

M0004/94B

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTRE DELEGUE AUX UNIVERSITES  
ET A LA RECHERCHE SCIENTIFIQUE

ECOLE NATIONALE POLYTECHNIQUE

**Mémoire**

المدرسة الوطنية المتعددة التقنيات  
BIBLIOTHEQUE — المكتبة  
Ecole Nationale Polytechnique

POUR L'OBTENTION DU DIPLOME DE  
MAGISTER  
EN GENIE-INDUSTRIEL

THEME

OPTIMISATION STOCHASTIQUE DE PROBLEMES NON  
LINEAIRES SANS CONTRAINTES ET A FONCTIONS  
DIFFERENTIABLES

Par

*M<sup>me</sup> HANYA KHERCHI*  
*Ingénieur d'Etat en Génie-Industriel*

Présenté devant le Jury:

Dr	O. BELMOKHTAR	Président
Dr	H. KERBACHE	Examinateur
Dr	A. OUABDESSALEM	Examinateur
Mr	T. LAMRAOUI	Examinateur
Melle	N. ABOUN	Rapporteur
Mr	M. SARI	Invité

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE**

**MINISTRE DELEGUE AUX UNIVERSITES  
ET A LA RECHERCHE SCIENTIFIQUE**

**ECOLE NATIONALE POLYTECHNIQUE**

**Mémoire**

المدرسة الوطنية المتعددة التقنيات  
المكتبة — BIBLIOTHEQUE  
Ecole Nationale Polytechnique

**POUR L'OBTENTION DU DIPLOME DE  
MAGISTER  
EN GENIE-INDUSTRIEL**

**THEME**

**OPTIMISATION STOCHASTIQUE DE PROBLEMES NON  
LINEAIRES SANS CONTRAINTES ET A FONCTIONS  
DIFFERENTIABLES**

**Par**

**M<sup>me</sup> HANYA KHERCHI**  
*Ingénieur d'Etat en Génie-Industriel*

**Présenté devant le Jury:**

<i>Dr</i> O. BELMOKHTAR	<i>Président</i>
<i>Dr</i> H. KERBACHE	<i>Examineur</i>
<i>Dr</i> A. OUABDESSALEM	<i>Examineur</i>
<i>Mr</i> T. LAMRAOUI	<i>Examineur</i>
<i>Melle</i> N. ABOUN	<i>Rapporteur</i>
<i>Mr</i> M. SARI	<i>Invité</i>

## REMERCIEMENTS

Au terme de ce modeste travail, je tiens à adresser mes plus grands remerciements.

- A ma Directrice de Thèse, M<sup>elle</sup> N. ABOUN, pour son aide et ses précieux conseils.

- A la présidente du Jury D<sup>r</sup> O. BELMOKHTAR, pour son soutien moral et son amitié qui m'ont profondément touchée.

- Au D<sup>r</sup> A. OUABDESSALEM pour ses nombreux conseils et sa grande patience et gentillesse qu'il sait témoigner à tous les étudiants qui le sollicitent.

- AU D<sup>r</sup> KERBACHE et à M<sup>r</sup> T. LAMRAOUI pour avoir accepté de faire partie du jury.

- A M<sup>r</sup> H. SARI pour avoir voulu répondre à notre invitation et faire partie du jury et surtout pour son aide considérable et ses précieux conseils.

- A toutes les personnes qui ont contribué de près ou de loin à l'élaboration de ce travail et en particulier au D<sup>r</sup> A. BOUZAHER qui a été à l'origine de cette thèse et à M<sup>r</sup> H. DJOUZI ingénieur au centre de calcul qui a pu établir l'ambiance de travail dont les étudiants avaient besoin.

- A mon époux et à ma mère. Leur affection et leur soutien moral ont pu me donner la volonté nécessaire pour aller au bout de mon travail.

## P L A N

	Pages
INTRODUCTION.	1
CHAPITRE 1 : GENERALITES SUR LA PROGRAMMATION DETERMINISTE.	5
1.1: INTRODUCTION	5
1.2: PROGRAMMATION LINEAIRE.	6
1.3: PROGRAMMATION NON LINEAIRE.	6
1.3.1: Programmation non linéaire avec contraintes.	6
A: Définitions et propositions.	6
B: Méthodes primales	6
C: Méthodes duales:	7
1.3.2: Programmation non linéaire sans contraintes.	8
A: Méthodes avec dérivées.	8
B: Méthodes sans dérivées.	8
1.4: CONCLUSION.	11
	11
CHAPITRE 2 : GENERALITES SUR LA PROGRAMMATION STOCHASTIQUE.	13
2.1: PROGRAMMATION LINEAIRE STOCHASTIQUE.	14
2.1.1: Approches actives.	15
A: Chance constrained programming.	15
B: Fat formulation.	15
C: Problèmes à recours.	16
2.1.2: Approche passive.	
2.2: PROGRAMMATION NON LINEAIRE STOCHASTIQUE.	19
2.2.1: Approches actives.	21
A: Approximations techniques.	21
B: Méthode du quasi-gradient stochastique.	23
B.1: Méthode du sous gradient déterministe.	23
B.2: Méthode du quasi-gradient stochastique.	24
a: Gradient d'une fonction aléatoire.	25
a.1: Cas où la fonction est différentiable.	25
a.2: Cas où la fonction est non différentiable.	28
a.3: Cas où la fonction est quelconque.	30

b: Choix du pas.	31
b.1: Ratio de la fonction estimée à la longueur du chemin de l'algorithme.	34
b.2: Ratio de l'estimation du gradient.	34
b.3: Ratio du chemin progressif.	35
b.4: Ratio des directions.	35
2.3: CONCLUSION.	36
CHAPITRE 3 : METHODE DE RESOLUTION ET RESULTATS.	37
3.1: APPROCHE ETUDIEE	37
3.2: METHODOLOGIE DE TRAVAIL.	37
3.3: METHODES DE RESOLUTION ET RESULTATS.	40
3.3.1: CONSIDERATIONS LIEES A LEURS IMPLEMENTATION.	40
A: LE PAS DE PROGRESSION.	41
B: LE CRITERE D'ARRET.	41
C: LE POINT INITIAL.	41
D: ERREUR DE LA SOLUTION.	43
3.3.2: METHODE DES GRADIENTS CONJUGUES.	43
A: ALGORITHME.	44
B: CONSIDERATIONS LIEES A SON IMPLEMENTATION.	45
C: MISE EN OEUVRE ET RESULTATS.	47
D: SYNTHESE ET CONCLUSIONS.	49
3.3.3: METHODE DE HOOK ET JEEVES.	50
A: ALGORITHME.	51
B: CONSIDERATIONS LIEES A SON IMPLEMENTATION.	52
C: MISE EN OEUVRE ET RESULTATS.	53
D: CONCLUSION DE L'ANALYSE	62
CONCLUSION	63
BIBLIOGRAPHIE	
ANNEXE 1	
ANNEXE 2	
ANNEXE 3	
ANNEXE 4	

SYMBOLES ET NOTATIONS

المدرسة الوطنية المتعددة التقنيات  
 المكتبة — BIBLIOTHEQUE  
 Ecole Nationale Polytechnique

$\cap$	intersection.
$\cup$	union.
$\in$	appartient à.
$\forall$	quelque soit.
$\exists$	il existe.
Opt	optimum: indique un minimum ou un maximum. $\text{Min } f(x) = - \text{Max } (- f(x)).$
$A^T$	matrice transposée de la matrice A.
$\langle x, y \rangle$	produit scalaire des vecteurs x et y.
$\nabla f(x)$	gradient total de f(x).
$\partial f(x)$	gradient partiel de f(x).
$[x]$	vecteur x.
$[x-y]$	vecteur x-y.
$\  \cdot \ $	norme.
$V_x$	voisinage de x.
$f \in C^1$	f est différentiable une fois et continue sur C.
$f \in C^2$	f est différentiable deux fois et continue sur C.
$H(x)$	matrice hessienne: matrice des dérivées secondes.
$E_x$	esperance mathématique de x.
$\Omega$	référentiel, espace des épreuves ou des résultats possibles.
$\omega$	point élément ou épreuve.
$F(\Omega)$	ensemble de tous les événements liés à l'expérience.
$(\Omega, F(\Omega), P)$	espace des probabilités.
$P$	mesure de probabilité correspondante.
$\xi$	vecteur des paramètres aléatoires.
$\xi_i$	i ème composante du vecteur $\xi$ .
$\xi^l$	l'observation l du vecteur $\xi$ .
$\xi_{i,k}$	i ème composante de la l ième observation du vecteur $\xi$ à l'itération k.
$s$	nombre d'observations.
$k$	nombre d'itérations.
$\theta_k$	pas de déplacement à l'itération k.
$A$	algorithme

INTRODUCTION

## INTRODUCTION

Un grand nombre de problèmes économiques de décision sont sujets à des phénomènes aléatoires. Nous pouvons citer entre autres, les problèmes de gestion de production et des stocks, d'économie politique, de transport, etc...

