire à minimiser $\max_{1 \leq i \leq N} (t_i - d_i)$ où t_i est la date où la tâche i est effectivement terminée, sous les contraintes que toutes les tâches doivent être réalisées chacune une fois et une seule.

Les exemples précédents ne représentent que quelques situations où le problème posé est un problème de genre voyageur de commerce. Il existe d'autres variantes, le lecteur intéressé est invité à consulter les ouvrages [8], [12] et [20].

Mais il faut mentionner que l'importance du TSP ne provient pas seulement de la diversité de son apparition, mais du fait qu' il représente un exemple typé pour d' autres problèmes de l'optimisation combinatoire.

2.3 Approches de résolution

Le TSP est l'un des problèmes de l'optimisation combinatoire pour lesquels un algorithme efficace (temps de calcul polynomial) n'existe pas. Tous les algorithmes, connus jusqu'à nos jours, nécessitent un temps de calcul qui augmente exponentiellement en fonction du nombre N de villes.

Théoriquement le problème peut être résolu par la génération des $(N - 1) !$ circuits.

Cependant, une telle approche, est extrêmement inefficace, par exemple un réseau de 20 villes seulement a $19 !$ circuits, de l'ordre de 10^{17} ; l'algorithme d'énumération nécessite des années de calcul [28].

Il existe deux autres approches de résolution :

2.3.1 Approches dites exactes

J.D.C, Little en 1963 [17]; a été parmi les premiers chercheurs qui ont présenté une méthode rigoureuse d'optimisation s'appliquant à ce problème.

D'autres méthodes ont été proposées, parmi celles-ci on trouve les travaux de Bertier ([4], 1965), Roy ([22], 1966), Held et Karp etc...

Le principe de ces approches consiste à déterminer par une méthode donnée:

- La méthode des plans sécants (*cutting planes*).
- Relaxation lagrangienne.
- Méthode des sous gradients, etc...

des bornes inférieurs utilisées ensuite dans un algorithme de séparation et évaluation (*branch and bound*).

Les techniques de programmation dynamique et la programmation en nombres entiers s'attaquent à des problèmes de petites tailles [26].

2.3.2 Approches heuristiques

A cause de la difficulté de la résolution optimale du TSP , les méthodes exactes ont une portée pratique insuffisante. C'est la raison pour laquelle nous nous sommes concentrés beaucoup plus sur les approches heuristiques.

Les plus connues peuvent être réparties en trois classes principales:

- Procédures de construction de tournées (*tour construction procedures*).
- Procédures d'affinage de tournées (*tour improvement procedures*).
- Procédures mixtes (*composite procedures*).

a. Procédures de construction de tournées

On y retrouve:

◊ *Nearest neighbor procedure* (Rosenkrantz, Stearns et Levis [21]; 1977) dont le principe est le suivant:

Etape1: démarrer la tournée à partir d'un nœud quelconque ;

Etape2: trouver le sommet le plus proche du dernier sommet de la tournée et l'ajouter à celle-ci ;

Etape3: répéter l'étape 2 jusqu'à épuisement des sommets puis joindre le sommet initial au sommet final .

◊ *Clark and Wright Savings* [7]: basée sur le principe de l'économie; cette procédure comme suit :

Etape1: retenir un sommet quelconque; noté 1 ;

Etape2: calculer les économies: $S_{ij} = C_{1i} + C_{1j} - C_{ij}$, pour $i,j=2,3,\dots,N$ et les classer par ordre décroissant ;

Etape3: relier les sommets selon l'ordre ainsi établi .

◊ *Insertion procedures*: une procédure d'insertion considère une tournée formée de « k » sommets, $2 \leq k \leq N$ à la $k^{\text{ème}}$ itération, et détermine le prochain sommet à insérer (étape de sélection) puis l'endroit d'insertion (étape d'insertion).

Il existe une longue série de procédures d'insertion qu'il serait fastidieux d'énumérer ici; parmi celles-ci :

- *Nearest neighbor.*

- *Cheapest insertion.*

- *Arbitrary insertion.*

- *Farthest insertion, etc...* pour plus de détail, voir [14].

b. Procédures d'affinage de tournées

Les meilleurs heuristiques pour le TSP sont probablement celles qui consistent à échanger (permuter) des branches ou arcs dans une tournée.

Introduites initialement par Lin [15 ,16] et généralisées ensuite, en collaboration avec Kernighan. Les procédures d'affinage; *2-opt* et *3-opt* (*r-opt* dans le cas général) agissent comme suit:

Etape1: trouver une tournée initiale (généralement de manière aléatoire) ;

Etape2: améliorer la solution (tournée) en substituant à un ensemble de *r*-arcs, un autre ensemble de même nombre d'arcs ;

Etape3: répéter l'étape 2 jusqu'à ce qu'il n'y ait plus de changement.

La solution est dite alors *r-optimale*.

Plus la valeur de *r* est élevée, meilleure est la solution ,mais l'effort de calcul augmente considérablement par la même occasion aussi [25], il est rare que *r* prenne des valeurs supérieurs à 3 dans la pratique .

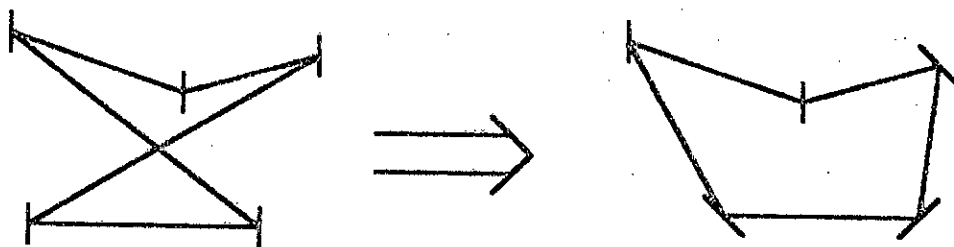


Fig.2.1 : Illustration d'une opération 2-opt.

◇ Procédures *or-optimale* : développée par Or [18] . Cette procédure est similaire en principe à *r - opt* , mais considère à la fois toutes les opérations d'échange portant sur deux arcs ou jonctions et celles conduisant à insérer une chaîne d'au plus trois sommets du graphe .

Elle plus rapide que la première [18].

c - Procédures mixtes:

Le principe de base est le suivant:

Etape1: obtenir une tournée initiale en utilisant l'une des procédures de construction;

Etape2: lui appliquer la procédure *2-opt*;

Etape3: lui appliquer la procédure *3-opt*;

Il existe plusieurs variantes qui consistent à ne pas passer par l'une des étapes (2) ou (3).

2.4 Historique de l'approche statistique du TSP

Une fois que l'une des méthodes heuristiques est mise en oeuvre, un autre problème non moins important apparaît, il s'agit de l'analyse de cette heuristique et de la solution qu'elle donne .

Plusieurs procédés sont alors proposés. Parmi ceux-ci, l'analyse empirique des heuristique, qui renferme entre autres, l'estimation statistique de la solution optimale par intervalle de confiance.

Parmi ces différentes procédures, on trouve [11]:

- La procédure donnée par Clough. Il a appliqué la théorie des valeurs extrêmes au problème d'estimation de l'intervalle de confiance de l'optimum global d'une fonction à plusieurs variables en utilisant la méthode de Monte-Carlo. L'intervalle est donné par la formule:

$$\text{Prob} \left[\left(\frac{n+c}{n-1} \right) \mu - \left(\frac{1+c}{n-1} \right) \frac{1}{n} \sum_{i=1}^n X_i \leq s \leq \mu \right] \approx 1 - \left(\frac{n-1}{n-c} \right)^{n-1}$$

avec: n : taille de l'échantillon ;

- c : réel quelconque ;
- s : solution optimale ;
- x : solutions données par l'heuristique 2-opt ;
- $\mu = \text{Min} \{x_i / 1 \leq i \leq n\}$.

- Mann et Al donnent une procédure où ils utilisent l'approximation par la loi Béta.

L'intervalle de confiance ainsi construit dépend de l'approximation.

- Une autre procédure, donnée par Robson et Whitlock, consiste à construire un intervalle de confiance par une troncature de la valeur inférieure d'une distribution uniforme.

Soit X_i la $i^{\text{ème}}$ plus petite valeur. On pose $\mu = \bar{X}_i$, l'intervalle de confiance est obtenu à un niveau de $100(1 - \alpha) \%$

$$\text{Prob} \left[\mu - \left(\frac{1 - \alpha}{\alpha} \right) (x_2 - \mu) \leq s \leq \mu \right] \approx 1 - \alpha$$

α : est le risque de se tromper .

L'estimation statistique de la longueur optimale des circuits a ensuite été traitée par de nombreux auteurs tels que: Golden (1977), Ladrinois(1982), Derigs (1983) et autres.

Tous ces différents chercheurs ont fait appel au théorème de Fisher, Tippet et Fréchet cité dans le premier chapitre.

« Prudence est mère de sûreté »

CHAPITRE III

COMMENT GENERER DES RESEAUX ALEATOIRES ?

- Méthodes de génération des nombres aléatoires .
- Création d'échantillons fictifs .
- Test de l'hypothèse aléatoire .
 - Procédure(1) .
 - Résultats et interprétation .

3 - Comment générer des réseaux aléatoires ?

Générer un réseaux aléatoire dans un plan (o, x, y) revient à générer les coordonnées des sommets appartenant à ce réseau d'une façon aléatoire .Concernant les longueurs des arcs , nous avons choisi les distances euclidiennes .

Dans ce chapitre, nous nous proposons d'examiner d'une part les modes classiques de génération des nombres au hasard, et, d'autre part, le mécanisme de création d'échantillons fictifs d'une loi de probabilité donnée .

3.1 Méthodes de génération des nombres au hasard

Il existe deux sortes de méthodes pour construire les nombre au hasard : les méthodes arithmétiques ,basées sur divers suites et les méthodes physiques .

3.1.1 Procédés arithmétiques

Ils sont sujets à caution, et les nombres ainsi générés sont souvent appelés nombres « quasi-aléatoires » [29] .

a. Méthode de Von Neumann

On choisit un nombre que l'on élève au carré. On prélève alors une partie centrale du résultat , que l'on élève à son tour au carré, et on itère le raisonnement.

Malheureusement ces nombres ne sont pas aléatoires car ils dépendent du départ choisi.

b. Méthode de Lehmer

On définit une suite (U_n) , $n \in \mathbb{N}$ de nombres entiers.

U_0 : entier positif arbitraire.

$$U_{n+1} = k U_n \text{ (modulo } m \text{)}$$

avec $m = 2^{31} - 1$ et $k = 23$ (m est appelé nombre de Mersenne).

c. Suites de Fibonacci

On définit une suite (U_n) , $n \in \mathbb{N}$

$$U_{n+2} = U_{n+1} + U_n \text{ (modulo } m \text{)}.$$

U_0, U_1 : entiers positifs arbitraires.

3.1.2 Procédés physiques

La table de Fisher et Yates donne 15000 chiffres obtenus des décimales de rang 15 à 19 des tables logarithmes.

La table de Kendall et Bobington Smith donne 100 000 chiffres obtenus par éclairage à intervalles irréguliers d'un disque divisé en 10 secteurs, chacun portant un chiffre de 0 à 9.

3.2 Création d'échantillons fictifs

Nous avons à notre disposition une table de nombres au hasard; et nous voulons en déduire un échantillon de nombres suivant une loi de probabilité donnée.

La méthode consiste à utiliser la fonction de répartition de cette loi.

Justification

Soit X une variable aléatoire réelle, $F(x)$ sa fonction de répartition :

$F(x) = P(X < x)$ supposée continue strictement croissante. On définit une variable aléatoire $Y = F(X)$.

Lemme (1):

Y suit une loi uniforme sur $[0, 1]$.

Lemme (2):

Les fractiles de Y sont les transformés par F des fractiles de X .

Soit x_α le fractile d'ordre α de x :

$$P(X < x_\alpha) = \alpha$$

$$P(X < x_\alpha) = P[F(X) < F(x_\alpha)]$$

D'où:

$$\alpha = P[Y < F(x_\alpha)]$$

et $F(x_\alpha)$ est le fractile d'ordre α de Y .

Application

On tire un échantillon (y_1, y_2, \dots, y_N) d'une loi uniforme sur $[0, 1]$ et on en déduit l'échantillon (x_1, x_2, \dots, x_N) qui est régi par la loi de fonction de répartition F , avec :
 $x_i = F^{-1}(y_i)$.

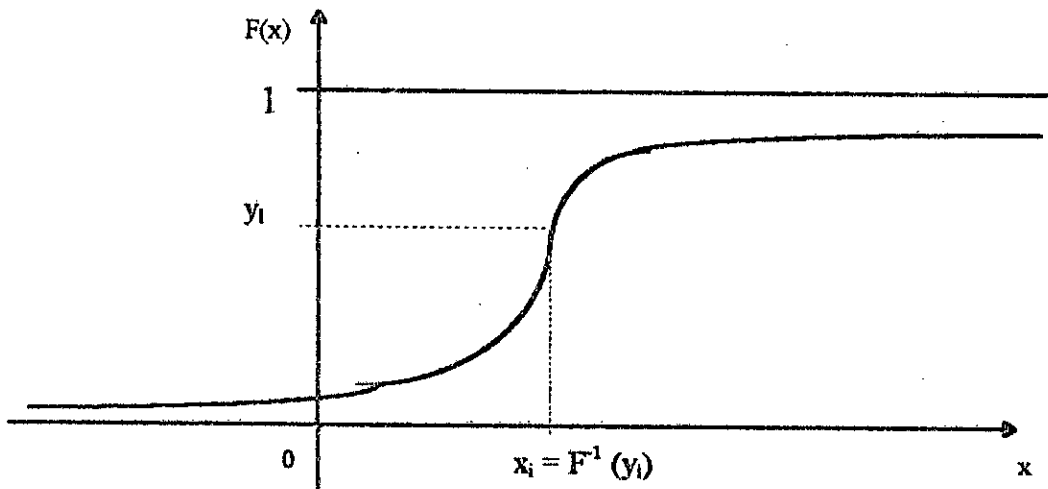


Fig.3.1: La fonction de répartition de la loi uniforme sur $[0, 1]$

Dans le cas d'une variable discrète, il n'existe généralement pas de valeur x_α correspondante exactement à une valeur de α .