Dans les problèmes de programmation mathématique déterministes, les données sont supposées parfaitement définies. Or en pratique, ceci n'est pas souvent vérifié. Ainsi, nous ne pouvons pas définir avec exactitude les prix du marché par exemple.

La question qui se pose alors est:

Les modèles déterministes sont-ils applicables dans ce cas? sinon dans quels cas le seront-ils?

Pour ces problèmes de décision il sera essentiel de définir:

- Les phénomènes aléatoires et leur(s) influence(s) sur la décision.
- Le(s) modèle(s) mathématique(s) adéquat(s), capable(s) de tenir compte du phénomène aléatoire.
- La politique optimale pour une version déterministe de ces modèles.
- Les quantités d'informations nécessaires sur les distributions de probabilités, afin d'établir une solution optimale.

Définition d'un programme mathématique stochastique:

Soient l'ensemble des réels  $\mathbb{R}^n$  et  $S$  un sous-ensemble de  $\mathbb{R}^n$ .

Soit  $f$  une fonction de  $\mathbb{R}^n$  dans  $\mathbb{R}$ .

Soit  $g_i, i \in I$ , une fonction de  $\mathbb{R}^n$  dans  $\mathbb{R}$ .  $I$  ensemble d'indices fini.

Soit enfin le programme mathématique déterministe formulé comme suit:

Il s'agit de trouver  $\hat{X} \in S$  tel que  $f(\hat{X}) = \text{opt} \{f(X)\}$   
 $X \in S$



$$\text{Où } S = \bigcap_{i \in I} D_i$$

$$D_i = \{X \in \mathbb{R}^n / g_i(X) \leq 0\}$$

Supposons, maintenant, que la fonction  $f(X)$  ou que les fonctions  $g_i(X)$  contiennent des paramètres aléatoires. Soit  $\xi$  le vecteur de ces paramètres,  $\xi$  de dimension  $d$  ( $d \leq n$ ).

Il s'agira alors de trouver  $\hat{X} \in S$  tel que  $f(\hat{X}) = \underset{X \in S}{\text{opt}} \{f(X, \xi)\}$

Où

$$S = \bigcap_{i \in I} D_i$$

avec

$$D_i = \{X \in \mathbb{R}^n / g_i(X, \xi) \leq 0\}$$

et  $\xi \in (\Omega, \mathcal{F}(\Omega), p)$ , espace de probabilités.

La théorie de la décision et l'analyse statistique ont montré que pour ce type de problème, la fonction objectif est exprimée comme étant l'espérance mathématique de la fonction à optimiser:

$$F(X) = E[f(X, \xi)]$$

$$F(X) = \int_{\Omega} f(x, \xi(\omega)) P(d(\omega))$$

Où  $\Omega$  est l'espace de probabilités

$P$  est la mesure de probabilités correspondante

Le programme stochastique s'écrit comme suit:

Trouver  $X \in S$ ,  $S \subset \mathbb{R}^n$ , qui optimise  $F(X)$

Soit encore  $\underset{X \in S}{\text{opt}} F(X) = \underset{X \in S}{\text{opt}} \int_{\Omega} f(x, \xi(\omega)) P(d(\omega))$

Le programme stochastique est exprimé en termes d'intégrales multiples. Ainsi, il peut-être résolu en faisant appel à des routines de programmation non linéaire en combinaison avec des routines d'intégration multiple, pour évaluer la fonction  $f$  et son gradient (ou sous-gradient selon le cas).

Cependant, les routines d'intégration disponibles sont à une dimension. A deux dimensions, commencent à apparaître certaines difficultés numériques (problèmes de convergences, propagation des erreurs, augmentation du temps de calcul,...) ; pour la dimension trois, les routines disponibles sont rapportées uniquement à certains cas spéciaux.

En effet, les problèmes impliquent souvent un nombre de variables aléatoires allant de 5 à 100, ce qui rend impossible l'utilisation de ces routines.

Beaucoup d'efforts de recherche ont été conduits dans ce sens afin d'élaborer des routines d'intégrations multiples qui reposent surtout sur les techniques d'échantillonnage; ceci implique nécessairement une génération de nombres pseudo ou quasi-aléatoires.

Bien que ces recherches ne soient pas abandonnées ([SZA 86] [GAS 84] [NIE 86]) on s'oriente beaucoup vers l'élaboration de procédures de solutions pour l'optimisation stochastique. Ces procédures peuvent être divisées en deux catégories :

- Méthodes de descente, reposant sur des directions déterminées par des estimations stochastiques du sous-gradient de  $F$ .
- Méthodes d'approximations qui remplacent la distribution originale  $P$  par une nouvelle distribution plus simple du point de vue numérique.

Toutefois, nous pourrions considérer une méthode de résolution, s'inspirant des travaux établis en programmation linéaire stochastique: La méthode passive ou du "Wait and See". Elle consiste à étudier le comportement de l'optimum sous l'effet aléatoire des paramètres. C'est cette dernière méthode qui fera l'objet de notre étude.

Par cette étude, nous espérons établir l'existence d'une loi probabiliste générale gérant le comportement de l'optimum. Ce résultat peut être intéressant lorsque le coût des méthodes actives devient trop élevé ou que leur application est trop complexe. Le décideur pourra alors disposer d'informations nécessaires pour la décision telles que la moyenne et la variance et même la loi de distribution de l'optimum.

L'étude s'articule sur trois chapitres et est menée de la manière suivante:

Dans le chapitre I nous présentons un état de l'art en programmation mathématique et en particulier en programmation déterministe qui nous permettra d'introduire la programmation stochastique et de mieux la comprendre.

Le chapitre II est une rétrospective des travaux réalisés en matière de programmation stochastique. Nous passons en revue les différentes approches de résolution des problèmes de programmation stochastique de même que nous introduirons l'approche du Wait and See qui fait l'objet de notre étude.

Le chapitre III définit la méthodologie de travail. Deux variantes sont exposées et leurs résultats numériques sont donnés et analysés.

Enfin nous concluons par quelques suggestions destinées à améliorer et enrichir notre étude.

CHAPITRE 1:

GENERALITE SUR LA PROGRAMMATION DETERMINISTE

## 1.1: INTRODUCTION:

L'approche scientifique, pour la prise de décision, est appelée "Recherche Opérationnelle" [BRA 77]. La programmation mathématique constitue l'élément le plus utilisé dans ce domaine.

En 1947, lorsque B. DANTZIG a développé la méthode du simplexe, La programmation mathématique et particulièrement la programmation linéaire, ont connu leur essor.

D'une façon très générale, un programme mathématique est un problème d'optimisation sous contraintes de la forme:

$$\begin{aligned} & \text{Optimiser} && f(X) \\ \text{Sous contraintes} && g_i(X) \leq 0 && i=\overline{1,m} \\ && X \in S, S \subset \mathbb{R}^n && \end{aligned} \quad (1)$$

$X \in \mathbb{R}^n$ ; il a pour composantes  $x_1, x_2, \dots, x_n$  qui sont les variables de décision et inconnues du problème.

La fonction  $f$  est appelée fonction objectif ou économique.

L'ensemble des conditions  $g_i(X) \leq 0$  ( $i=\overline{1,m}$ ) et  $X \in S$ , sont les contraintes du problème.

Lorsque les fonctions  $f$  et  $g_i$  ( $i=\overline{1,m}$ ) sont toutes linéaires, le modèle sera un modèle de programmation linéaire. Par contre, si les fonctions  $f$  et/ou  $g_i$  ( $i=\overline{1,m}$ ) sont non linéaires, le modèle sera un modèle de programmation non linéaire.

Dans certains cas,  $f$  ou  $g_i$  ( $i=\overline{1,m}$ ) sont caractérisées par des distributions de probabilités. Si  $f$  et  $g_i$  ( $i=\overline{1,m}$ ) sont toutes linéaires, alors le modèle sera un modèle de programmation stochastique linéaire. Dans le cas contraire, le modèle sera un modèle de programmation stochastique non linéaire.

En programmation mathématique nous distinguerons donc:

- La programmation mathématique déterministe.
- La programmation mathématique stochastique.

Chacune de ces classes se compose des deux sous classes suivantes:

- Les modèles linéaires.
- Les modèles non linéaires.

## 1.2: PROGRAMMATION LINEAIRE:

Soit le problème de programmation mathématique (1).

Si les fonctions  $f$  et  $g_i$  ( $i=\overline{1,m}$ ) sont déterministes, le problème (1) est un problème de programmation mathématique déterministe.

Dans ce cas nous distinguons deux classes de programmation:

- La programmation linéaire
- La programmation non linéaire.

Un problème de programmation mathématique linéaire est formulé comme suit:

$$\begin{aligned} \text{Opt} \quad & f(x) = C \cdot X = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \\ \text{S.c} \quad & A \cdot X \leq B \\ & X \geq 0 \end{aligned} \quad (2)$$

$n$  nombre de variables de décision.

$m$  nombre de contraintes ( $m \leq n$ ).

$(x_1, x_2, \dots, x_n)^T$  composantes du vecteur des variables de décision,  $X \in \mathbb{R}^n$

$(c_1, c_2, \dots, c_n)$  composantes du vecteur ligne des coûts,  $C \in \mathbb{R}^n$ .

$(b_1, b_2, \dots, b_m)$  composantes du vecteur des ressources,  $b \in \mathbb{R}^m$ .

$a_{ij}$  ( $i=\overline{1,m}$  et  $j=\overline{1,n}$ ) éléments de la matrice des coefficients technologiques,  $A(m \times n)$ .

## 1.3: PROGRAMMATION NON LINEAIRE:

### 1.3.1: PROGRAMMATION NON LINEAIRE AVEC CONTRAINTES: [BAA 77][MIN 83]

A: Définitions et propositions:

Soit le problème (1) de programmation mathématique, dans le cas non linéaire et d'une minimisation, il se formule comme suit:

$$\begin{aligned} \text{Min} \quad & f(X) \\ \text{S.c} \quad & g_i(X) \leq 0 \quad i=\overline{1,m} \\ & X \in S, S \subset \mathbb{R}^n \end{aligned} \quad (3)$$

Les fonctions  $f$  et  $g_i (i=\overline{1,m})$  sont supposées continues et différentiables.

Nous noterons  $NL_x$  l'ensemble des solutions de (3):

$$NL_x = \{ X \in \mathbb{R}^n / g_i(X) \leq 0, i=\overline{1,m} \}$$

Soit  $\mathcal{L}(X,\lambda)$  la fonction de Lagrange,  $\lambda \in \mathbb{R}^m$ :

$$\mathcal{L}(X,\lambda) = f(X) + \sum_{i=1}^m \lambda_i g_i(X)$$

On appelle dual de (3) (par opposition (3) est appelé problème primal), le problème suivant:

$$\text{Max } W(\lambda) = \text{Max}_{\lambda \geq 0} \{ \text{Min}_{X \in S} \mathcal{L}(X,\lambda) \} \quad \text{et } \lambda \in \mathbb{R}^m \quad (4)$$

Nous avons alors:

$$\forall \lambda \geq 0, W(\lambda) \leq W(\lambda^*) \leq f(X^*)$$

Où  $\lambda^*$  est la valeur optimale du problème dual (4).

$X^*$  est la valeur optimale du problème primal (3).

Il existe en programmation non linéaire contrainte deux grandes familles de méthodes de résolution:

- Les méthodes primales.
- Les méthodes duales.

**B: Les méthodes primales:**

Comme leur nom l'indique, elles opèrent directement sur le problème primal. Elles ont la caractéristique:

- d'être itératives.
- d'assurer une décroissance monotone de la fonction à minimiser (ou une croissance monotone de la fonction à maximiser).

Elles présentent donc l'avantage de procurer, à chaque instant du processus itératif, une solution approchée satisfaisant les contraintes.

L'inconvénient de ces méthodes est qu'elles ont une mauvaise convergence globale.

Parmi ces méthodes nous citerons [MIN 83] [BAA 77]:

- La méthode de ZOUTENDIJK, basée sur les directions admissibles (ne converge pas tout le temps).
- La méthode des contraintes actives.
- Les méthodes du gradient projeté (ROSEN).
- La méthode du gradient réduit (WOLFE).
- La méthode du gradient généralisé (ABADIE).
- La méthode de linéarisation.

**C: Méthodes duales:**

Elles utilisent la notion de dualité. Elles opèrent sur le dual du problème donné. Elles ont la caractéristique:

- d'être robustes.
- d'avoir une convergence globale facile à obtenir.

L'inconvénient de ces méthodes est que la solution du primal n'est obtenue qu'en fin de convergence car :  $\hat{w}(\lambda^*) \leq f(x^*)$

Parmi ces méthodes, nous citerons les méthodes des pénalités, qui ont pour principe de remplacer le problème primal de programmation non linéaire avec contraintes (4) par un problème de programmation non linéaire sans contraintes. Nous citons, parmi ces méthodes [MIN 83]:

- Méthodes des pénalités extérieures.
- Méthodes des pénalités intérieures ou barrières.

### 1.3.2: PROGRAMMATION NON LINÉAIRE SANS CONTRAINTES:

**A: Méthodes avec dérivées:**

Soit le problème (1) de programmation mathématique, dans le cas non linéaire et sans contraintes, il se formule comme suit:

$$\begin{aligned} \text{Opt } f(x) \\ x \in \mathbb{R}^n \end{aligned} \quad (5)$$



$f$  est une fonction à valeurs réelles  $f: \mathbb{R}^n \rightarrow \mathbb{R}$

Dans le cas d'une minimisation, nous devons déterminer un point  $X^*$  de  $\mathbb{R}^n$  tel que:

$$\forall X \in \mathbb{R}^n \quad f(X^*) \leq f(X) \quad X^* \text{ est alors un minimum global de } f \text{ sur } \mathbb{R}^n. \quad (6)$$

Si l'inégalité (6) est stricte :  $f(X^*) < f(X)$ ,  $\forall X \in \mathbb{R}^n$ ,  $X \neq X^*$ , le minimum global  $X^*$  est alors unique.

### 1: Conditions nécessaires et suffisantes d'optimalité:

Soit le problème (5),  $X^*$  est optimum de  $f$  sur  $\mathbb{R}^n$  s'il vérifie les conditions d'optimalité.

#### Définition:

Soit  $S$  un ensemble de  $\mathbb{R}^n$ . Soit  $X \in S$  et soit  $d \in \mathbb{R}^n$ . On dit que le vecteur  $d$  est une direction admissible au point  $X$  si:

$$\exists \bar{\alpha} \in \mathbb{R}, \bar{\alpha} > 0 \text{ tel que } X + \alpha d \in S, \forall \alpha, 0 \leq \alpha \leq \bar{\alpha}.$$

#### Condition nécessaire du premier ordre:

Soient  $S \subset \mathbb{R}^n$  et  $f$  une fonction de  $S$  dans  $\mathbb{R}$ .  $f \in C^1$  sur  $S$ . Si  $X^*$  est un minimum local de  $f$  sur  $S$  alors:

$$\forall d \in \mathbb{R}^n, d \text{ direction admissible de } \mathbb{R}^n \text{ au point } X^* \text{ nous avons:}$$

$$\nabla f(X^*) \cdot d \geq 0.$$

#### Condition nécessaire du second ordre:

Soient  $S \subset \mathbb{R}^n$  et  $f$  une fonction de  $S$  dans  $\mathbb{R}$ .  $f \in C^2$  sur  $S$ . Si  $X^*$  est un minimum local de  $f$  sur  $S$  alors pour toute direction  $d \in \mathbb{R}^n$  admissible au point  $X^*$  nous avons:

$$1- \quad \nabla f(X^*) \cdot d \geq 0.$$

$$2- \text{ Si } \nabla f(X^*) \cdot d = 0 \text{ alors } d^T \nabla^2 f(X^*) d \geq 0.$$

Condition suffisante:

Soient  $S \subset \mathbb{R}^n$  et  $f$  une fonction de  $S$  dans  $\mathbb{R}$ .  $f \in C^2$  sur  $S$ . Soit  $X^*$  un point intérieur de  $S$ . Nous supposons de plus que  $\nabla f(X^*) = 0$  et que la matrice  $\nabla^2 f(X^*)$  est définie positive ( $\forall Y \in \mathbb{R}^n, Y \neq 0, Y^T \nabla^2 f(X^*) Y > 0$ ). Alors  $X^*$  est un minimum local strict de  $f$ .

Pour beaucoup de problèmes d'optimisation sans contraintes, les principales méthodes connues ne permettent pas la détermination d'un optimum global. Souvent, ces méthodes convergent vers un optimum local, c'est à dire un point qui vérifie (6) dans un voisinage de  $X^*$ .