Notons aussi qu'on peut tirer l'échantillon de n'importe quelle loi de probabilité (Normale, Poisson, ...).

3.3 Test de l'hypothèse aléatoire

Plutôt que de générer un échantillon fictif suivant une loi de probabilité (loi uniforme), il est plus fréquent d'avoir à tester l'hypothèse qu'un échantillon (x_1, x_2, \dots, x_N) est effectivement aléatoire, c'est-à-dire, par exemple, ne provient pas d'un modèle donné.

L'un des tests les plus efficaces est le suivant : l'échantillon (x_1, x_2, \dots, x_N) étant considéré comme ordonné par l'indice i ; $i = 1, 2, \dots, N$; on appelle point critique ou point de retournement un point x_k tel que :

$$x_k > x_{k-1} \quad \text{et} \quad x_k > x_{k+1}$$

ou

$$x_k < x_{k-1} \quad \text{et} \quad x_k < x_{k+1}$$

On peut établir (Kendall et Stuart, [13]) que le nombre Z de points critiques est une variable aléatoire qui converge en loi, sous l'hypothèse que, l'échantillon est effectivement aléatoire ; vers une loi normale :

$$(Z - \eta) / \gamma \text{ loi } N(0, 1)$$

avec :

$$\eta = (2N - 4) / 3$$

$$\gamma^2 = (16N - 29) / 90$$

Si l'on observe z points critiques ; on rejettera l'hypothèse aléatoire si :

$$\left| \frac{z - \eta}{\gamma} \right| > f$$

f : fractile d'ordre $1 - \alpha / 2$ de $N(0, 1)$.

Afin de vérifier que les réseaux, sur lesquels portera notre étude par la suite, sont générés d'une façon aléatoire, nous avons procédé comme suit :

a - Procédure (1) : Test de l'hypothèse aléatoire

- 1 - Fixer le nombre de villes $N=10$ villes (nœuds) ;
 - 2 - générer un échantillon fictif des abscisses (x_1, x_2, \dots, x_N) suivant une loi uniforme sur $[0, 1]$;
 - 3 - calculer z_1 le nombre de points critiques;
 - 4 - calculer η et γ ;
 - 5 - fixer un seuil de rejet α et déterminer la valeur correspondante f à partir de la table de la loi $N(0, 1)$;
 - 6 - si $t_1 = |(z_1 - \eta) / \gamma| > f$ rejeter l'hypothèse aléatoire avec le risque de se tromper α ;
 - 7- aller à (2), remplacer abscisses par ordonnées, z_1 par z_2 , t_1 par t_2 et répéter les étapes (3), (4), (5) et (6) ;
 - 8 - $N = N + 5$, si $N \leq 100$ aller à l'étape (2) .
- Sinon stop.

b - Résultats

Nous avons fait tourner cette procédure avec 8 séries de problèmes, chacune comporte 19 problèmes de tailles allant de 10 à 100 villes avec un pas de 5.

Les résultats de la première série sont présentés dans le tableaux 3.1 les autres sont dans l'annexe (3) .

Le lecteur intéressé trouvera les coordonnées de chaque problème dans la *disquette d'accompagnement* [annexe 7] .

N	η	γ	z_1	t_1	z_2	t_2
10	5.333	1.206	5	0.276	4	1.105
15	8.667	1.531	10	0.871	8	0.435
20	12.000	1.798	13	0.556	12	0.000
25	15.333	2.030	14	0.657	15	0.164
30	18.667	2.239	19	0.149	16	1.191
35	22.000	2.429	24	0.823	21	0.412
40	25.333	2.606	24	0.512	26	0.256
45	28.667	2.771	27	0.601	32	1.203
50	32.000	2.927	27	1.708	30	0.683
55	35.333	3.075	34	0.434	31	1.409
60	38.667	3.216	38	0.207	41	0.725
65	42.000	3.352	42	0.000	40	0.596
70	45.333	3.482	50	1.340	51	1.628
75	48.667	3.607	43	1.571	46	0.739
80	52.000	3.728	53	0.268	45	1.878
85	55.333	3.846	55	0.087	58	0.693
90	58.667	3.960	57	0.421	53	1.431
95	62.000	4.070	63	0.246	61	0.246
100	65.333	4.178	63	0.558	61	1.037

Tableau-3.1- Le test de l'hypothèse aléatoire (série n° 1)

Le risque de se tromper $\alpha = 0.05$ et le fractile d'ordre $1 - \alpha / 2$; $f = 1.960$ tiré de la table de la loi $N(0, 1)$.

Les résultats montrent que tous les 152 réseaux sur lesquels portera notre étude par la suite sont générés d'une façon aléatoire avec le risque de se tromper $\alpha = 0.05$.

« Estimer ne coûte rien .
Estimer incorrectement coûte cher »
vieux proverbe chinois.

CHAPITRE IV

ESTIMATION STATISTIQUE DE LA SOLUTION OPTIMALE PAR INTERVALLE DE CONFIANCE .

- Résultat de Fisher, Tippett et Fréchet
- Choix de l'heuristique.
- Intervalle de confiance de la solution optimale
 - Procédure (2).
 - Résultats et interprétations.

4. Estimation statistique de la solution optimale par intervalle de confiance en utilisant la loi de Weibull.

Les procédures d'estimation de la solution optimale par intervalle de confiance, citées dans le chapitre (II) présentent des inconvénients et paraissent inefficaces [26].

Une autre procédure plus efficace développée par Golden [10], fait appel à un résultat fondamental établi par Fisher , Tippet et Fréchet en 1928, dans le cadre de la théorie de la distribution statistique des valeurs extrêmes , afin de déterminer plus exactement une valeur estimée de l'optimum.

4.1 Résultat de Fisher, Tippet et Fréchet

Soit « n » échantillons, chacun de taille « s » tirés de façon aléatoire de l'ensemble de tous les circuits possibles (PM) qui est borné par dessous. Dans chaque échantillon « i » ; $i = 1, 2, \dots, n$ il y a une valeur minimale x_i , et la plus petite valeurs μ parmi les $n.s$ observations est la plus petite des n valeurs minimales x_i .

$$\mu = \text{MIN} \{x_i / 1 \leq i \leq n\}.$$

Fisher , Tippet et Fréchet ont démontré que lorsque n est suffisamment grand , la distribution des x_i tend vers la loi de Weibull [Fig IV. 1].

Dans le cas du TSP où le nombre de villes à visiter est N , la population mère (PM) est composée de $(N-1) !$ circuits avec un coût total non inférieur au *paramètre de position \hat{a}* qui représente *la solution optimale estimée du TSP*.

Les paramètres de la loi de Weibull a , b et c sont estimés en utilisant le logiciel UNIFIT [annexe 1] à partir de l'échantillon $(X_1 , X_2 , \dots , X_n)$ obtenu par l'heuristique choisie .

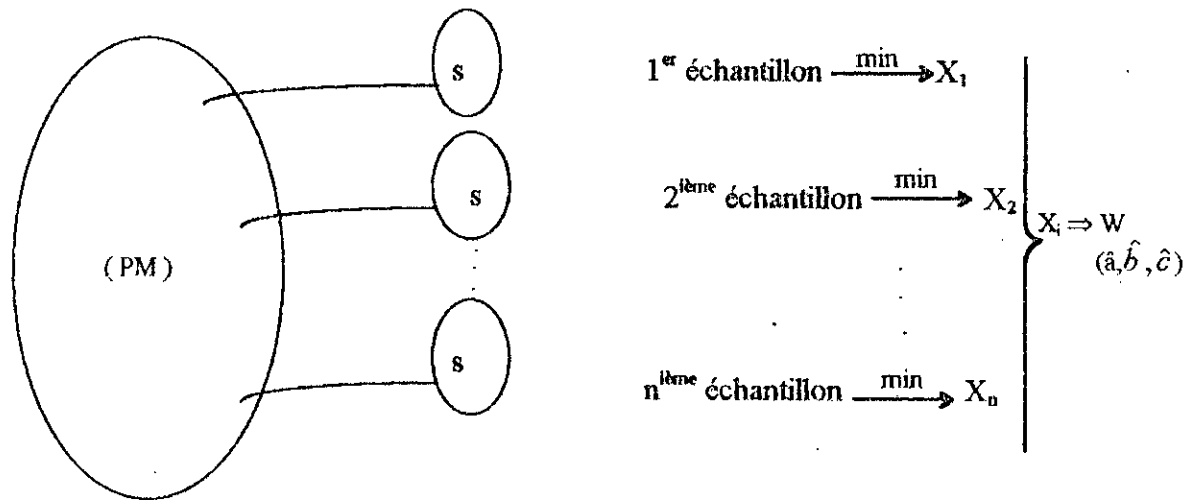


Fig IV . 1 -Description du résultat de Fisher, Tippett et Fréchet

4.2 Choix de l'heuristique

En prenant en compte les critères de choix d'un heuristique [14]:

- la rapidité (temps d'exécution);
- la flexibilité;
- la facilité de l'implémentation;
- et la simplicité;

et en prenant en considération les conclusions suivantes de Golden et Alt [10]:

- ◇ Plusieurs procédures de construction de tournées peuvent trouver une solution avec 5 % à 7 % de l'optimalité, ces procédures sont utilisées quand une solution effective est désirée.
- ◇ Les procédures d'affichage, en particulier *2-opt* et *3-opt* citées dans le chapitre II avec des solutions initiales aléatoires, opèrent avec approximation la même efficacité et efficacité que la meilleure des procédures de construction .

Chap. IV Estimation statistique de la solution optimale par intervalle de confiance

- ◇ La procédure mixte peut trouver une solution du TSP avec 2 % à 3 % de l'optimalité avec une grande régularité . Cependant cette procédure de calcul demande un temps de calcul sensiblement élevé, comparé aux procédures d'affinage et de construction .
- ◇ Pour trouver efficacement une solution du TSP avec 1% à 2% de l'optimalité, il faut répéter l'application de la procédure mixte.

L'heuristique que nous avons choisi et qui va être utilisé dans tout ce qui suit est la procédure 2-opt.

Et comme c'est noté dans le chapitre II, la première étape de cette procédure consiste à trouver par l'une des méthodes de constructions, citées avant, une tournée initiale.

L'inconvénient majeur de ces dernières est qu'elles donnent un nombre restreint de tournées (circuits) possibles, cet inconvénient nous a orienté vers le choix des tournées initiales aléatoires.

Principe de calcul de la procédure 2-opt.

Soit $H = \{ A_1, A_2, \dots, A_N \}$ le circuit hamiltonien initial.

$X = \{ (A_i, A_j) \}$, ensemble d'arcs de H substitué par Y .

$Y = \{ (A_p, A_q) \}$ s'il y a amélioration de la solution initiale, $H' = (H - X) \cup Y$: la nouvelle tournée raffinée.

Il y a $N(N-1)/2 - N = N(N-3)/2$ tournées H' pour un H donné.

Pour chacune de celles - ci :

soit δ : l'amélioration réalisée.

$$\delta = W(H) - W(H'),$$

$W(H)$: longueur du circuit H .

On cherche parmi les circuits H' celui qui maximise δ .

Si $\delta_{\max} \leq 0$, nous avons la solution finale donnée par 2-opt .

Si non, nous utilisons la solution correspondante comme initiale, et nous répétons la procédure 2-opt jusqu'à ce que δ_{\max} soit non positif.

4.3 Intervalle de confiance de la solution optimale du TSP

La procédure de construction de l'intervalle de confiance est obtenue en s'appuyant sur la fonction de répartition de la loi de Weibull qui nous permettra d'écrire:

$$\text{Prob} [\mu \leq a + b] = 1 - \text{Prob} [\mu > a + b]$$

$$\text{Prob} [\mu \leq a + b] = 1 - \text{Prob} [\text{Min} \{ X_i / 1 \leq i \leq n \} > a + b]$$

$$\text{Prob} [\mu \leq a + b] = 1 - [1 - F_{X_1}(a + b)][1 - F_{X_2}(a + b)] \dots [1 - F_{X_n}(a + b)]$$

$$\text{Prob} [\mu \leq a + b] = 1 - \exp(-n).$$

n : étant la taille de l'échantillon.

ou :

$$\text{Prob} [\mu - b \leq a \leq \mu] = 1 - \exp(-n).$$

L'intervalle de confiance est ainsi défini avec un risque de se tromper $\exp(-n)$, il permet de comparer la qualité de la solution des différentes heuristiques.

Et le rapport b / μ représentant la largeur relative de l'intervalle, nous permet d'avoir une idée sur l'efficacité de l'heuristique utilisée.

En effet, plus b / μ est petit, plus l'heuristique utilisée est efficace [10].

Dans ce chapitre notre objectif consiste à retrouver les résultats de Golden [10], pour cela nous avons procédé comme suit :

Chap. IV Estimation statistique de la solution optimale par intervalle de confiance

n - Procédure (2) : Estimation de la solution optimale en utilisant la loi de Weibull

Phase (1) : Construction de l'échantillon

- 1 - Fixer N le nombre de villes à visiter ;
- 2 - générer un réseau aléatoirement ;
- 3 - fixer un point de départ aléatoirement ;
- 4 - construire une tournée initiale aléatoire à partir du point de départ choisi ;
- 5- appliquer *2-opt* pour avoir la solution X_i , $i = 1, 2, \dots, n$; n étant la taille de l'échantillon.
- 6 - répéter les étapes 4, 5 et 6 n fois, pour obtenir l'échantillon (X_1, X_2, \dots, X_n) .

Phase (2) : Utilisation de UNIFIT

UNIFIT nous permet, à partir de l'échantillon (X_1, X_2, \dots, X_n) d'ajuster les données à l'une des 13 distributions continues.