- Méthodes numériques d'optimisation avec dérivées:

Les méthodes utilisées sont en général des procédures itératives engendrant une suite de points  $X^0, X^1, \dots, X^k$  qui converge vers l'optimum local de  $f$ . A chaque étape  $k$ , le point  $X^{k+1}$  est défini comme suit:

$$X^{k+1} = X^k + \vartheta_k d^k$$

Où  $\vartheta_k$  est le pas de déplacement. Il peut-être prédéterminé ou choisi de façon à minimiser la fonction de  $\vartheta$ ,  $\psi(\vartheta) = f(X^k + \vartheta d^k)$ :

$$\psi(\vartheta_k) = \min_{\vartheta \geq 0} \psi(\vartheta)$$

$d^k$  est une direction de déplacement qui peut-être:

- Soit le gradient de  $f$  en  $X^k$ ,  $d^k = -\nabla f(X^k)$ .
- Soit choisie de façon arbitraire, à condition que ce soit une direction de descente, c'est à dire vérifiant la condition :  $\nabla f^T(X^k) d^k < 0$ .

Parmi ces méthodes nous citons [BAA 77]:

- Méthode de Newton (cas de minimisation): la direction de descente est choisie telle que:

$$d^k = [\nabla^2 f(X^k)]^{-1} \nabla f(X^k)$$

- Méthodes des directions conjuguées pour la résolution des fonctions quadratiques et non quadratiques. Nous avons:

- a- La méthode de Davidon-Fletcher-Powell (1963).
- b- La méthode des gradients conjugués de Fletcher et Reeves (1964).
- c- La méthode de Zangwill (1967).

**B: Méthodes sans dérivées:**

Contrairement aux méthodes précédentes, ces méthodes n'utilisent pas la dérivée de la fonction  $f$ .

Elles procèdent de la manière suivante:

Soient  $X$  un vecteur et  $d$  une direction de descente.  $f$  est alors optimisée suivant la direction  $d$ . La méthode la plus simple dans ce cas est celle des coordonnées cycliques:

**Méthode des coordonnées cycliques:**

Cette méthode recherche l'optimum le long de  $n$  directions  $e^1, e^2, \dots, e^n$  où  $e^i$  est un vecteur unitaire. L'inconvénient de cette méthode est qu'elle ne converge pas vers l'optimum lorsque la fonction n'est pas différentiable. Dans ce dernier cas, la méthode de HOOK et JEEVES est applicable, elle est une amélioration de la méthode précédente.

#### 1.4: Conclusion:

Soit  $A$  un algorithme. C'est une application définie sur un ensemble  $S$ , qui à tout point  $X \in S$  fait correspondre un sous ensemble de  $S$ :

$$X^{k+1} = A(X^k)$$

Soit la norme d'ordre  $p$ ,  $p \in \mathbb{R}$ , notée  $\| \cdot \|_p$

On appelle taux de convergence d'un algorithme le rapport suivant:

$$v = \lim_{k \rightarrow \infty} \frac{\| X^{k+1} - X^* \|}{\| X^k - X^* \|_p}$$

Si  $p = 1$  : -  $v < 1$ , convergence linéaire.

-  $v = 0$ , convergence superlinéaire.

Si  $p = 2$  : -  $\nu$  quelconque, convergence quadratique.

Si  $p > 2$  : -  $\nu$  quelconque meilleure convergence. Plus  $p$  est grand, meilleure est la convergence.

La méthode des gradients conjugués et la méthode de hook et Jeeves présentent une convergence superlinéaire.

Le principal avantage de ces méthodes est qu'elles ne peuvent diverger. A partir d'un point  $X^0 \in \mathbb{R}^n$ , elles consistent à générer une suite  $X^1, X^2, X^3, \dots$  telle que la suite  $f(X^1), f(X^2), f(X^3), \dots$  soit monotone. En général ces méthodes convergent vers un optimum de la fonction  $f(X)$ , mais cet optimum peut être relatif, c'est là leur principal inconvénient. Ainsi est-il nécessaire, pour ces méthodes, après avoir obtenu un optimum de  $f(X)$  de vérifier si en ce point  $X^*$ , on a  $f'(X^*)=0$ , sinon cet optimum est relatif. Il faut alors changer de point de départ et recommencer une nouvelle recherche de l'optimum. Pour cela, dans les algorithmes qui seront présentés ultérieurement, il ya toujours une phase de réinitialisation.

CHAPITRE 2:

GENERALITE SUR LA PROGRAMMATION STOCHASTIQUE

**2: GENERALITES SUR LA PROGRAMMATION STOCHASTIQUE:**

Soit le problème de programmation stochastique suivant:

$$\begin{aligned} & \text{Opt } f(X, \xi) \\ & \text{S.c } g_i(X, \xi) \leq 0 \quad i=\overline{1, m} \\ & X \in S, S \subset \mathbb{R}^n \end{aligned}$$

$X$  est le vecteur des variables de décision,  $X \in \mathbb{R}^n$ .

$\xi$  est le vecteur des paramètres aléatoires,  $\xi \in \mathbb{R}^d$ .

De même qu'en programmation déterministe, nous distinguons:

- La programmation stochastique linéaire, où les fonctions  $f$  et  $g_i (i=\overline{1, m})$  sont toutes linéaires.
- La programmation stochastique non linéaire, où les fonctions  $f$  et/ou  $g_i (i=\overline{1, m})$  sont non linéaires.

Pour les problèmes de programmation stochastique, nous distinguons deux catégories de méthodes de résolution:

- Les approches passives (wait and see).
- Les approches actives (here and now).

**Approches passives:**

Il s'agit d'observer les variables aléatoires, puis de résoudre un ensemble de problèmes déterministes, afin d'étudier le comportement de l'optimum. Cette étude permettra de déduire la loi de probabilité de l'optimum, ainsi que sa moyenne et sa variance.

**Approches actives:**

Elles consistent à agir sur les variables aléatoires avant leur observation, afin de résoudre un problème déterministe et déduire la valeur de l'optimum.

## 2.1: PROGRAMMATION STOCHASTIQUE LINEAIRE

De même qu'en programmation déterministe, où la programmation linéaire a connu un champ d'application très vaste [BER 68] [CHA 58] [WIL 66] [ZIE 74], vu la simplicité de modélisation et la rapidité de résolution du problème, la programmation stochastique linéaire a elle aussi vu cette grande évolution.

Soit le problème de programmation linéaire suivant:

$$\begin{aligned} \text{Opt } f(X) &= C^T X \\ \text{S.c } g_i(X) &\leq 0 \quad i = \overline{1, m} \\ X &\geq 0 \end{aligned} \quad (1)$$

Si les fonction  $f$  et  $g_i (i = \overline{1, m})$  sont des fonctions aléatoires, le modèle le plus satisfaisant sera celui qui tiendra compte du risque. Ceci est obtenu en considérant (1) comme un problème probabiliste, donc formulé comme suit:

$$\begin{aligned} \text{Opt } f(X, \xi) \\ \text{S.c } g_i(X, \xi) &\leq 0 \quad i = \overline{1, m} \\ X \in \mathbb{R}^n, \quad X &\geq 0 \end{aligned} \quad (2)$$

$f$  et  $g_i (i = \overline{1, m})$  sont des fonctions aléatoires.

$\xi$  est le vecteur des paramètres aléatoires.

$\xi \in \mathbb{R}^d$ ,  $d \leq n$ ,  $d \leq m$ .

Les méthodes de résolution de cette classe de problèmes sont nombreuses et variées. De plus, les recherches sont surtout orientées vers les méthodes actives.

Parmi les approches actives nous retiendrons:

- "The chance constrained programming" de Charnes, Symond et Cooper [CHA 58].
- "The fat formulation" de Mandansky [MAN 62].
- "Programming with recourse" ou les problèmes à étages. Cet approche a fait l'objet de beaucoup de travaux du fait de sa large application à de nombreux problèmes pratiques.

Parmi les approches passives, nous retiendrons:

- "Stochastic programming" de Tintner [TIN 55].

### 2.1.1: Approches actives:

#### A: Chance constrained programming:

Nous supposons que l'aléa n'intervient que dans les contraintes. Le problème est formulé alors comme suit:

$$\begin{aligned} \text{Opt } f(X) &= C^T X \\ \text{S.c } g_i(X, \xi) &\geq 0 \quad i=\overline{1, m} \\ X &\geq 0 \end{aligned} \quad (3)$$

L'hypothèse émise est que chaque contrainte doit être satisfaite avec un certain taux  $\alpha_i$ . La formulation devient alors:

$$\begin{aligned} \text{Opt } f(X) &= C^T X \\ \text{S.c } X &\in U \end{aligned} \quad (4)$$

$$\text{Où } U = \{ X \in \mathbb{R}^n / P(g_i(X, \xi) \geq 0) = \alpha_i \}$$

Où  $\xi$  a une distribution de probabilité  $P$  connue.

(4) sera résolu comme un simple programme linéaire déterministe.