- 1 - Estimer les paramètres de la loi de Weibull ; noter les \hat{a} , \hat{b} et \hat{c} ;
- 2 - appliquer le test de Kolmogorov-Smirnov à un seuil α pour comparer la fonction de répartition ajustée à la fonction de répartition théorique de la loi de Weibull ;
- 3 - si le test accepte l'ajustement au seuil choisi ; calculer l'intervalle de confiance $[\lambda, \mu]$ tel que :
 $\lambda = \mu - \delta$, le risque de se tromper étant e^{-n} .
Si non aller à l'étape (3) de la phase 1.

Phase (3) :

- 1 - Répéter les phases 1 et 2 pour des problèmes de tailles différentes ;
- Stop.

Chap. IV Estimation statistique de la solution optimale par intervalle de confiance

b - Résultats

Les résultats de calculs de cette procédure sont présentés dans le tableau IV. 1 ainsi que ceux trouver par Golden (tableau IV. 2).

N	n	μ	\hat{a}	\hat{b}	\hat{c}	λ	K-S	Val-Cri*
30	25	435.45	435.40	17.46	1.45	417.99	0.75	1.33
40	50	521.57	521.56	33.61	1.33	487.96	1.11	1.33
50	50	569.61	567.93	48.62	1.25	520.99	0.97	1.33
60	50	623.28	621.85	44.75	1.81	578.53	0.82	1.33
70	50	710.49	710.47	36.49	1.71	674.00	0.75	1.33
80	50	730.59	730.38	39.37	1.91	691.22	1.09	1.33
90	50	755.10	754.49	38.79	2.24	716.31	0.65	1.33
100	50	791.63	791.27	54.53	1.56	737.10	0.66	1.33

Tableau - IV .1 - Résultats expérimentaux de la procédure (2).

N	n	μ	\hat{a}	\hat{b}	\hat{c}	λ	K-S	Val-Cri
70	25	659.39	650	37.78	1.95	621.61	0.083	0.27
80	25	700.46	695	37.77	1.56	662.69	0.097	0.27
90	25	747.24	740	35.74	2.01	711.50	0.111	0.27
100	25	785.73	720	117.42	4.83	668.31	0.114	0.27
110	30	832.44	770	111.65	6.18	720.79	0.106	0.24
120	30	915.36	870	82.30	4.90	833.06	0.086	0.24
130	30	917.56	900	91.72	2.00	825.84	0.084	0.24

Tableau - IV. 2 - Résultats expérimentaux trouvés par Golden dans l'article [10].

- N : nombre de villes .
- n : taille de l'échantillon .
- μ : borne supérieure de l'intervalle de confiance de la solution optimale (solution heuristique) .
- \hat{a} , \hat{b} et \hat{c} : paramètre de la loi de Weibull .
- λ : borne inférieure de l'intervalle de confiance de la solution optimale .
- K-S : statistique de Kolmogorov-Smirnov .
- Val-Cri : valeur critique .

(*) : Valeurs calculées par le logiciel UNIFIT.

c - Interprétation des résultats de calculs

Les résultats du tableau IV.1 nous permettent essentiellement de répondre aux trois questions suivantes :

- (i) Qu'elle est l'efficacité et la précision de la technique proposée c'est-à-dire l'heuristique 2-opt ?
- (ii) Est-ce que l'hypothèse que les solutions heuristiques obtenues par 2-opt tendent approximativement vers la loi de Weibull ?
- (iii) Que peut-on dire en comparant les résultats du tableau IV.1 calculés par la procédure (2) à ceux trouvés par Golden [10] ?

Les résultats obtenus tendent à confirmer notre intuition ; des réseaux de 30 à 100 nœuds ont été générés (chap III) aléatoirement dans un carré de surface⁽¹⁾ 10 000 unités. Le temps d'exécution est très grand pour une exécution totale sur 486 DX4 100 MH. L'heuristique 2-opt converge dans tous les cas traités vers une solution du problème. Les paramètres a , b et c de la loi de Weibull ont été estimés et les observations du test de Kolmogorov - Smirnov tombent sous la valeur critique en un niveau de signification $\alpha = 0.05$.

Une comparaison entre le tableau IV.1 et le tableau IV. 2 montrent que les solutions sont du même ordre bien qu'on ait travaillé sur des réseaux générés aléatoirement .

(1) : Nous avons choisi la même surface que celle utilisée par Golden .

« Si tu as dix pas à faire vers quelqu'un ,
les neuf premiers pas ne sont que la moitié du chemin ».
proverbe arabe.

CHAPITRE V

APPLICATION DE LA METHODE DES MOMENTS

- Calcul des moments.
- Objectif.
- Etude de la variation de « K » en fonction de la taille de l'échantillon.
 - Procédure(3).
 - Résultats et interprétations.
- Etude de la variation de « W » d'un problème à un autre.
 - Procédure (4).
 - Résultats et interprétations

5-Application de la méthode des moments

Dans ce dernier chapitre , nous nous sommes contentés ; dans un premier temps , d'exposer les principaux résultats concernant le calcul des moments de l'ensemble des circuits hamiltoniens générés par un nombre de villes donné.

Et par la suite nous essayerons d'établir une corrélation entre la solution estimée, en utilisant la loi de Weibull (chap IV) et les moments ainsi calculés.

5.1 Calcul des moments [16]

a - Caractéristiques statistiques

$$\begin{matrix}
 & A_1 & A_2 & \dots & A_j & \dots & A_N \\
 \begin{matrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ \cdot \\ A_i \\ \cdot \\ \cdot \\ \cdot \\ A_N \end{matrix} & \begin{pmatrix} 0 & C_{12} & \dots & C_{1j} & \dots & C_{1N} \\ C_{21} & 0 & \dots & C_{2j} & \dots & C_{2N} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ C_{i1} & C_{i2} & \dots & C_{ij} & \dots & C_{iN} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ C_{N1} & C_{N2} & \dots & C_{Nj} & \dots & 0 \end{pmatrix}
 \end{matrix}$$

Considérons l'ensemble des circuits hamiltoniens reliant N villes A_1, A_2, \dots, A_N . Soit C_{ij} le coût du trajet reliant A_i à A_j . Les coûts respectifs des différents trajets sont représentés dans la matrice ci-dessus.

• Remarques

- 1- Nous pouvons, pour fixer les idées, supposer que la ville de départ et d'arrivée est A_1 , par suite, le nombre total de circuits est le nombre de permutations des nombres 2, 3, ..., N c'est donc $(N-1)!$.
- 2- Deux termes quelconques d'un circuit n'appartiennent ni à la même ligne, ni à la même colonne.
- 3- A chaque terme de la 1^{ère} ligne, par exemple C_{1h} , on peut associer $(N-2)$ termes de la h^{ième} ligne.
- 4- Le 1^{er} terme appartient à la 1^{ère} ligne, un exclu, le dernier appartient à la 1^{ère} colonne, un exclu.
- 5- Chaque terme de la 1^{ère} ligne est répété $(N-1)! / (N-1) = (N-2)!$ fois, il en est de même de chaque terme de chacune des lignes.
- 6- Le coût d'un circuit comprend N termes, satisfaisant aux conditions précédentes.

b - Moyenne de la distribution des circuits (moment d'ordre 1)

Soit x_i le coût du i^{ème} circuit, la moyenne des différents circuits s'écrit:

$$m = \frac{1}{(N-1)!} \sum_{i=1}^{(N-1)!} x_i$$

En tenant compte des remarques précédentes, on a:

$$m = \frac{1}{(N-1)!} (N-2)! \sum_{i \neq j} c_{ij} = \frac{1}{(N-1)} \sum_{i \neq j} c_{ij}$$

Ainsi, la moyenne des coûts des $(N-1)!$ circuits est égale au produit par $1 / (N-1)$ de la somme des termes de la matrice considérée.

Phase (3): L'ajustement.

1- Utiliser le logiciel CURVE [annexe (2)] pour définir la fonction $k = f (n)$.

-Stop.

c - Résultats

Un réseau de taille $N = 50$ villes a été généré.

Les résultats sont présentés dans le tableau (V -1) ainsi que le graphe [fig V-1] illustrant la fonction $k = f (n)$.

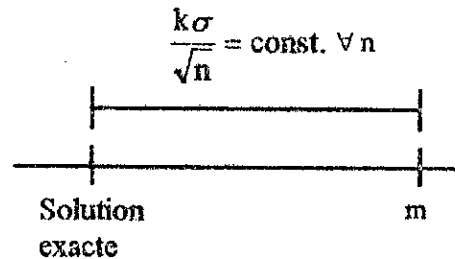
n	μ	\hat{a}	\hat{b}	λ	K-S	K	W
50	556.520	556.504	32.060	524.596	1.086	103.724	14.669
60	559.262	559.213	32.467	526.795	0.740	113.464	14.648
70	558.964	558.955	35.071	523.893	1.216	122.571	14.650
80	559.262	559.260	33.854	525.407	1.047	131.013	14.648
90	559.733	559.492	30.673	526.588	0.500	138.944	14.646
100	559.262	559.258	32.076	527.185	1.001	146.480	14.648
110	559.262	559.219	31.687	527.575	0.847	153.630	14.648
120	556.520	556.443	34.477	522.042	0.769	160.693	14.669
130	556.520	556.506	33.680	522.839	0.959	167.249	14.669
140	557.508	557.505	29.663	527.844	1.205	173.473	14.661
150	557.508	557.505	29.535	527.973	1.322	179.561	14.661
160	559.203	559.203	28.559	530.644	1.122	185.284	14.648
170	557.779	557.778	33.570	524.208	0.864	191.130	14.659
180	558.000	557.977	31.518	526.782	1.228	196.513	14.657
190	559.262	559.258	31.409	527.853	1.139	201.905	14.648
200	557.508	557.507	33.158	524.350	0.977	207.339	14.661
210	559.262	559.260	30.046	529.215	1.211	212.266	14.648
220	557.508	557.469	31.570	525.938	0.926	217.464	14.661
230	557.508	557.507	32.940	524.568	1.039	222.347	14.661
240	556.520	556.506	32.980	523.540	0.674	227.247	14.69
250	558.000	557.980	32.586	525.414	0.974	231.755	14.57

Tableau V . 1 - La variation de K en fonction de la taille de l'échantillon (n)

(problème de taille $N = 50$ villes , $m = 2481$, $\sigma = 131.197$)

Analytiquement

Si l'on dispose du coût minimal (la solution exacte) parmi les différents coûts des circuits et de leur moyenne , il est évident que l'écart entre la moyenne et le coût minimal est constant .



On peut conclure alors que l'heuristique utilisé c'est - à - dire *2-opt* converge dans tous les cas traités vers une solution proche de l'optimum .

La borne inférieure donnée par la méthode des moments devient :

$$LI = m - \frac{k\sigma}{\sqrt{n}} = m - W\sqrt{n}\frac{\sigma}{\sqrt{n}} = m - W\sigma$$

Et nous avons en examinant les résultats de l'annexe (4), remarqué que la valeur de W varie d'un problème à un autre.

Il est intéressant d' étudier cette variation lorsqu'on fait varier N , m et σ . Dans cet esprit une autre série d'essais expérimentaux a été effectuée de la façon suivante:

5.4 Etude de la variation de W d'un problème à un autre lorsque la taille de l'échantillon est fixe

a- Procédure (4) : Variation de W en fonction de N

N	m	σ	\hat{a}	k-s	W
10	492.087	45.892	+	+	3.943 ^(*)
15	704.802	71.334	+	+	5.828 ^(*)
20	1211.576	105.649	+	+	7.645 ^(*)
25	1415.011	109.87	471.099	1.226	8.591
30	1509.266	100.734	459.892	0.691	10.417
35	1905.637	126.452	480.349	0.965	11.271
40	2048.303	119.368	544.524	0.576	12.598
45	2153.684	121.519	506.738	0.656	13.553
50	2481.000	131.197	556.504	1.086	14.669
55	2958.760	153.144	579.458	1.105	15.536
60	2871.594	136.267	618.822	0.737	16.532
65	3378.552	156.964	655.155	1.154	17.350
70	3716.084	162.714	700.083	0.563	18.535
75	4130.633	177.095	671.075	0.976	19.535
80	4230.011	173.785	732.118	0.762	20.127
85	4213.11	159.363	785.732	0.502	21.507
90	4463.582	170.160	722.745	0.666	21.914
95	5081.287	189.617	748.871	1.141	22.848
100	5052.696	184.181	766.601	0.878	23.271

Tableau V. 2 - La variation de W en fonction du nombre de villes (série n° 1)

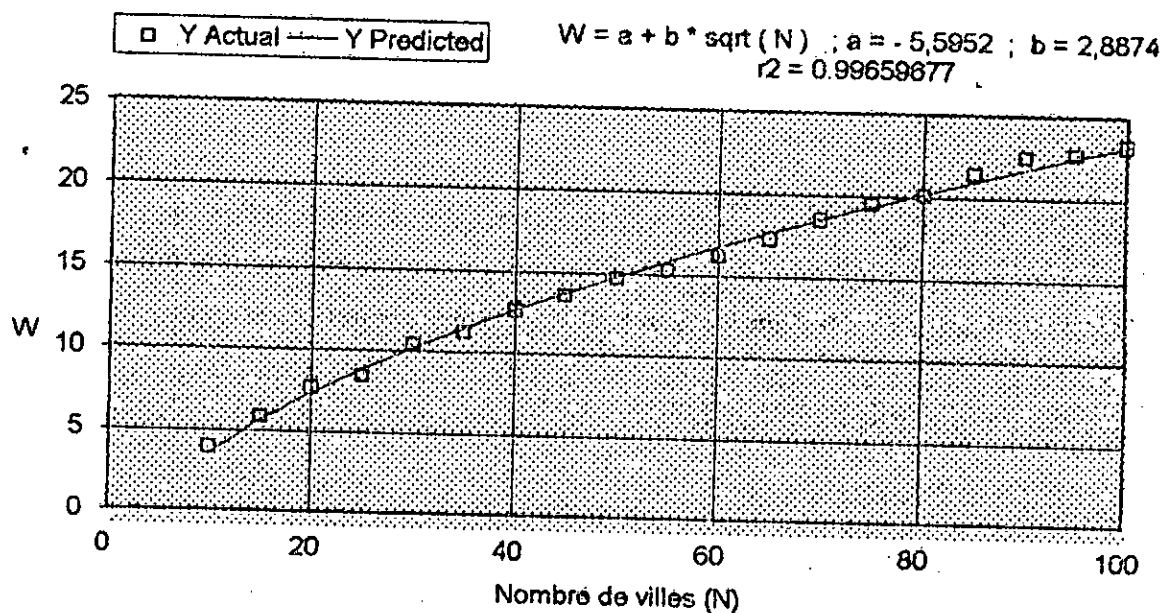


Fig V.2- La variation de W en fonction du nombre de villes (série n° 1)

(+): La solution (μ) est exacte .