#### B: Fat formulation:

En considérant toujours le problème (3), Pour toute valeur observée  $\xi^l$  de  $\xi$ , nous choisissons  $g_i(X, \xi^l) \geq 0$ , nous aurons alors:

$$\begin{aligned} \text{Opt } f(X) \\ \text{S.c } X \in G \end{aligned} \quad (5)$$

$$\text{Où } G = \{ X \in \mathbb{R}^n / g_i(X, \xi^l) \geq 0 \quad i=\overline{1, m} \quad l=\overline{1, L} \}$$

Où  $L$  est le nombre de valeurs possibles pour le vecteur aléatoire  $\xi$ .

Si  $L$  est fini, le problème est discret; (5) est alors un problème de programmation déterministe que nous pouvons résoudre.



Si  $L$  est infini et s'il existe un nombre  $L_1$  fini de points  $\xi^1, \xi^2, \dots, \xi^{L_1}$  tel que :

$$P(\xi = \xi^{L_1}) = 1$$

Nous nous ramenons au cas où  $L$  est fini.

L'inconvénient dans cette approche est que  $G$  peut être vide, ou contenir très peu d'éléments. Le domaine risque alors d'être non borné. Ceci rendra la résolution du problème (5) impossible.

C: Problèmes à recours:

Soit le problème de programmation linéaire suivant:

$$\text{Opt } f(X) = C^T X$$

$$\text{S.c } A X = b$$

$$X \geq 0$$

(6)

$X \in \mathbb{R}^n$ ,  $C \in \mathbb{R}^n$ ,  $b \in \mathbb{R}^m$  et  $A(m \times n)$ .

Supposons que le vecteur  $b$  et/ou le vecteur  $C$  soient aléatoires. Pour un vecteur de décision  $X$  donné, les contraintes du problème (6) peuvent ne pas être satisfaites. Dans les problèmes avec recours, nous supposons que nous avons accès à une action de recours, pour rendre les contraintes consistantes. Il est alors possible d'apporter des actions correctrices après l'observation des paramètres aléatoires. Lorsque les contraintes sont violées, il est prévu des pénalités à cet effet. Soient alors:

$C$  : Vecteur des coûts.  $C \in \mathbb{R}^n$

$X$  : Vecteur des variables de décision.  $X \in \mathbb{R}^n$ .

$A X = b$  : Contraintes des ressources.  $A(n \times m)$ ,  $b \in \mathbb{R}^m$ .

$\xi$  : Vecteur des paramètres aléatoires.  $\xi \in \mathbb{R}^d$ ,  $d \leq n$ .

$W$  : l'output

$T$  : Matrice technologique, contient la transformation linéaire d'activités en produits finis  $W = T X$ .

En général, pour les problèmes stochastiques, l'objectif est d'optimiser l'espérance mathématique de la fonction objectif aléatoire sous certaines contraintes..

Pour les problèmes à recours et pour un cas de minimisation, l'objectif est de minimiser l'espérance mathématique du coût total composé de deux parties:

1/- le coût de la production  $C^T X$ .

2/-le coût de la pénalité pour la violation des contraintes.

Soit le produit  $i$ ,  $i=1, \overline{m}$ , l'output  $\omega_i$  est comparé avec la valeur observée  $\xi_i$ . Tout déséquilibre entre  $\xi_i$  et  $\omega_i$  est pénalisé comme suit:

Si  $\xi_i < \omega_i$  : La pénalité est  $q_i (\omega_i - \xi_i) = q_i y^-$   
représente l'excès de produit.

Si  $\xi_i \geq \omega_i$  : La pénalité est  $q_i (-\omega_i + \xi_i) = q_i y^+$   
représente le manque de produit.

$q_i$ ,  $y^+$  et  $y^-$  appartiennent à  $\mathbb{R}^m$ .

Le problème est formulé alors comme suit:

$$\text{Opt } C^T X + Q^*(X, \xi)$$

$$\text{S.c } A X = b$$

$$X \geq 0$$

(7)

Où  $Q^*(X, \xi)$  minimise le problème suivant:

$$\text{Opt } q^+ y^+ + q^- y^- = Q(X, \xi)$$

$$\text{S.c } y^+ - y^- = \xi - T X = \xi - w$$

$$y^+ \geq 0$$

$$y^- \geq 0$$

(8)

Le problème (7) est appelé problème du premier étage.

Le problème (8) est appelé problème du second étage.

Les conditions d'optimalité et de faisabilité pour les problèmes à recours ont été étudiés par Wets [WET 72].

Nous avons à prendre une décision avant la réalisation de  $T$ . Il est en général impossible d'exiger que l'égalité (9) soit satisfaite pour chaque réalisation des paramètres stochastiques des contraintes.

$$TX = \xi$$

(9)

Le problème avec recours est une façon de contourner ces difficultés de modélisation. La décision de recours peut être interprétée comme une correction dans (9) et le coût de recours  $Q(X, \xi)$  comme une pénalité pour une divergence en (9).

La formulation d'un problème de programmation stochastique implique le calcul d'intégrales multiples, or les routines efficaces et disponibles traitent de problèmes à une dimension. Parmi les approches proposées, pour éviter le calcul d'intégrales multidimensionnelles, nous citerons:

- l'approche de Wets [WET 74], qui consiste à réduire cette intégration à  $m$  intégrations unidimensionnelles.
- L'approche de Qi Liqun [QI 86] qui consiste à définir deux sous programmes du problème (7)-(8) plus simples à résoudre.
- l'approche de Birge et Wets [BIR 89] qui consiste à remplacer la fonction objectif par une autre fonction qui la borne supérieurement pour un cas de maximisation; le problème est alors simplifié.

### 2.1.2: Approche passive:

Soit le problème de programmation stochastique linéaire (3).

Pour chaque  $\xi^l$  ( $l=1, L$ ) où  $L$  est le nombre de valeurs possibles de  $\xi$ , nous résolvons le problème suivant:

$$\text{Opt } f(X)$$

$$\text{S.c } X \in G_l$$

$$\text{Où } G_l = \{X \in \mathbb{R}^n, X \geq 0 / g_i(X, \xi^l) \geq 0 \quad i=1, m\} \quad l=1, L$$

L problèmes de programmation déterministe sont résolus. Ainsi nous générons la distribution de l'optimum, en nous basant sur celle de  $\xi$ . A partir de cette distribution, nous aurons les informations nécessaires en ce qui concerne la moyenne, la variance et éventuellement la distribution de loi de probabilité de l'optimum.

## 2.2: PROGRAMMATION NON LINEAIRE STOCHASTIQUE:

Soit le problème de programmation non linéaire suivant:

$$\begin{array}{l} \text{Opt } f(X) \\ X \in S \end{array}$$

Si  $S \subset \mathbb{R}^n$ , le problème est contraint.

Si  $S = \mathbb{R}^n$ , le problème est sans contraintes.

Si la fonction  $f$  et/ou les contraintes sont fonction de paramètres aléatoires, le problème devient un problème non linéaire stochastique. Nous aurons alors:

$$\begin{array}{l} \text{Opt } f(X, \xi) \\ X \in S \end{array}$$

$X$  est le vecteur des variables de décision.  $X \in \mathbb{R}^n$ .

$\xi$  est le vecteur des variables aléatoires.  $\xi \in \mathbb{R}^d$ .

En programmation stochastique, l'objectif est défini comme étant l'espérance mathématique de la fonction  $f(X, \xi)$ . Il est exprimé alors comme suit:

$$\begin{array}{l} \text{Opt} \\ X \in \mathbb{R}^n \end{array} F(X) = E_{\xi} [ f(X, \xi) ] = \int_{\Omega} f(X, \xi(\omega)) P(d(\omega)) \quad (10)$$

$\omega \in \Omega$  où  $(\Omega, \mathcal{F}(\Omega), P)$  est l'espace de probabilités

$P$  la mesure de probabilité correspondante.

Si  $\xi$  est un vecteur discret, pouvant prendre nombre fini de valeurs  $\xi^1, \xi^2, \dots, \xi^L$  avec des probabilités  $P_1, P_2, \dots, P_L$  tel que  $\sum_{l=1}^L P_l = 1$ . Nous pouvons écrire:

$$F(X) = \sum_{l=1}^L P_l f(X, \xi^l) \quad (11)$$

Si  $\xi = (\xi_1, \xi_2, \dots, \xi_d)$  a une densité de probabilité  $\varphi(\xi_1, \xi_2, \dots, \xi_d)$ , (10) est exprimé à l'aide d'une intégrale de REIMAN:

$$F(X) = \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} f(X, \xi) \varphi(\xi) d\xi_1 d\xi_2 \dots d\xi_d \quad (12)$$

Ainsi, le calcul de  $F(X)$  donné par (12), nécessite l'évaluation d'une intégrale multiple. S'il est difficile d'obtenir analytiquement l'intégrale, nous devons utiliser des méthodes numériques nécessitant un effort de calcul qui croît en même temps que le nombre de variables.