(*) : Valeur correspondante à μ .

L'annexe - 5 - contient les résultats des autres séries.

Le graphe de la figure (V-2) montre que la meilleure courbe d'ajustement est celle qui est définie par la fonction :

$$W = p + q\sqrt{N}$$

p : constante égale à -5.595

q : constante égale à 2.887

Avec un coefficient de corrélation $r \approx 0,99$ et de même pour les autres séries [annexe (5)].

c- Analyse des résultats

Les valeurs des constantes p et q correspondantes aux différentes séries traitées sont présentées dans le tableaux suivant:

Série traitée	p	q	r ²
série (1)	- 5,595	2,887	0,997
série (2)	- 5, 779	2,919	0,996
série (3)	- 5,458	2,882	0,995
série (4)	- 5, 148	2,830	0,995
série (5)	- 5,480	2,876	0,996
série (6)	- 5,416	2,862	0,994
série (7)	- 5,993	2,940	0,996
série (8)	- 5,432	2,868	0,995

Tableau V . 3 - Valeurs de p et q correspondantes aux séries

Nous remarquons qu'il existe une petite différence entre les valeurs de p et q correspondantes aux séries , qui peut être justifiée par le fait que les valeurs de W sont calculées avec l'erreur de l'heuristique utilisé [voir le chapitre 4] et celle de l'estimation des paramètres de la loi de Weibull .

Face à ce problème nous avons proposé des bornes empiriques des constantes p et q à un seuil $\alpha = 0,05$ à l'aide du logiciel CURVE.

5.5 Le modèle

En fin le modèle décrivant la borne inférieure de la solution optimale en fonction des moments de l'ensemble des circuits hamiltoniens reliant N villes est :

$$LI \approx m - (p + q\sqrt{N})\sigma \quad (***)$$

avec :

m : moyenne des coûts des différents circuits.

σ : l'écart type entre ces coût.

N : nombre de villes.

p : constante appartenant à [-6.159 ; - 4.843] au seuil $\alpha = 0.05$.

q : constante appartenant à [2.795 ; 2.972] au seuil $\alpha = 0.05$.

5.6 Vérification du modèle

Afin de comparer la solution que donne ce modèle à ceux obtenus en utilisant l'estimation par intervalle de confiance de la solution optimale, nous allons recalculer par la formule (***) les valeurs de la solution estimée (le paramètre de position \hat{a} de la loi de Weibull) pour tous les problème traités .

Concernant le choix des constante p et q , nous avons pris la moyenne arithmétique des valeurs du tableau (V-2) correspondantes .

Les résultats de calculs portant sur la 1^{ère} série⁽¹⁾ sont présentés dans le tableau suivant :

N	μ	\hat{a} (1)	b	λ	solution calculée (2)	déviaton approximative (2)-(1)]*100/(1)
10	311.136	+	+	+	327.835	5.367 *
15	289.018	+	+	+	408.679	7.148 *
20	403.890	+	+	+	451.890	7.456 *
25	471.11	471.099	18.305	452.805	439.851	6.633
30	460.263	459.892	18.700	441.563	476.426	3.595
35	480.395	480.349	21.596	458.799	449.117	6.501
40	544.588	544.524	26.522	518.066	542.808	2.163
45	507.336	506.738	30.337	476.999	476.472	5.971
50	556.52	556.504	32.060	524.596	532.962	4.228
55	579.471	579.458	46.640	532.830	532.469	8.109
60	619.024	618.822	35.111	583.912	583.207	5.755
65	655.163	655.155	31.054	624.316	599.380	8.513
70	701.943	700.083	58.293	643.649	692.336	1.106
75	671.291	671.075	46.975	624.316	689.710	2.777
80	732.315	732.118	32.511	699.804	711.102	2.871
85	785.859	785.732	26.896	758.963	859.754	9.478
90	730.330	722.745	72.394	657.936	751.909	4.036
95	748.873	748.871	33.710	715.163	803.090	7.240
100	766.690	766.601	52.416	714.274	762.697	0.509

Tableau V . 4 - L'erreur relative de la solution calculée par la formule (**)
(série n° 1)

Nous remarquons que dans presque tous les cas la déviaton de la solution calculée par l'équation (**) par rapport à la solution estimée ne dépasse pas les 10 % .

(1) : les résultats de calculs portant sur les autres séries sont présentés dans l'annexe -6 -

(*) : la solution μ est exacte .

(*) : erreur par rapport à μ .

• Propositions

- ◇ La distribution de l'ensemble des solutions du TSP tend à être approximativement unique lorsque le nombre de villes est entre 10 et 100 quelque soit la façon dont les villes sont distribuées.

- ◇ Un bon choix des constantes p et q peut donner des solutions où l'erreur par rapport aux bornes de l'intervalle de confiance de la loi de Weibull sera plus petite.

- ◇ Si l'on dispose des solutions exactes correspondantes aux problèmes de tailles suffisamment grandes, on peut éliminer évidemment les erreurs dues à l'heuristique et de l'estimation, la seule erreur qui restera dans le calcul de la solution par la méthode des moments est celle de l'approximation de la distribution de l'ensemble des solutions à une distribution unique quelque soit la distribution des villes.

CONCLUSION

CONCLUSION

Dans le cadre de ce mémoire, notre objectif était de vérifier le bien fondé de l'idée de M^r. OUABDESSELAM qui consiste à déterminer une borne inférieure de la solution optimale du TSP moyennant les caractéristiques de la population mère, formée par les circuits hamiltoniens (ensemble de solution possibles du TSP) à savoir la moyenne et l'écart type.

Dans ce sens nous avons repris les travaux de Golden [10] qui nous ont permis d'obtenir une estimation de la solution optimale par intervalle de confiance.

Par la suite considérant cette estimation comme étant la solution du problème; nous avons élaboré des procédures de calculs qui nous ont permis de déterminer par l'expérimentation l'équation décrivant la borne inférieure de la solution optimale du problème.

Néanmoins il serait souhaitable de continuer l'étude pour déterminer avec plus de précision les valeurs des constantes p et q , pour déterminer la distribution de la population mère et pourquoi pas pour vérifier ce résultat analytiquement.

BIBLIOGRAPHIE

Bibliographie

- [1] M. BABAS.
" Statistiques, files d'attente et simulation ".
OPU (novembre 1992).
- [2] E. BALAS, T.W. McGUIRE et P. TOTH.
" Statistical analysis of some traveling salesman problem algorithms ".
Management science research report 501 (1984).
- [3] M. BELLMORE et G.L. NEMHANS.
" The traveling salesman problem : a survey ".
Oper.res. vol 16 , pp 538 -557 (1968).
- [4] P. BERTIER.
" Procédures pour élaborer des tournées de distributions ".
METRA N° 4 , (1965).
- [5] L. CAIDI.
" Test d'efficacité d'algorithmes pour le problème du voyageur de commerce ".
PFE, Génie Industriel E.N.P D'ALGER (1995).
- [6] N. CHRISTOFIDES.
" The traveling salesman problem in graph theory an algorithmic approach ".
W , Rheinolt. New York , pp236 -281 (1975).
- [7] C. CLARK et J.W. WRIGHT
« Scheduling of vehicles from central depot to a number of delivery points » .
Oper.Rcs vol 12 , pp 568 - 581 , (1964)
- [8] Y. EDAN, T.FLASH, U.M. PEIPER, I.SHMULEVICH et SARIG.
" Near minimum time task planning for fruit picking robots ".
IEEE vol 7, pp 48 - 56 (1991).
- [9] B. GNEDENKO.
" Sur la distribution limite du terme maximum d'une série aléatoire ".
Ann.Math, 44 3. (1943).
- [10] B.L. GOLDEN.
" A statistical approach to the TSP ".
Networks, 7 pp 209 - 225 (1977b).

- [11] B.L. GOLDEN et F.B. ALT.
 " Interval estimation of a global optimum for large combinatorial problem " .
 Comn.Statist. B- Simulation Comput.7, pp 361-367 (1979).
- [12] M. GONDRAN.
 " Graphes et algorithmes " .
 Eyrolles,Paris (1985).
- [13] M. KENDALL et A.STUART.
 " The advanced theory of statistics " .
 Griffin (1976).
- [14] E.L. LAWLER, J.K.LENSTRAN, A.H.G.RINNOY KAN et D.B.SHMOYS.
 " The traveling salesman problem, a guide tour of combinatoriel optimization " .
 John Wiley et Sons Ltd (1987).
- [15] S. LIN
 « Computer solution to the traveling salesman problem » .
 Bell System Technology Journal , vol 44 , pp 2245 -2269 (1965).
- [16] S. LIN et B. W. KERNINGHAN
 « An effective heuristic algorithm for the traveling salesman problem » .
 Oper . Res , vol 21 pp 498 - 516 (1973).
- [17] J. D. LITTLE
 « An algorithm for the TSP » .
 O.R.S.A. 11 pp 972-989 (1963)
- [18] I. OR
 « Traveling salesman -type combinatorial problems and their relation to the logistic of regional blood banking » .
 Doctoral dissertation , Northwestern University , Evanston , Illinois (1976).
- [19] A.OUABDESSELAM.
 « Etude statistique de la distribution de cycles hamiltoniens »
 Séminaire , Institut Fourier , Grenoble (mai 1962).
- [20] ROSEAU.
 « Exercices et problèmes résolus de recherche opérationnelle » .
 Tome 3 , pp 241- 244 Masson (1985).

- [21] D. ROSENKRANTZ , R.STERNS et LEVIS
 « An analysis of several heuristics for the traveling salesman problem » .
 SIAM J. COMP , vol 6 , pp 563 - 581 (1977)
- [22] B.ROY.
 « Un modèle d'implantation et de cheminement » .
 Doc . SEMA (mai 1966) .
- [23] M.SAKAROVITCH.
 « Optimisation combinatoire , programmation discrète » .
 Hemann,Paris (1984) .
- [24] G.SAPORTA.
 « Probabilités, Analyse des données et Statistiques» .
 Tec.Halip(1990) .
- [25] M.SARI.
 « Investigation des modèles à coûts unitaires variables dans le cas
 d'un système de distribution à dépôt unique ou à plusieurs dépôts » .
 Thèse de Magister en G-Industriel , E.N.P d'Alger (1990) .
- [26] A.SKODARLI et A.TALAH.
 « Sur le problème du voyageur de commerce » .
 P.F.E , G- Industriel ,E.N.P d'Alger (1992) .
- [27] M.R.SPEIGEL.
 « Probabilités et Statistique ,cours et problèmes» .
 Série Schaum. Traduction Française R.JACOUD.
 Mc Graw-hill, Paris (1981) .
- [28] M.M.SYSLO, N.DEO et J.KOWALIK.
 « Discrete Optimisation Algorithms with Pascal programs » .
 Prentice-Hall (1983) .
- [29] P.TASSI.
 « Méthodes statistiques » .
 Economica. (1989)
- [30] L.VILLENEUVE.
 « Introduction to UNFIT » .
 G-Industriel, Ecole Polytechnique de Montréal, Canada (1988) .

ANNEXES

- Annexe 1 : Description de UNIFIT
- Annexe 2 : Description de CURVE
- Annexe 3 : Test de l'hypothèse aléatoire
- Annexe 4 : La variation de K en fonction de la taille de l'échantillon
- Annexe 5 : La variation de W en fonction du nombre de villes
- Annexe 6 : L'erreur relative de la solution calculée
- Annexe 7 : Disquette d'accompagnement

ANNEXE (1) : Description de UNIFIT

UNIFIT est un logiciel en mode interactif qui permet d'adopter échantillon de données qu'on a calculé expérimentalement, a une distribution de probabilité théorique. Ceci aide bien a représenter un cas réel dans un programme de simulation.

Son avantage est qu'il permet de trouver la meilleur représentation qui s'adapte a un cas et ceci en un peu temps .

UNIFIT est capable d'adopter un échantillon a l'une des 13 distribution continues ou des 5 distributions discrètes suivantes :

Lois continues et négatives :

- Expérimental
- Inverse gaussien
- Gamma
- Log normale
- Weibull
- Pearson type 5
- Pearson type 6

Lois continues (univers non bornées) :

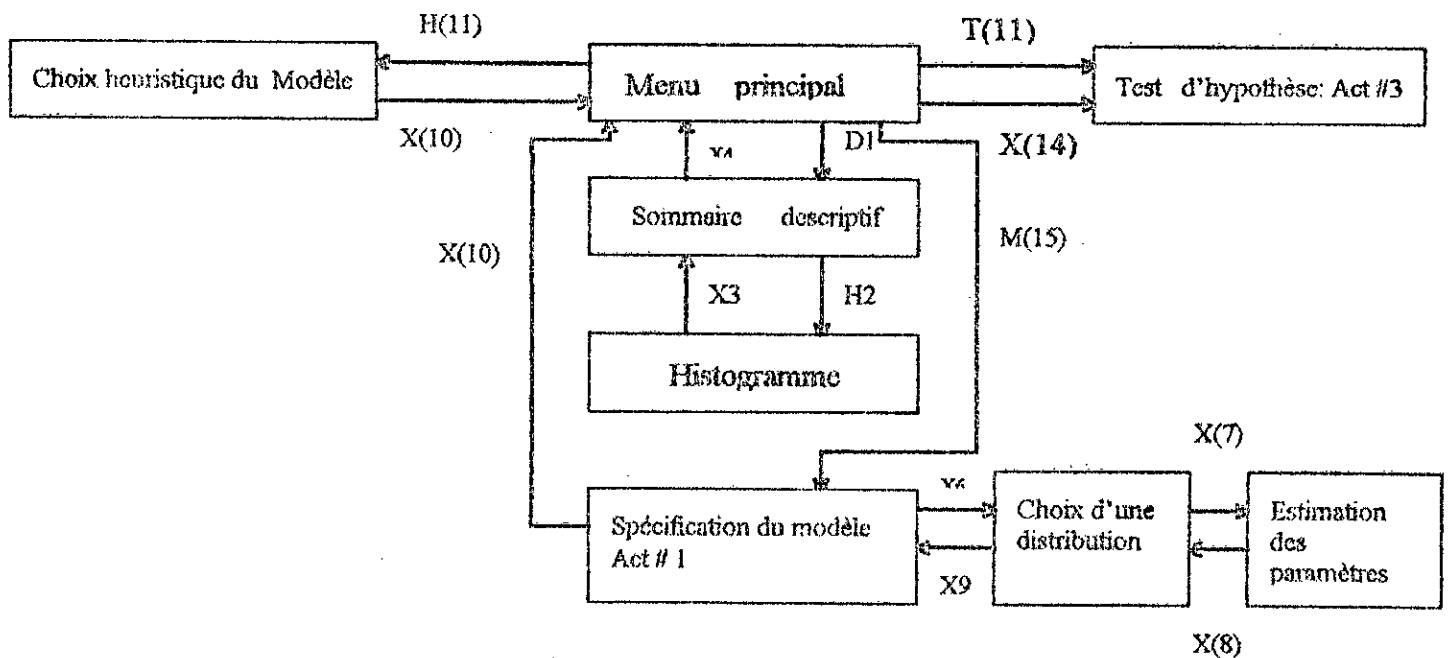
- Extream value type A
- logistic.
- Extream value type B
- normal

Lois continues (univers bornées):

- uniforme
- Beta

Lois discret :

- Binomial
- uniforme
- Géométrique
- Poisson négative.



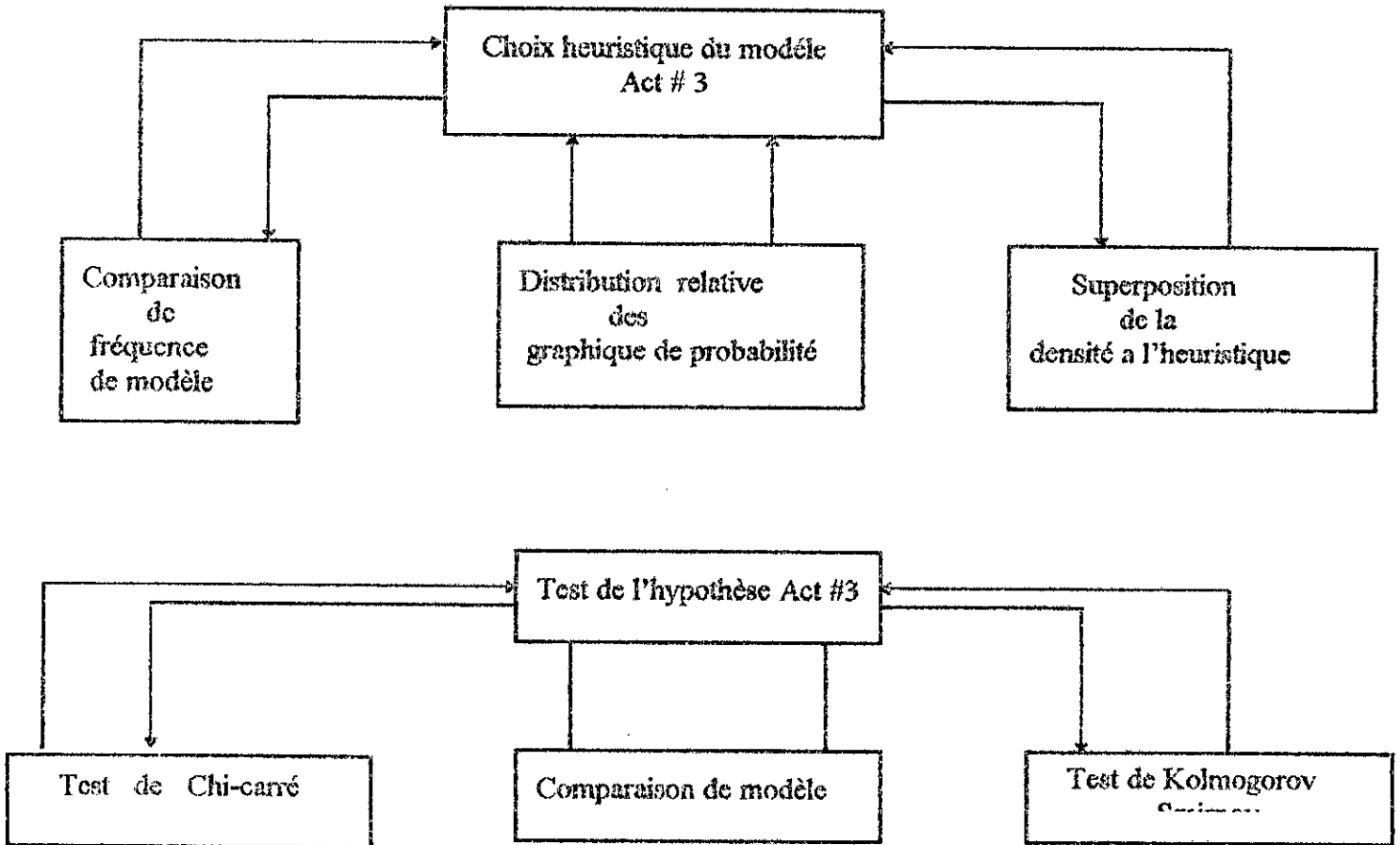


Schéma d'utilisation de UNIFIT

ANNEXE (2) : Description de CURVE

CURVE (sous WINDOWS, version 1.01 ; 1993) est un logiciel dont les possibilités et les performances en font un « superajusteur », tend par la multiplicité des fonction que par la précision de ces commandes et la possibilité d'automatiser les analyses .

En effet CURVE est :

- Un grapheur très puissant, aux performances remarquables
- Un bon tableur.

Quelques possibilités de CURVE

A partir d'un échantillon contenant un ensemble de valeurs d'une variable indépendante et un ensemble correspondant de valeurs d'une variable dépendante, CURVE permet d'exprimer la corrélation (relation) entre ces variables sous forme d'équations mathématiques.

Il offre aussi une liste d'équations de courbes (sélectionnées parmi 8000 équation connues par ce logiciel) qui s'ajustent aux mieux avec les données expérimentales , ordonnées dans l'ordre décroissant du coefficient de corrélation .

L'équation du 1^{er} rang est celle qui définit la meilleure courbe d'ajustement.

L'utilisateur de CURVE peut visualiser la courbe d'ajustement .

Notons aussi que CURVE offre la possibilité de construire un intervalle de confiance pour les constantes figurant dans l'équation de la courbe de l'ajustement au seuil 0.01, 0.05 et 0.10.

ANNEXE (3) : Test de l'hypothèse aléatoire

Série de problèmes n° 2

N	η	γ	z_1	t_1	z_2	t_2
10	5.333	1.206	6	0.553	6	0.553
15	8.667	1.531	7	1.089	7	1.089
20	12.000	1.798	14	1.112	15	1.668
25	15.333	2.030	15	0.164	17	0.821
30	18.667	2.239	18	0.298	16	1.191
35	22.000	2.429	23	0.412	24	0.823
40	25.333	2.606	21	1.663	25	0.128
45	28.667	2.771	25	1.323	27	0.601
50	32.000	2.927	35	1.025	36	1.366
55	35.333	3.075	35	0.108	34	0.434
60	38.667	3.216	43	1.347	40	0.415
65	40.000	3.352	42	0.000	41	0.298
70	45.333	3.482	51	1.628	38	2.106
75	48.667	3.607	44	1.294	53	1.201
80	52.000	3.728	55	0.805	51	0.268
85	55.333	3.846	52	0.867	48	1.907
90	58.667	3.960	61	0.589	51	1.936
95	62.000	4.070	61	0.246	60	0.491
100	65.333	4.178	60	1.277	65	0.080

Série de problèmes n° 3

N	η	γ	z_1	t_1	z_2	t_2
10	5.333	1.206	6	0.553	5	0.276
15	8.667	1.531	9	0.218	7	1.089
20	12.000	1.798	11	0.556	11	0.556
25	15.333	2.030	15	0.164	13	1.149
30	18.667	2.239	18	0.298	15	1.638
35	22.000	2.429	23	0.412	19	1.235
40	25.333	2.606	26	0.256	21	1.663
45	28.667	2.771	31	0.842	30	0.481
50	32.000	2.927	33	0.342	30	0.683
55	35.333	3.075	30	1.734	37	0.542
60	38.667	3.216	37	0.518	43	1.347
65	42.000	3.352	41	0.298	38	1.193
70	45.333	3.482	46	0.191	47	0.479
75	48.667	3.607	50	0.370	48	0.185
80	52.000	3.728	46	1.609	56	1.073
85	55.333	3.846	61	1.474	54	0.347
90	58.667	3.960	53	1.431	56	0.673
95	62.000	4.070	57	1.228	58	0.983
100	65.333	4.178	60	1.277	60	1.277

Série de problèmes n° 4

N	η	γ	z_1	t_1	z_2	t_2
10	5.333	1.206	5	0.276	5	0.276
15	8.667	1.531	9	0.218	10	0.871
20	12.000	1.798	11	0.556	12	0.000
25	15.333	2.030	15	0.164	18	1.313
30	18.667	2.239	18	0.298	17	0.745
35	22.000	2.429	23	0.412	22	0.000
40	25.333	2.606	22	1.279	27	0.640
45	28.667	2.771	29	0.120	30	0.481
50	32.000	2.927	33	0.342	33	0.342
55	35.333	3.075	34	0.434	37	0.542
60	38.667	3.216	38	0.207	36	0.829
65	42.000	3.352	42	0.000	45	0.895
70	45.333	3.482	48	0.766	46	0.191
75	48.667	3.607	54	1.478	45	1.017
80	52.000	3.728	48	1.073	52	0.000
85	55.333	3.846	54	0.347	56	0.173
90	58.667	3.960	65	1.600	55	0.926
95	62.000	4.070	57	1.228	59	0.737
100	65.333	4.178	63	0.558	61	1.037

Série de problèmes n° 5

N	η	γ	z_1	t_1	z_2	t_2
10	5.333	1.206	7	1.381	4	1.105
15	8.667	1.531	8	0.435	8	0.435
20	12.00	1.798	13	0.556	14	1.112
25	15.333	2.030	16	0.328	13	1.149
30	18.667	2.239	16	1.191	19	0.149
35	22.000	2.429	21	0.412	23	0.412
40	25.333	2.606	22	1.276	24	0.512
45	28.667	2.771	24	1.684	27	0.602
50	32.000	2.927	34	0.683	31	0.342
55	35.333	3.075	32	1.084	35	0.108
60	38.667	3.216	41	0.725	41	0.725
65	42.000	3.352	37	1.492	41	0.298
70	45.333	3.482	46	0.191	48	0.766
75	48.667	3.607	48	0.185	52	0.924
80	52.000	3.728	52	0.000	49	0.805
85	55.333	3.846	54	0.347	55	0.087
90	58.333	3.960	61	0.589	61	0.589
95	62.000	4.070	60	0.491	64	0.491
100	65.333	4.178	72	1.596	64	0.319

Série de problèmes n° 6

N	η	γ	z_1	t_1	z_2	t_2
10	5.333	1.206	06.000	0.553	5.000	0.276
15	8.667	1.531	11.000	1.524	8.000	0.435
20	12.000	1.798	13.000	0.556	14.000	1.112
25	15.333	2.030	12.000	1.642	16.000	0.328
30	18.667	2.239	17.000	0.745	20.000	0.596
35	22.000	2.429	23.000	0.412	20.000	0.823
40	25.333	2.606	27.000	0.640	22.000	1.279
45	28.667	2.771	26.000	0.962	31.000	0.842
50	32.000	2.927	33.000	0.342	34.000	0.683
55	35.333	3.075	40.000	1.518	34.000	0.434
60	38.667	3.216	35.000	1.140	36.000	0.829
65	42.000	3.352	40.000	0.597	39.000	0.895
70	45.333	3.482	42.000	0.957	40.000	1.532
75	48.667	3.607	53.000	1.201	45.000	1.017
80	52.000	3.728	51.000	0.268	52.0000	0.000
85	55.333	3.846	55.000	0.087	59.0000	0.953
90	58.667	3.96	54.000	1.179	59.0000	0.084
95	62.000	4.07	56.000	1.474	64.0000	0.491
100	65.333	4.178	61.000	1.037	73.0000	1.835

Série de problèmes n° 7

N	η	γ	z_1	t_1	z_2	t_2
10	5.33	1.206	07.000	1.381	07.000	1.381
15	8.667	1.531	08.000	0.435	07.000	1.089
20	12.000	1.798	13.000	0.556	15.000	1.668
25	15.333	2.030	15.000	0.164	12.000	1.642
30	18.667	2.239	18.000	0.298	20.000	0.596
35	22.000	2.429	20.000	0.823	22.000	0.000
40	25.333	2.606	26.000	0.256	21.000	1.663
45	28.667	2.771	30.000	0.481	24.000	1.684
50	32.000	2.927	31.000	0.342	37.000	1.708
55	35.333	3.075	37.000	0.542	32.000	1.084
60	38.667	3.216	34.000	1.451	36.000	0.829
65	42.000	3.352	46.000	1.193	38.000	1.193
70	45.333	3.482	48.000	0.766	50.000	1.340
75	48.667	3.607	47.000	0.462	51.000	0.647
80	52.000	3.728	53.000	0.268	52.000	0.000
85	55.333	3.846	51.000	1.127	56.000	0.173
90	58.667	3.960	57.000	0.421	60.000	0.337
95	62.000	4.070	57.000	1.228	64.000	0.491
100	65.333	4.178	67.000	0.399	59.000	1.516