L'application directe des méthodes de programmation non linéaire aux problèmes de programmation stochastique non linéaire nécessite le calcul de la forme (10) en chaque point  $X^k$ ,  $k=0,1,\dots$ , généré par l'algorithme non linéaire. Elle peut aussi nécessiter le calcul du gradient  $\nabla F(X^k)$ ,  $k=0,1,\dots$ , ce qui est encore plus difficile que le calcul de la fonction  $F(X)$  elle-même. On note:

$$\nabla F(X) = \int_{\Omega} \nabla_X f(X, \xi(\omega)) P(d(\omega)) \quad (13)$$

Où  $\nabla_X f(X, \xi)$  représente le gradient de  $f(X, \xi)$  en  $X$ .

Si  $\xi$  est discret:

$$\nabla F(X) = \sum_{l=1}^L P_l \nabla_X f(X, \xi^l) \quad (14)$$

Si  $\xi$  est continu:

$$\nabla F(X) = \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} \nabla_X f(X, \xi) \varphi(\xi) d\xi_1 d\xi_2 \dots d\xi_d \quad (15)$$

De même qu'en programmation stochastique linéaire, nous distinguerons deux catégories de méthodes de résolution des problèmes de programmation stochastique non linéaire:

- 1- Approches actives: nous présenterons les méthodes suivantes:
  - Les approximations techniques.
  - La programmation stochastique du quasi-gradient.
- 2- Approche passive du wait and see.

Dans notre exposé, nous n'allons pas différencier la programmation non linéaire stochastique avec contraintes de la programmation sans contraintes, car la procédure est la même [WET 84]. Nous apporterons des précisions en cas de nécessité.

### 2.2.1: APPROCHES ACTIVES:

#### A: APPROXIMATIONS TECHNIQUES:

Cette approche consiste à remplacer le problème original par un problème plus simple. Nous approximons le vecteur  $\xi$  par un autre vecteur  $\tilde{\xi}$ , pour lequel l'intégrale (10) sera facilement calculable.

Si nous choisissons  $\tilde{\xi}$  discret, nous obtenons la forme (11).

Soit  $\Xi \subset \mathbb{R}^d$  le plus petit fermé dans  $\mathbb{R}^d$  tel que  $P[\xi \in \Xi] = 1$

Soit  $C$  une collection finie de sous-ensembles  $\Omega_i$ ,  $i=1, \overline{L\omega}$ , de  $\Xi$  tels que:

$$\bigcup_{i=1}^L \Omega_i = \Xi \quad (16)$$

$$\Omega_i \cap \Omega_j = \emptyset \quad i \neq j, \quad i=1, \overline{L\omega}, \quad j=1, \overline{L\omega} \quad (17)$$

Pour toute partition  $C$  de  $\Xi$  nous pouvons écrire:

$$F(X) = \int_{\Xi} f(X, \xi) P(d\xi) = \sum_{i=1}^L \int_{\Omega_i} f(X, \xi) P(d\xi) \quad (18)$$

En particulier, lorsque  $\xi$  est continu de distribution connue  $\varphi$ :

$$F(X) = \sum_{i=1}^{L\omega} \int_{\Omega_i} \dots \int f(X, \xi) \varphi(\xi) d\xi_1 d\xi_2 \dots d\xi_d \quad (19)$$

L'approximation sur  $\Omega_i$  est établie comme suit:

$$\int_{\Omega_i} f(X, \xi) P(d\xi) \approx f(X, \xi^i) \int_{\Omega_i} P(d\xi) = f(X, \xi^i) P(\xi \in \Omega_i) \quad (20)$$

Où  $\xi^i$  est une sélection représentative de  $\Omega_i$ .

Nous approximations donc  $f(X, \xi)$  par une fonction en  $\xi$  qui est constante dans chaque sous-ensemble  $\Omega_i$ ,  $i=1, L, \omega$ . Nous écrivons alors:

$$F(X) = \sum_{i=1}^{L\omega} P_i f(X, \xi^i) \quad (21)$$

$$\text{Avec } P_i = P(\xi \in \Omega_i)$$

Ceci revient, finalement à approximer  $\xi$  par un vecteur  $\tilde{\xi}$  atteignant les valeurs  $\xi^i$  avec des probabilités  $P_i$ ,  $i=1, L, \omega$ . L'approximation obtenue en (21) est équivalente à la relation (10).

Cependant, il reste difficile de déterminer la partition qui procure la meilleure précision de l'approximation. En effet, la division de  $E$  en pièces  $\Omega_i$ ,  $i=1, L, \omega$  de plus en plus petites, dans le souci d'une plus grande précision, peut mener à des situations plus complexes du point de vue numérique

Supposons  $\tilde{\xi} = (\xi^1, \xi^2, \dots, \xi^{10})$ . Si l'ensemble de définition de chaque  $\xi^j$  est divisé en dix intervalles, nous aurons  $10^{10}$  sous ensembles  $\Omega_j$  de  $E$ . Ce nombre est au delà de n'importe quelle capacité de calcul.

Si nous considérons alors  $\xi^i$  comme suit :

$$\xi^i = E[\xi(\omega) / \xi(\omega) \in \Omega_i] \quad (22)$$

$$\text{Avec } P(\xi(\omega) \in \Omega_i) = 1 \quad (23)$$

Nous obtenons alors, non seulement une bonne précision de l'approximation, mais aussi des informations qui peuvent aider à affiner la partition si la précision n'est pas suffisante.

En effet, si la fonction est linéaire nous aurons en (20) une égalité stricte. Ceci implique qu'une division additionnelle du sous-ensemble  $\Omega_i$  est inutile pour apporter la précision d'approximation au point  $X$ .

Si  $f(X, \xi^l)$  est non linéaire dans  $\Omega_l$ , l'approximation peut être plutôt grossière. De ce fait, la densité de partitionnement dans  $\mathcal{E}$  dépendra de la linéarité ou la non linéarité de  $f(X, \xi^l)$ .

Généralement, les propriétés de la fonction ne sont, préalablement, pas connues. Certaines informations sont obtenues au cours de la résolution d'un problème approximé. En outre, les propriétés de  $f(X, \xi^l)$  changent lorsque le vecteur  $X$  varie. Le but est de trouver la meilleure partition pour un vecteur  $X$ , tendant vers la solution optimale du problème. L'approximation de la solution du problème original et la construction d'une partition de  $\mathcal{E}$  sont mutuellement liées [KAL 88] comme indiqué dans l'approche suivante:

- 1- Choisir un partitionnement initial qui satisfait (16) et (17).
- 2- Choisir  $\xi^l \in \Omega_l$  et des probabilités  $P_{l,i=1,L,\omega}$  d'après (22) et (23).
- 3- Résoudre le problème approximé.
- 4- A la solution  $X^{*k}$ , on analyse la précision de l'approximation en investiguant les propriétés de la fonction  $f(X^{*k}, \xi^l)$  dans chaque sous-ensemble  $\Omega_{l,i=1,L,\omega}$ . Nous choisissons, parmi ces derniers, ceux qui doivent être encore divisés si la précision n'est pas suffisante. On retourne au pas (2).

## B: METHODE DU QUASI-GRADIENT STOCHASTIQUE:

### B.1: Méthode du Sous-gradient déterministe [ZOW 85]:

Supposons la fonction déterministe  $f(X)$  différentiable.

Soit  $X^k$  le point courant de l'itération.

Soit  $\vartheta_k$  le pas d'itération,  $\vartheta_k > 0$ .

La méthode du gradient déterministe, dans le cas d'une minimisation, est donnée par la forme itérative suivante:

$$X^{k+1} = X^k - \vartheta_k \nabla_x f(X^k) \quad (24)$$

$-\nabla_x f(X^k)$  est la direction de descente au point  $X^k$   
(cas de minimisation).



Supposons que la fonction  $f(X)$  soit non différentiable en  $X^i$ . Nous définissons alors une sous différentielle de  $f$  comme suit:

$$\partial f(X^i) = \text{conv} \{ \text{gr} \in \mathbb{R}^n / \text{gr} = \lim_{X^i \rightarrow X} \nabla f_X(X^i) \}$$

conv dénote l'espace convexe, où le gradient de  $f$  existe partout. Dans ce cas,  $f$  est différentiable en  $X$ . Les éléments de  $\partial f(X)$  sont appelés sous gradients. En tout point  $X$ , nous connaissons alors  $f(X)$  et un vecteur  $\text{gr} \in \partial f(X^i)$ . Ainsi, la relation (24) devient:

$$X^{k+1} = X^k - \theta_k \text{gr}_k / \|\text{gr}_k\| \quad \text{Où } \text{gr}_k \in \partial f(X^k) \text{ et } \theta_k > 0.$$

## B.2: Méthode du quasi-gradient stochastique:

Soit le problème stochastique suivant:

$$\text{OPT } F(X) \quad \text{Où } F(X) = E_{\xi} [f(X, \xi)]. \quad (25)$$

$X \in S$

$X$  est le vecteur des variables de décision.

$S$  l'espace des contraintes.

$\xi$  est le vecteur des paramètres aléatoires appartenant à l'espace probabiliste  $(\Omega, F, P)$ .  $\xi \in \mathbb{R}^d$ ,  $d \leq n$ .

$P$  la mesure de probabilité correspondante.

L'algorithme du quasi-gradient stochastique se formule alors comme suit:

$$X^{k+1} = \text{Proj}_{X \in \Omega} (X^k - \rho^k v^k) \quad (26)$$

$$\text{Proj}(X) = \text{Min}_{x \in \Omega} \|X^k - \rho^k v^k\|$$

$X^k$  est le point courant

$\rho^k$  est le pas de déplacement.

$v^k$  est la direction aléatoire qui peut être une estimation statistique du gradient.

La principale difficulté de la méthode du quasi-gradient stochastique est le comportement oscillatoire dû à l'aléa qu'elle présente quelquefois. Il est difficile d'affirmer alors, que l'algorithme a atteint le voisinage de la solution optimale. Il est préférable dans ce cas d'utiliser des méthodes itératives pour le choix du pas et de la direction de déplacement.

Un autre inconvénient de cet algorithme est la lenteur de sa convergence. Le taux de convergence peut être représenté comme suit:

$$\|x^* - x^k\| \approx C / \sqrt{s}$$

$x^*$  est la solution optimale.

$s$  le nombre d'observations du vecteur  $\xi$  des paramètres aléatoires.

$C$  est le taux de convergence.

Il est plus simple, pour de tels algorithmes d'atteindre un voisinage de la solution optimale, plutôt que la solution optimale elle même.

L'algorithme (26) nécessite le choix du pas  $\rho^k$  et le choix de la direction aléatoire  $v^k$  appelée quasi-gradient stochastique de la fonction.

Soient les fonctions  $F(X)$  et  $f(X, \xi)$  considérées en (25),  $\xi$  étant un vecteur aléatoire, elles le sont aussi. Quel est alors le gradient de la fonction aléatoire  $F(X)$ ?

a: Gradient d'une fonction aléatoire  $F(X)$ :

a.1: Cas où  $F(X)$  est différentiable:

1ère approche:

Elle consiste à utiliser les différences finies.

A l'itération  $k$  nous avons:

$$\phi^k = \sum_{i=1}^n \frac{f(X^k + \sigma_k e_i, \xi_{i,k}^1) - f(X^k, \xi_{i,k}^2)}{\sigma_k} e_i \quad (27)$$

ou encore les différences finies centrées:

$$\phi^k = \sum_{i=1}^n \frac{f(X^k + \sigma_k e_i, \xi_{i,k}^1) - f(X^k - \sigma_k e_i, \xi_{i,k}^2)}{2\sigma_k} e_i \quad (28)$$

Où  $\xi_{i,k}^1$  et  $\xi_{i,k}^2$  sont les  $i$ èmes composantes de deux vecteurs  $\xi$  à l'itération  $k$ .

$e_i$   $i$ ème composante du vecteur unitaire de base de  $\mathbb{R}^n$ .

$\sigma_k$  doit être choisi convenablement pour assurer la convergence en probabilité de (27) et (28).

Il est suffisant de prendre  $\sigma_k$  tel que:

la série  $\sum_{i=1}^{\infty} \frac{(\rho^k)^2}{\sigma_k^2}$  soit convergente

Dans le cas où  $\xi_{i,k}^1$  et  $\xi_{i,k}^2$  sont identiques, alors  $\forall \sigma_k \rightarrow 0$ , (27) et (28) convergent.

L'avantage de cette méthode est qu'elle atteint rapidement le voisinage de l'optimum, si  $\sigma_k$  est choisi convenablement.

Dans le cas des problèmes sans contraintes Gaivoronski [GAI 88] propose la démarche suivante pour le choix de  $\sigma_k$ :

1- Prendre  $\sigma_k$  tel que  $\sigma_k \approx \text{Max } 0.1 \parallel X^1 - X^2 \parallel$   
 $\{X^1 \in S, X^2 \in S\}$

2- Calculer  $X^{k+1} = \text{Proj}_S (X^k - \rho^k \phi^k)$ , où  $\phi^k$  est calculé à partir de (27) ou (28).

L'estimation du gradient est alors donnée par:

$$v^{k+1} = (1 - \alpha_k) v^k + \alpha_k \phi^k \quad (29)$$

où  $0 \leq \alpha_k \leq 1$ .

3- Prendre:

$$\alpha_{k+1} = \begin{cases} \alpha_k & \text{Si } v^k \geq \beta_1 \sigma_k \\ \beta_2 \sigma_k & \text{Sinon} \end{cases} \quad (30)$$

$\beta_1$  et  $\beta_2$  sont arbitrairement choisis au début du processus itératif.

Si  $\xi_{i,k}^1 \equiv \xi_{i,k}^2$ , ce processus converge. Sinon il produit une bonne approximation de solution.

L'inconvénient majeur de cette approximation pour les différences finies est que la variance du quasi-gradient augmente lorsque  $\sigma_k$  diminue. Ceci implique que l'algorithme converge plus lentement. Ce problème peut-être évité en augmentant le nombre des échantillons utilisés pour obtenir  $\phi^k$  lorsque l'on s'approche de la solution optimale. Lorsque la variance de  $\phi^k$  est très grande, ce dernier est parfois normalisé.

Un autre inconvénient de cette approximation est qu'elle nécessite, à chaque itération, l'évaluation de la fonction  $F(X)$  comme suit:

$n+1$  fois dans le cas (27)

$2n$  fois dans le cas (28)

où  $n$  est la dimension du vecteur  $X$ .

2ème approche:

Pour éviter de tels calculs, nous pouvons avoir recours à l'approche suivante qui s'inspire de la recherche aléatoire de la direction de descente développée dans le cas de l'optimisation déterministe. Nous aurons:

$$\phi^k = \sum_{i=1}^{M_k} \frac{f(X^k + \sigma_k h_i, \xi_{i,1}^k) - f(X^k, \xi_{i,2}^k)}{\sigma_k} h_i \quad (31)$$

$h_i$  ième composante du vecteur distribué uniformément sur la sphère unitaire.

$M_k$  Nombre de points aléatoires. Il sera déterminé suivant les facilités de calcul disponibles. Ainsi, il serait intéressant d'augmenter  $M_k$  tant que  $\sigma_k$  décroît.

$\sigma_k$  est le pas donné dans la recherche aléatoire. Il sera choisi de la même façon que dans l'approximation des différences finies.

#### a.2: Cas où $F(X)$ est non différentiable:

##### 1ère approche:

Lorsqu'il est possible d'obtenir le gradient de la fonction  $F(X)$  aux points fixés  $X$  et  $\xi$ , nous pourrions prendre comme gradient à l'étape  $k$  du processus considéré, le gradient de la fonction  $F(X)$ :

$$\nabla F(X) = f_X(X^k, \xi^k)$$

$$\text{D'où } \phi^k = f_X(X^k, \xi^k) \quad (32)$$

Où  $\xi^k$  est une observation du paramètre  $\xi$  à l'étape  $k$  de l'algorithme.

Si de plus, la détermination du gradient et l'observation des paramètres aléatoires ne nécessitent pas un grand effort de calcul, nous pourrions prendre la moyenne pour un nombre  $N$  donné d'observations.

Soit:

$$\phi^k = \frac{1}{N} \sum_{i=1}^N f_X(X^k, \xi^i) \quad (33)$$

Là aussi, un choix convenable de  $N$  s'impose. Une manière simple de procéder est proposée par Ermeliou [ERM 83].

1- Commencer avec  $N$  petit ( $N=1$ , par exemple).

2- Augmenter  $N$  si le point courant est dans un voisinage de l'optimum.

2ème approche:

Si  $F(X)$  est non différentiable, mais satisfait la condition suivante:

Soit un point  $Y \in S$  tel que:

$$F(X,r) = \int_S F(X+y) dH(X,r) \quad (34)$$

Soit la boule  $B$  de rayon  $r$  et de centre  $(0,0)$ .

$H(X,r)$  est la mesure de probabilité sur  $B(0,r)$ .

Nous supposons que  $H(X,r)$  a une densité non nulle en dehors de la boule  $B(0,r)$ .