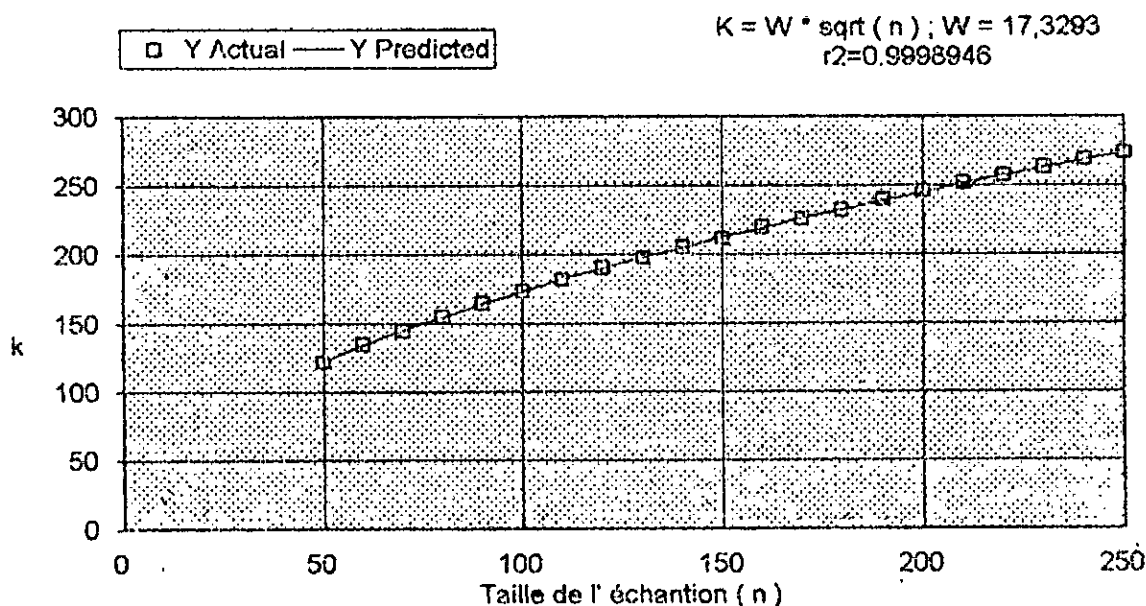
Série de problèmes n° 8

N	η	γ	z_1	t_1	z_2	t_2
10	5.333	1.206	07	1.381	6	0.553
15	8.667	1.531	07	1.089	07	1.089
20	12.000	1.798	12	0.000	13	0.556
25	15.333	2.030	15	0.164	14	0.657
30	18.667	2.239	22	1.489	19	0.149
35	22.000	2.429	20	0.823	23	0.412
40	25.333	2.606	25	0.128	21	0.663
45	28.667	2.771	28	0.241	26	0.962
50	32.000	2.927	33	0.342	33	0.342
55	35.333	3.075	35	0.108	35	0.108
60	38.667	3.216	38	0.207	43	1.347
65	42.000	3.352	40	0.597	40	0.597
70	45.333	3.482	42	0.957	48	0.766
75	48.667	3.607	48	0.185	51	0.647
80	52.000	3.728	48	1.073	46	1.607
85	55.333	3.846	55	0.087	53	0.607
90	58.667	3.960	59	0.084	60	0.337
95	62.000	4.070	63	0.246	67	1.228
100	65.333	4.178	58	1.755	69	0.878

ANNEXE (4) : La variation de k en fonction de la taille de l'échantillon

Problème de taille N = 65 villes (m = 2871 ; $\sigma = 130.269$)

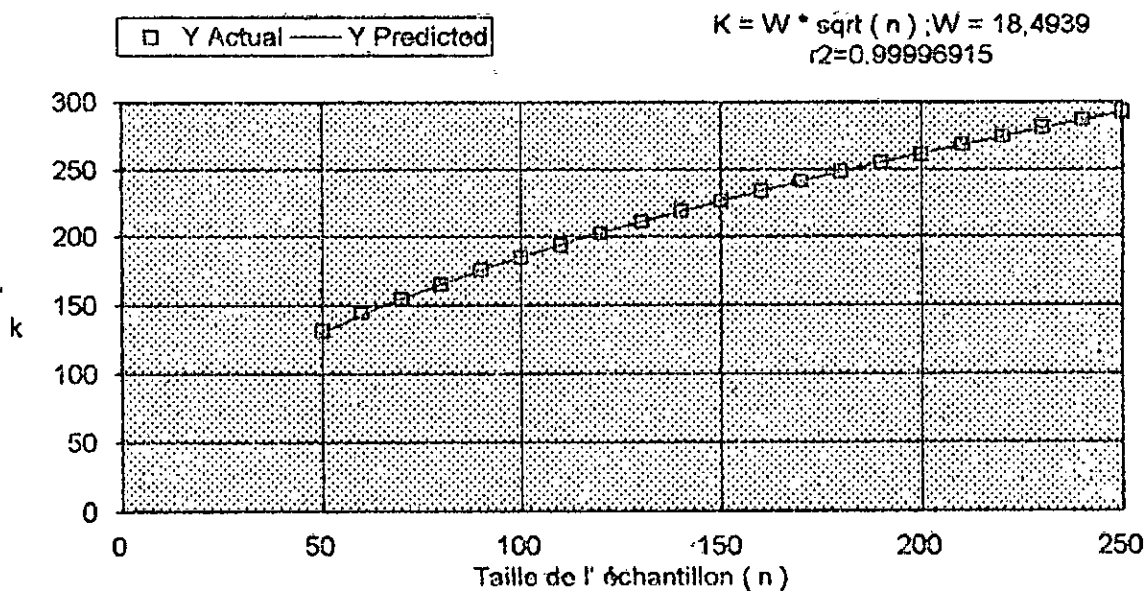
Taille de l'échantillon n	(μ) solution heuristique	$\hat{\mu}$	$\hat{\sigma}$	λ	K-S	K	W
50	626.325	626.300	25.541	600.783	0.973	121.876	17.236
60	619.024	618.822	35.111	583.912	0.737	133.953	17.293
70	621.771	621.764	28.750	593.020	0.767	144.499	17.271
80	619.715	619.475	34.476	585.238	0.654	154.630	17.288
90	617.607	616.662	36.090	581.517	0.582	164.215	17.310
100	619.840	619.824	33.925	585.915	0.562	172.855	17.285
110	612.487	612.059	43.613	568.873	0.886	181.917	17.345
120	613.413	613.404	40.292	570.121	0.852	189.893	17.335
130	617.325	617.277	40.036	577.288	0.565	197.717	17.341
140	615.777	615.717	37.023	578.754	0.650	204.898	17.317
150	620.450	620.419	31.562	588.887	0.398	211.648	17.281
160	611.718	610.837	43.971	567.746	0.436	219.519	17.355
170	619.024	618.956	33.760	585.264	0.578	225.463	17.292
180	613.021	619.924	40.197	572.823	0.734	231.894	17.285
190	607.717	606.534	48.721	558.996	0.479	239.664	17.387
200	607.717	607.526	49.223	558.493	0.787	245.789	17.380
210	609.661	609.271	44.767	564.894	0.636	251.657	17.366
220	613.387	613.384	38.643	574.744	0.680	257.119	17.335
230	613.570	613.546	37.677	575.893	0.942	262.879	17.334
240	607.717	607.445	45.659	562.057	0.992	269.258	17.381
250	618.023	618.022	36.625	581.398	0.993	273.527	17.299



La variation de K en fonction de la taille de l'échantion
(problème de taille N = 65 villes)

Problème de taille N = 70 villes (m = 3716 ; $\sigma = 162.714$)

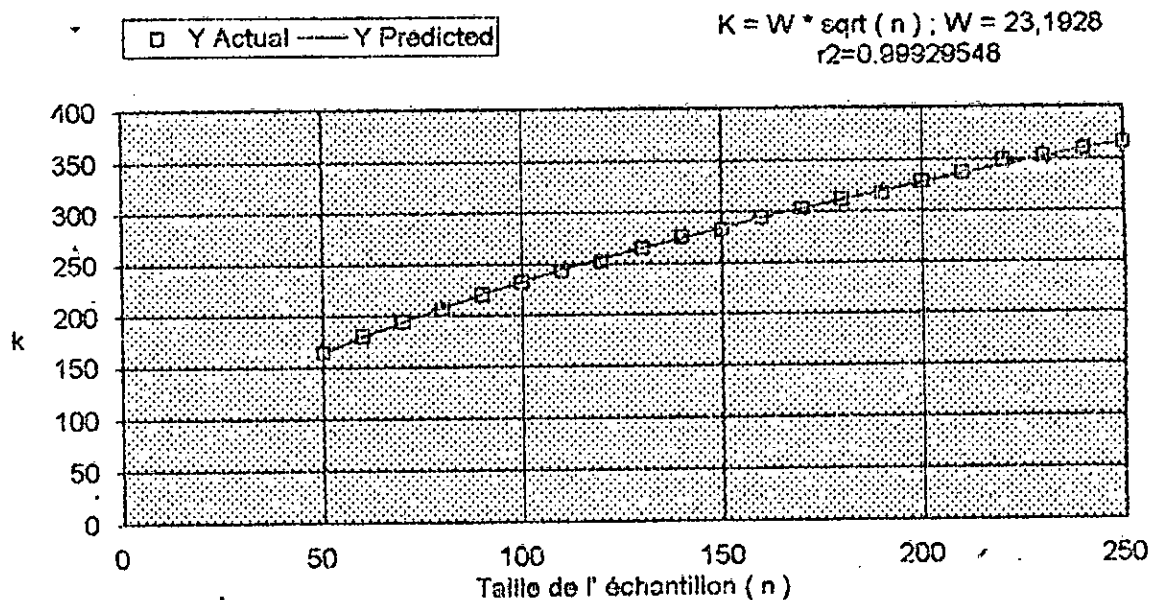
Taille de l'échantillon n	(μ) solution heuristique	\hat{a}	\hat{b}	λ	K-S	K	W
50	703.460	701.950	55.523	647.937	0.938	130.985	18.524
60	701.943	700.083	58.293	643.649	0.563	143.576	18.535
70	713.247	713.220	46.461	666.786	1.105	154.404	18.455
80	708.346	708.303	47.304	661.042	0.753	165.335	18.485
90	711.924	711.906	44.895	667.029	1.106	175.155	18.463
100	712.479	712.471	44.969	667.509	1.023	184.594	18.459
110	715.008	714.983	46.901	668.107	1.042	193.442	18.444
120	711.559	711.342	49.451	662.107	0.708	202.289	18.466
130	705.963	705.950	51.865	654.098	1.053	210.927	18.499
140	704.138	704.121	57.523	646.615	0.908	219.022	18.510
150	707.493	707.480	52.048	655.445	1.058	226.457	18.490
160	706.782	706.767	55.883	650.898	1.220	233.939	18.494
170	705.521	705.362	53.936	651.585	0.665	241.251	18.503
180	705.551	705.436	57.753	647.898	0.689	248.239	18.502
190	709.188	709.072	52.409	656.778	0.541	254.734	18.480
200	711.559	711.342	49.908	661.651	0.600	261.154	18.466
210	703.567	703.408	56.272	647.295	0.787	268.310	18.515
220	708.327	708.304	63.029	655.298	1.312	274.178	17.485
230	703.460	703.315	55.356	648.104	0.967	280.805	18.515
240	704.193	704.193	54.028	650.164	0.765	286.761	18.510
250	703.481	703.363	57.553	645.928	1.050	292.754	18.515



La variation de K en fonction de la taille de l'échantillon (problème de taille N = 70 villes)

Problème de taille N = 100 villes ($m = 5356.492$; $\sigma = 199.212$)

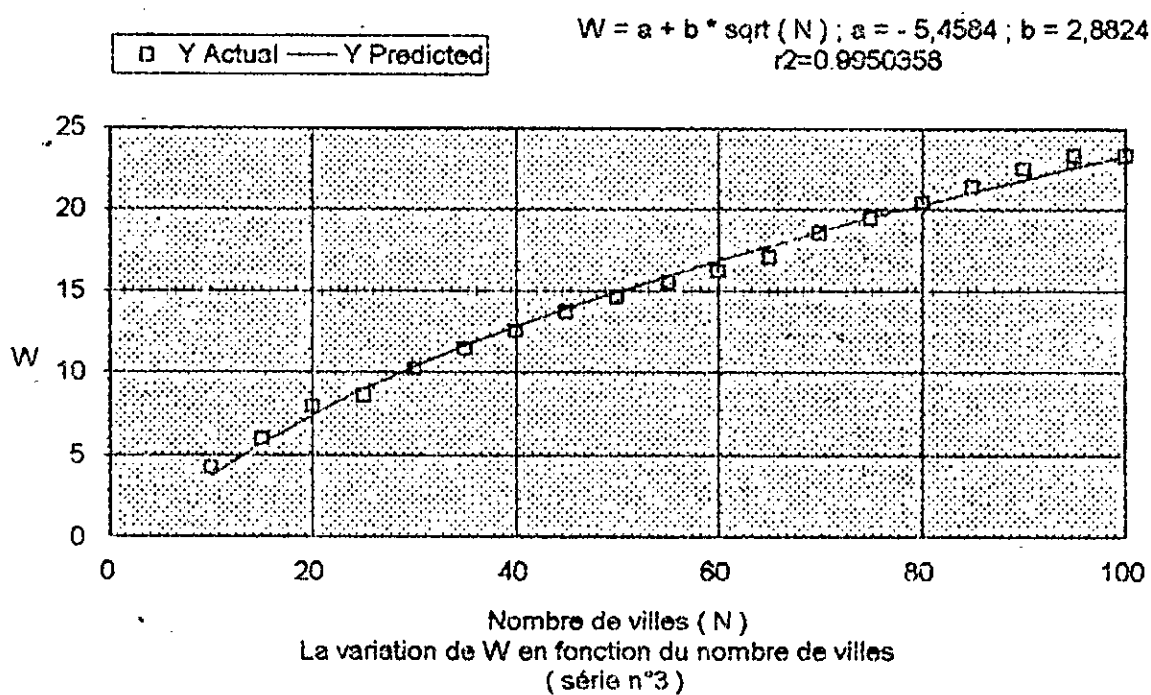
Taille de l'échantillon (n)	solution heuristique μ	$\hat{\alpha}$	$\hat{\beta}$	λ	K-S	K	W
50	843.951	834.729	100.936	743.015	0.738	160.501	23.158
60	833.534	831.988	100.075	733.459	0.936	175.926	23.206
70	832.192	829.165	96.974	735.218	0.622	190.141	23.197
80	851.518	850.581	72.981	778.538	0.722	202.307	23.980
90	836.720	836.033	90.880	745.840	0.937	215.272	23.144
100	838.355	835.914	92.450	745.904	0.932	226.923	23.144
110	844.762	844.695	82.517	762.244	0.891	237.536	23.062
120	861.073	861.072	86.586	774.487	0.941	247.198	23.000
130	833.534	833.112	101.998	731.536	1.052	258.888	23.216
140	832.192	826.277	106.881	725.311	0.421	269.071	23.247
150	845.321	845.293	81.471	763.842	0.897	277.346	23.054
160	833.534	832.796	109.902	730.632	1.034	287.235	23.220
170	840.274	839.966	84.765	755.508	0.843	295.606	23.096
180	833.534	833.145	100.625	732.909	0.950	304.636	23.209
190	848.556	848.554	78.469	770.086	1.119	311.917	23.022
200	830.889	828.769	103.522	727.367	1.127	321.425	23.236
210	830.889	828.543	103.344	727.545	0.869	329.379	23.236
220	847.275	847.137	78.141	769.134	1.099	335.745	23.027
230	830.889	829.767	103.269	727.620	1.077	344.614	23.236
240	830.570	826.731	105.827	724.743	0.963	352.262	23.250
250	836.720	836.705	90.971	745.749	1.158	368.734	23.145



La variation de K en fonction de la taille de l'échantillon (problème de taille N = 100 villes)

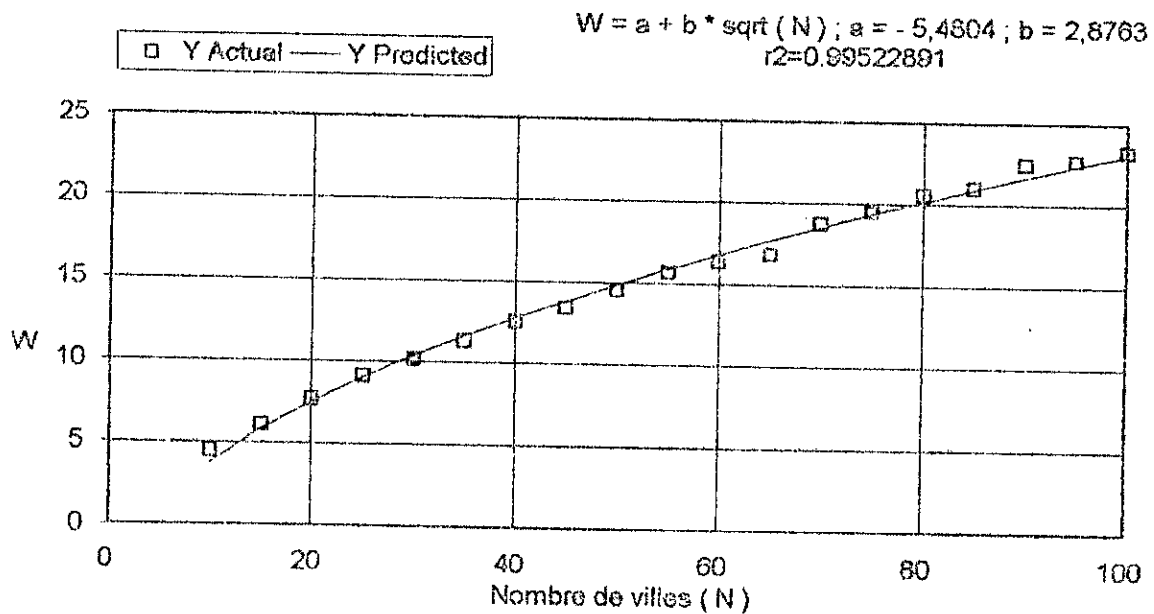
Série de problèmes n° 3

N	m	σ	\hat{a}	k-s	W
10	540.295	62.782	+	+	3.770
15	784.663	77.761	322.434	1.299	5.944
20	976.754	91.583	279.842	1.243	7.609
25	1313.63	107.174	390.734	1.1087	8.611
30	1550.243	103.221	498.152	0.854	10.192
35	1858.044	117.808	510.07	0.916	11.442
40	2036.543	124.650	473.402	0.722	12.54
45	2196.884	122.981	509.065	0.657	13.724
50	2526.447	138.370	506.374	0.868	14.599
55	2905.83	147.292	613.819	0.540	15.561
60	3326.982	162.687	629.831	0.972	16.578
65	3574.358	167.013	612.283	0.432	17.735
70	3687.019	186.845	664.74	1.040	18.581
75	4070.203	175.930	702.724	0.573	19.473
80	4273.627	174.221	714.634	0.760	20.428
85	4449.703	172.464	751.993	1.064	21.44
90	4876.169	186.845	760.007	0.949	22.020
95	4702.806	174.130	819.065	0.600	22.781
100	4932.498	164.604	857.587	0.753	23.357



Série de problèmes n° 5

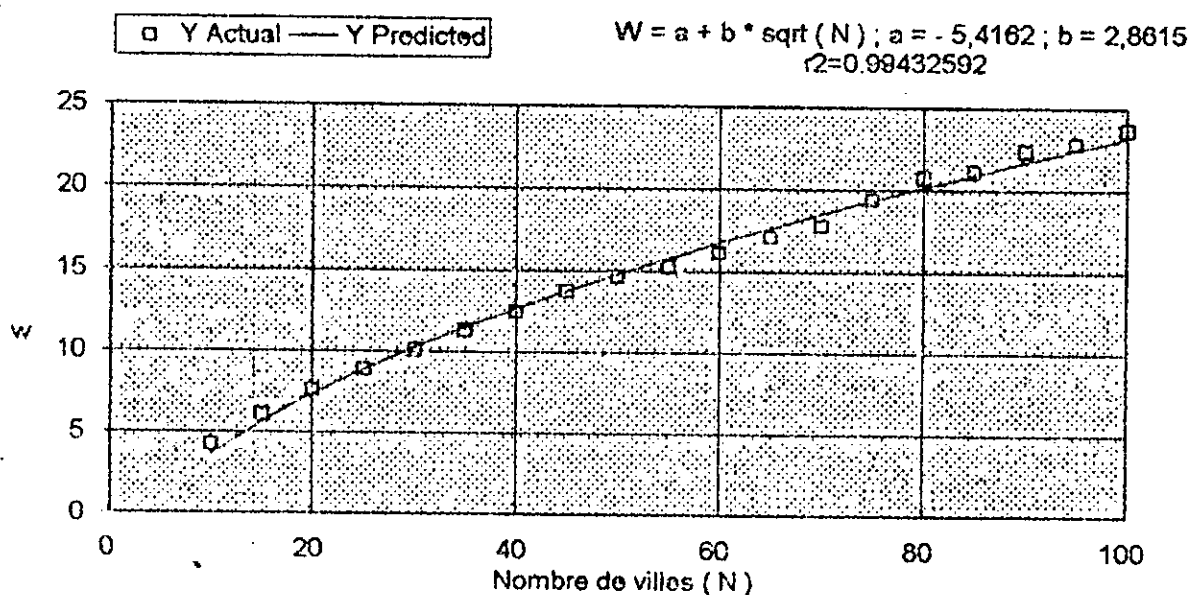
N	m	σ	\hat{a}	k-s	W
10	455.733	54.462	+	+	3.999
15	928.226	92.313	381.414	1.006	5.923
20	1024.477	80.293	420.534	0.925	7.746
25	1477.737	115.207	441.228	1.055	8.997
30	1484.3	105.467	418.720	1.049	10.104
35	2031.358	138.089	464.245	0.913	11.348
40	2107.218	127.220	519.602	1.097	12.479
45	2354.544	133.725	532.228	0.809	13.627
50	2722.305	149.999	547.182	0.839	14.501
55	2757.984	138.238	584.847	0.644	15.720
60	3273.331	163.445	631.859	1.149	16.773
65	3254.338	150.175	636.687	0.525	17.431
70	3711.319	165.338	680.600	0.971	18.330
75	4020.867	173.138	696.086	0.629	19.203
80	4321.885	177.726	750.301	0.452	20.096
85	4312.652	171.085	713.901	0.717	21.035
90	4457.280	163.215	761.085	0.721	22.033
95	4926.273	180.598	781.009	0.586	22.953
100	5235.567	185.331	832.506	0.523	23.218



La variation de W en fonction du nombre de villes
(série n° 5)

Série de problèmes n° 6

N	m	σ	$\hat{\mu}$	K-S	W
10	455.733	54.462	+	+	4.228
15	908.226	921.316	+	+	5.839
20	1051.169	86.202	392.735	0.843	7.638
25	1189.909	84.168	446.622	1.387	8.831
30	1343.02	91.226	422.363	0.917	10.092
35	1850.565	122.495	467.605	1.079	11.290
40	2171.921	133.564	514.18	0.605	12.411
45	2093.052	114.595	526.193	0.488	13.673
50	2649.517	143.400	552.661	1.173	14.622
55	2958.741	152.392	598.079	0.656	15.490
60	3316.385	158.669	663.007	0.688	16.722
65	3673.468	169.629	673.83	0.552	17.094
70	3884.534	175.324	665.843	0.752	18.358
75	3834.107	162.366	678.974	0.995	19.432
80	4081.638	164.81	763.242	0.667	20.134
85	4313.007	168.294	759.992	0.880	21.112
90	4672.648	181.843	695.077	0.460	21.873
95	5017.438	185.053	788.234	0.690	22.854
100	5276.153	189.082	802.919	0.850	23.657



La variation de W en fonction du nombre de villes
(série n° 6)

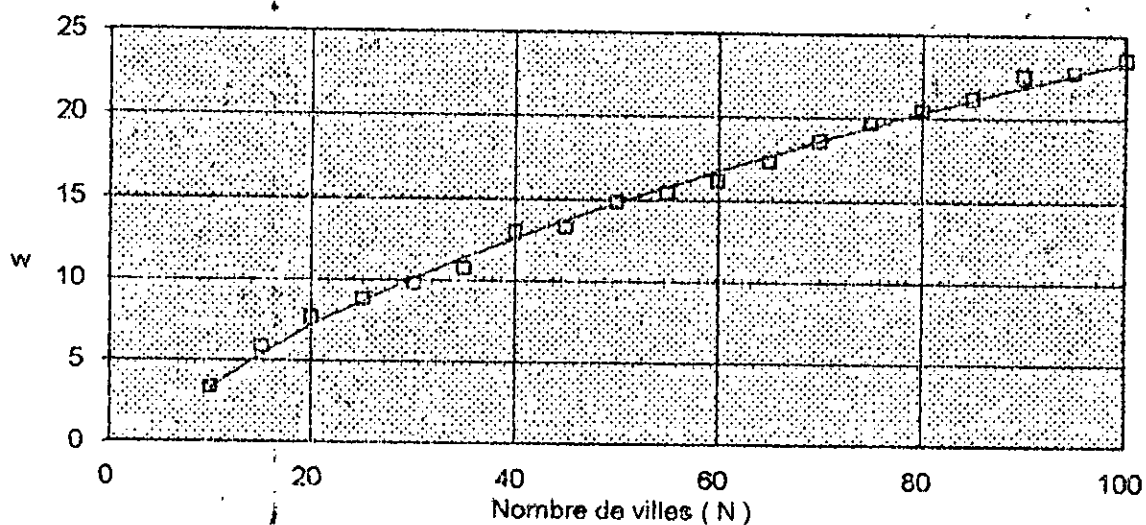
Série de problèmes n° 7

N	m	σ	\hat{a}	k-s	W'
10	433.877	61.393	+	+	3.396
15	711.039	68.865	306.664	1.883	5.872
20	1084.039	90.716	388.229	1.330	7.670
25	1327.516	102.932	425.526	1.030	8.763
30	1699.572	129.800	449.615	1.287	9.629
35	1978.620	133.517	477.413	1.591	11.243
40	1860.247	105.817	487.644	0.903	12.971
45	2221.423	124.277	526.011	0.084	13.642
50	2588.953	137.072	556.216	1.172	14.830
55	3109.025	159.681	644.189	0.788	15.436
60	3183.914	151.755	659.586	0.728	16.634
65	3287.129	152.331	644.033	0.552	17.351
70	3685.605	163.268	645.226	0.686	18.622
75	3743.301	156.698	661.122	0.775	19.669
80	4305.699	176.211	705.047	0.713	20.433
85	4584.746	182.849	708.161	0.582	21.201
90	4434.140	167.020	769.168	0.590	21.943
95	4843.119	176.265	824.110	1.032	22.801
100	5398.187	194.743	817.064	1.028	23.524

□ Y Actual — Y Predicted

$$W = a + b \cdot \sqrt{N} ; a = -5,9928 ; b = 2,9400$$

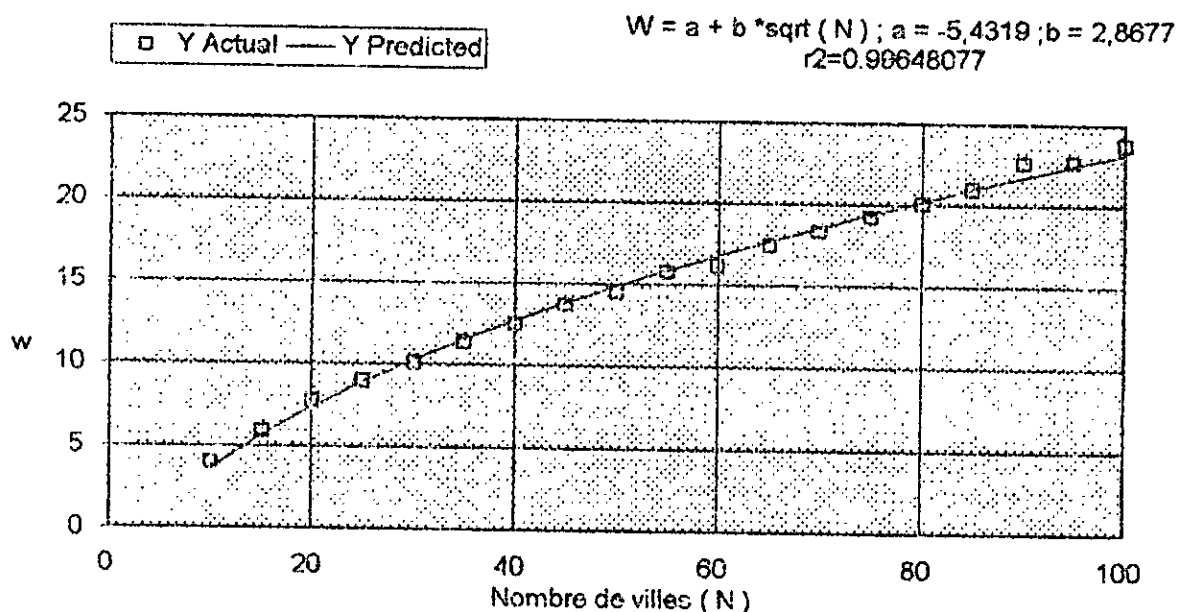
$$r^2 = 0,99607196$$



La variation de W en fonction du nombre de villes
(série n° 7)

Série de problèmes n° 8

N	m	σ	\hat{a}	k-s	W
10	542.726	53.431	338.236	1.880	3.827
15	858.005	74.069	424.273	0.944	5.856
20	1121.835	95.526	396.437	1.309	7.594
25	1386.766	109.576	473.926	1.265	8.841
30	1538.897	105.316	473.926	1.079	10.112
35	1829.621	114.445	526.985	1.141	11.382
40	2187.281	131.294	524.833	0.463	12.662
45	2504.956	145.410	545.028	0.937	13.478
50	2641.800	139.579	567.919	0.788	14.858
55	2853.035	143.493	577.935	0.932	15.552
60	3096.870	151.440	594.351	0.817	16.525
65	3496.060	165.650	567.804	1.292	17.677
70	3701.326	163.267	659.798	0.549	18.629
75	3877.288	164.252	668.047	0.481	19.538
80	4111.563	167.811	676.776	0.707	20.468
85	4168.366	163.843	744.368	0.935	20.898
90	4795.255	179.570	736.392	0.884	22.046
95	4788.388	175.115	763.718	0.805	22.983
100	4817.509	172.168	784.637	0.445	23.424



La variation de W en fonction du nombre de villes
 (série n° 8)

ANNEXE (6) : Erreur relative de la solution calculée

Série de problèmes n° 2

N	μ	\hat{a}	\hat{b}	λ	S-C	D-A
10	280.917	+	+	+	287.716	2.420
15	322.091	303.4	28.118	293.118	323.487	6.620
20	431.016	430.941	19.274	411.741	454.843	5.546
25	440.773	440.772	9.114	431.658	449.834	2.056
30	476.2	476.127	19.737	456.462	481.006	1.024
35	476.756	475.878	26.128	450.628	461.782	2.962
40	519.661	519.273	30.404	489.257	499.080	3.888
45	547.43	547.428	22.592	524.838	510.521	9.995
50	564.174	564.166	20.579	543.595	514.616	8.780
55	636.216	634.847	34.176	602.040	617.467	2.737
60	635.335	634.789	40.260	555.075	574.545	9.490
65	662.962	662.914	33.404	629.558	592.806	10.575
70	630.610	628.749	49.002	581.608	691.361	9.958
75	644.338	643.043	52.706	591.631	707.170	9.972
80	729.864	729.841	35.193	694.671	771.687	5.733
85	667.95	664.717	55.615	612.335	625.792	5.856
90	737.417	736.007	76.796	660.621	779.759	5.945
95	803.375	801.922	55.608	747.767	856.674	6.827
100	815.514	814.735	48.207	767.226	783.145	3.877

Série de problèmes n° 3

N	μ	\hat{a}	\hat{b}	λ	S-C	D-A
10	303.602	+	+	+	315.592	3.949
15	332.152	322.434	18.464	313.688	347.015	7.624
20	331.553	279.842	60.992	270.56	303.124	8.319
25	392.256	390.734	24.228	368.027	362.214	7.299
30	498.269	498.152	17.133	481.135	491.903	1.254
35	510.078	510.07	17.297	492.781	501.089	1.762
40	473.405	473.402	13.922	459.482	453.987	4.101
45	509.14	509.065	27.585	481.555	499.494	1.879
50	506.667	505.374	21.002	485.665	471.993	6.805
55	613.866	613.819	16.176	597.69	572.253	6.771
60	629.926	629.831	37.413	592.512	534.829	5.547
65	613.187	612.283	30.29	582.896	617.261	0.813
70	666.66	664.74	45.689	629.047	664.387	0.053
75	702.724	702.724	34.218	668.506	651.916	7.230
80	714.728	714.634	41.547	673.181	745.889	4.373
85	752.038	751.993	48.644	703.389	820.673	9.133
90	760.017	760.007	48.000	712.017	800.549	5.335
95	819.134	819.065	33.272	785.862	845.418	3.217
100	862.296	857.587	55.709	824.29	868.876	1.316

Série de problèmes n° 4

N	μ	$\hat{\delta}$	λ	\hat{b}	S-C	D-A
10	253.678	+	+	+	259.073	3.579
15	285.503	+	+	+	312.559	9.476
20	349.281	338.281	325.735	23.546	461.356	6.821
25	462.022	462.006	448.395	13.627	483.794	4.716
30	468.833	466.755	436.050	32.782	450.720	3.435
35	488.047	488.025	460.176	27.875	469.295	3.837
40	495.856	495.855	466.978	28.877	475.049	4.203
45	510.789	509.963	481.170	29.618	498.243	2.296
50	594.917	594.897	570.315	24.598	548.316	7.828
55	600.593	600.556	576.176	24.416	574.949	4.264
60	587.833	586.033	555.920	31.913	542.505	6.814
65	675.192	673.712	626.456	48.736	623.084	7.515
70	661.232	660.823	623.342	37.89	680.772	3.018
75	701.685	701.607	671.721	29.964	705.345	0.532
80	705.171	704.785	672.343	32.828	755.222	7.156
85	747.177	746.838	713.691	33.486	740.294	0.876
90	741.198	741.181	701.842	39.356	745.834	0.627
95	780.075	779.882	742.822	37.252	805.192	3.245
100	852.262	851.725	773.680	78.582	849.548	0.255

Série de problèmes n° 5

N	μ	$\hat{\delta}$	\hat{b}	λ	S-C	D-A
10	238.275	+	+	+	260.808	9.456
15	381.640	381.414	4.931	376.709	408.678	7.148
20	420.806	420.534	11.689	409.117	433.884	3.174
25	441.421	441.228	6.550	434.871	455.010	3.123
30	421.173	418.720	21.481	399.692	402.932	3.770
35	464.644	464.245	24.993	439.651	440.799	5.050
40	519.644	519.602	31.389	488.255	492.034	5.306
45	533.553	532.228	33.709	499.844	508.865	4.388
50	547.249	547.182	37.015	510.234	495.091	9.517
55	585.455	584.847	32.872	552.583	567.852	2.906
60	632.208	631.859	46.024	586.184	628.448	0.539
65	636.718	636.687	25.045	611.674	595.371	6.489
70	680.609	680.600	25.514	655.155	638.808	6.140
75	695.213	696.086	37.239	657.974	656.828	5.504
80	750.366	750.301	27.357	723.009	723.176	3.615
85	718.722	713.901	53.437	665.285	712.639	0.176
90	716.365	761.085	43.427	717.938	797.097	4.731
95	781.218	781.009	45.939	735.279	851.565	9.034
100	833.446	832.506	64.885	768.561	818.782	1.648

Série de problèmes n° 6

N	μ	$\hat{\delta}$	$\hat{\beta}$	λ	S-C	D-A
10	225.437	+	+	+	242.912	7.751
15	369.12	+	+	+	908.226	5.294
20	392.735	392.735	13.903	378.832	417.113	6.207
25	446.769	446.622	14.693	432.076	442.724	0.872
30	422.472	422.363	15.891	406.581	407.667	3.479
35	467.616	467.605	20.094	447.522	439.624	5.984
40	514.55	514.18	25.757	488.793	476.193	7.454
45	526.397	526.193	25.124	501.273	511.406	2.809
50	553.031	552.661	51.226	501.805	520.287	5.857
55	598.095	598.079	31.766	486.32	544.364	8.981
60	663.951	663.007	39.434	624.517	651.710	1.703
65	674.202	673.83	38.1156	636.086	670.053	0.560
70	666.09	665.843	41.101	624.989	626.451	5.916
75	679.011	678.974	36.130	642.881	679.366	0.057
80	763.740	763.242	36.948	726.792	744.460	2.460
85	760.076	759.992	29.792	730.284	771.723	1.543
90	696.036	695.077	55.506	640.53	706.136	1.591
95	788.582	788.234	41.103	747.479	842.215	6.848
100	803.177	802.919	82.838	720.339	871.998	8.603

Série de problèmes n° 7

N	μ	$\hat{\delta}$	$\hat{\beta}$	λ	S-C	D-A
10	225.388	+	+	+	214.145	4.988
15	318.066	306.664	23.869	294.194	323.459	5.476
20	389.361	388.229	9.871	379.490	416.780	7.354
25	426.602	425.526	22.262	404.340	413.758	2.765
30	449.929	449.615	15.501	434.428	423.762	5.750
35	477.413	477.413	15.956	431.457	440.723	7.685
40	487.655	487.644	25.994	461.661	516.795	5.977
45	526.030	526.011	35.110	490.920	556.145	5.730
50	556.233	556.216	33.501	522.732	553.682	0.453
55	644.202	644.189	29.381	614.821	579.167	10.009
60	659.805	659.586	34.732	625.073	625.352	5.190
65	644.107	644.033	39.631	604.476	589.988	8.391
70	645.305	645.226	39.695	605.610	651.562	0.981
75	661.922	661.122	45.757	616.165	698.688	5.682
80	705.157	705.047	52.819	652.337	737.667	4.626
85	708.473	708.161	36.385	672.086	737.192	4.099
90	770.411	769.168	51.757	718.653	790.959	2.833
95	824.123	824.110	35.664	798.569	866.174	5.104
100	817.146	817.064	81.064	736.082	862.175	5.521

Série de problèmes n° 8

N	μ	\hat{a}	\hat{b}	λ	S-C	D-A
10	338.416	338.236	3.791	334.618	351.489	3.863
15	425.241	424.273	12.183	413.058	441.136	3.974
20	396.959	396.437	6.812	390.147	419.244	5.753
25	420.241	473.926	14.294	405.947	414.237	0.949
30	474.011	473.926	17.705	456306	459.077	3.133
35	526.986	526.985	24.179	502.807	511.402	2.957
40	524.838	524.833	14.946	509.892	520.373	0.850
45	545.111	545.028	34.326	510.785	498.000	8.627
50	567.996	567.919	26.705	541.201	569.305	0.247
55	568.380	578.430	32.453	546.08 0	553.598	6.856
60	594.358	594.351	38.700	555.658	579.646	0.296
65	567.808	567.804	46.340	521.468	563.096	0.829
70	660.763	659.798	41.417	619.346	667.301	1.137
75	668.551	668.047	35.609	632.942	685.902	2.672
80	679.791	676.776	60.590	619.241	713.619	5.444
85	744.387	744.368	44.709	699.678	720.741	3.173
90	736.439	736.392	63.015	673.424	778.323	5.694
95	766.719	763.718	62.385	704.334	837.389	9.646
100	789.643	784.637	97.977	691.666	807.320	2.891

(S-C) : Solution calculée.

(D-A) : Déviation approximative = $[(s-c)-\hat{a}] * 100/\hat{a}$ en valeur absolue.

ANNEXE : (7)
disquette d'accompagnement

- Pour ne pas perdre de temps à élaborer et rentrer les programmes.
- Pour démarrer directement une application en tapant son nom.
- Pour vérifier les résultats.

Le lecteur intéressé retrouve sur disquette PC les programmes de calculs ainsi que les coordonnées correspondantes aux sommets (villes) de chacune des 152 problèmes traités dans ce mémoire .

Contenu de la disquette : tous les fichiers ".pas" et ".dat" correspondant aux programmes et aux coordonnées respectivement. La disquette contient deux répertoires:

1- Programmes \ contient les fichiers suivants:

- Coordonn.pas
- Test-alé.pas
- Moyenne.pas
- 2-opt.pas

2- Coordonnées \ contient les sous répertoires suivants:

-Série "i" \ contient les fichiers suivants :

- cord "N" .dat

Telle que $i = 1, 2, \dots, 8$.

$N = 10, 15, \dots, 100$. avec un pas de "5" .