$F(X,r)$  est différentiable et  $F(X,r)$  tend uniformément vers  $F(X)$  sur tout ensemble compact lorsque  $r$  tend vers 0.  $\lim_{r \rightarrow 0} F(X,r) = F(X)$

Il n'est plus nécessaire à présent de calculer l'intégrale (34). Il suffira de calculer  $\phi^k$  selon (27) ou (33), mais au point  $X^k + Y^k$  plutôt qu'au point  $X^k$ , où  $Y^k$  est une variable aléatoire distribuée selon  $H(y,r)$  et  $\text{opt } F(X) = \lim_{r \rightarrow 0} \text{opt } F(X,r)$  [ERM 73]. Dans ce cas (27) devient:

$$\phi^k = \sum_{i=1}^n \frac{f(X^k + Y^k + \sigma_k e_i, \xi_{i,k}^1) - f(X^k + Y^k, \xi_{i,k}^2)}{\sigma_k} e_i \quad (35)$$

La distribution  $H(X, r_k)$  est, à l'itération  $k$ , la distribution uniforme sur un cube à  $n$  dimension d'arête  $r_k$ . Si nous voulons obtenir une convergence en probabilité, nous choisirons  $r_k$  tel que:

$$\sigma_k / r_k \xrightarrow[k \rightarrow \infty]{} 0 \quad \text{et} \quad (r_k - \sigma_{k+1}) / \rho_k \xrightarrow[k \rightarrow \infty]{} 0$$

Cette approche qui consiste à rendre la fonction  $F(X)$  différentiable a l'avantage d'améliorer le comportement de  $F(X)$ . En effet, supposons que  $F(X)$  puisse s'écrire comme la somme de deux fonctions  $D(X)$  et  $R(X)$ .

$$F(X) = D(X) + R(X)$$

Où  $D(X)$  a un optimum global.

$R(X)$  a un comportement oscillatoire.

La différentiation peut aider à contourner l'effet oscillatoire, ce dernier pouvant orienter le processus d'optimisation vers un optimum local, loin de l'optimum global.

Ainsi, il est parfois utile de différencier la fonction objectif  $F(X)$  afin d'obtenir le gradient de  $f(X, \xi)$ . Dans ce cas, nous prenons:

1-  $r_k$  très grand pour les quelques premières itérations.

2- Diminuer  $r_k$  lorsque l'on s'approche de la solution optimale.

Tout ce qui a été dit concernant le choix du paramètre  $\sigma_k$  des différences finies, peut être appliqué pour le choix de  $r_k$ . La seule différence est que si  $r_k$  augmente, ceci n'entraîne pas l'augmentation de la variance de  $\phi^k$ . Par ailleurs il est préférable d'avoir  $\sigma_k < r_k$ .

### a.3: Cas où $F(X)$ est quelconque:

Toutes les approches pour l'approximation du gradient, proposées ci-dessus, sont basées sur l'information obtenue au point courant. Une façon d'utiliser les informations obtenues durant toutes les itérations précédentes du processus d'optimisation est la forme (29). Le vecteur  $v^k$ , ainsi obtenu, peut être utilisé en (26).

Les autres techniques qui utilisent les informations antérieures sont les suivantes:

$$v^k = \frac{1}{M_k} \sum_{i=k-M_k+1}^k \phi^i \quad (36)$$

$M_k$  étant la taille de la mémoire,  $M_k$  est fixé.

Ou bien:

$$v^k = A_k \phi^k \quad (37)$$

Où  $\phi^k$  est choisi selon l'une des techniques proposées ci dessus  
 $A_k$  matrice hessienne.

b: Choix du pas:

La théorie de la convergence [POL 66] suggère que toute série vérifiant:

$$\rho_k > 0, \sum_{k=1}^{\infty} \rho_k = \infty, \sum_{k=1}^{\infty} \rho_k^2 < \infty$$

Peut être utilisée comme séquence de pas.

L'inconvénient du choix du pas avant les itérations est que les informations accumulées durant le processus itératif, ne sont pas utilisées.

On donne alors à l'utilisateur la possibilité de changer les paramètres, en se basant sur le comportement de la fonction objectif  $F(X)$  qui sera estimée de la manière suivante:

$$\text{Soit } \hat{F}(X) = \frac{1}{L} \sum_{i=1}^L f(X, \xi^i) \quad (38)$$

$s$  est le nombre d'observations.

$\hat{F}(X)$  étant un estimateur de  $F(X)$ .

Une implantation possible d'une telle approche, appelée approche interactive [GAI 88], peut être la suivante:

1-Choisir le pas constant pour un certain nombre d'itérations de départ (10 à 20).  $\hat{F}(X^k)$  et  $X^k$  sont calculés à chaque itération.

2-Décider d'une valeur du pas lorsque l'un des trois cas suivants se présente:



i- Lorsque  $\hat{F}(X^k)$  et les composantes de  $X^k$  présentent des sauts aléatoires, le pas courant est alors trop grand, nous devons le diminuer.

ii- Lorsque  $\hat{F}(X^k)$  et les composantes de  $X^k$  ont un comportement régulier (une tendance), le pas courant est convenable. Nous changeons le pas dès que  $\hat{F}(X^k)$  ou/et les composantes de  $X^k$  commencent à osciller.

iii- Lorsque  $\hat{F}(X^k)$  et les composantes de  $X^k$  changent très lentement, Le pas est trop petit, il est nécessaire de l'augmenter.

L'inconvénient dans cette approche est que l'utilisateur perd trop de temps lorsque l'ordinateur effectue le calcul pour une observation.

Pour cela, une autre approche est proposée [CHE 82], l'approche adaptative automatique, dans laquelle le choix du pas est aussi basé sur les informations obtenues sur tous ou quelques points  $X^i$ ,  $i=1, k$ , du processus d'optimisation.

Soit  $\varphi(\bar{X}^k, u^k)$ . Où  $\varphi$  est l'estimateur des performances de l'algorithme.

$\bar{X}^k$  est la séquence  $(X^1, X^2, \dots, X^k)$  des points obtenus durant le processus jusqu'à l'itération  $k$ .

$u^k$  l'ensemble des paramètres utilisés dans l'estimation.

Ces mesures de performance doivent en général vérifier une ou plusieurs des propriétés suivantes:

1-  $\varphi(\bar{X}^k, u^k)$  augmente lorsque l'algorithme montre un comportement régulier, c'est à dire que  $\hat{F}(X^k)$  décroît ou que  $X^k$  montre une tendance distincte.

2-  $\varphi(\bar{X}^k, u^k)$  diminue ou change de signe si  $\hat{F}(X^k)$  ne change pas de signe ou si  $X^k$  commence à osciller.

3-  $\varphi(\bar{X}^k, u^k)$  augmente loin de la solution et diminue au voisinage immédiat de la solution.

La valeur de la fonction objectif et de son gradient peuvent être utilisées comme indicateurs pour l'évaluation de la performance. Le principal avantage est qu'elles n'utilisent pas un grand nombre d'observations (souvent une seule). Bien que l'estimation soit pauvre au début des itérations, elle devient plus précise avec l'évolution du processus d'optimisation. Ainsi (38) peut être un estimateur et aussi un cas particulier de (39):

$$\hat{F}(X^{k+1}) = (1-\gamma^k) \hat{F}(X^k) + \gamma^k f(X^k, \xi^k) \quad (39)$$

Toute observation  $\mu^k$  avec la propriété (40) peut être utilisée en (39) à la place de  $f(X^k, \xi^k)$ :

$$E(\mu^k / X^1, X^2, \dots, X^{k-1}) = F(X^k) + d^k \quad \text{avec} \quad d^k \rightarrow 0 \quad (40)$$

Ainsi, nous aurons:

$$\hat{F}(X^{k+1}, \mu^k) = (1-\gamma^k) \hat{F}^k + \gamma^k E(\mu^k / X^1, X^2, \dots, X^k)$$

Nous pouvons aussi utiliser la valeur moyenne de  $\hat{F}(X^k)$  comme estimation:

$$\hat{F}(X^k) = \frac{1}{M_k} \sum_{i=1}^{M_k} f(X^i, \xi^i) \quad (41)$$

Parmi les différentes règles pour le choix du pas, les suivantes sont proposées:

b.1: ratio de la fonction estimé à la longueur du chemin de l'algorithme:

On choisit  $\rho_0, \alpha_1, \alpha_2$  ( $\alpha_1 > 0, \alpha_2 > 0$ ),  
 Une séquence  $M_k$ ,  
 Un entier  $\hat{M}$ .

Après chaque  $\hat{M}$  itérations, la valeur du pas est revue comme suit:

$$1- \text{ Calculer } \varphi^1(\bar{X}^k, u^k) = \frac{\hat{F}(X^{k-M_k}) - \hat{F}(X^k)}{\text{long}(k, M_k)} \quad (42)$$

$$\text{long}(k, M_k) = \sum_{i=k-M_k}^{k-1} \|X^{i+1} - X^i\| \quad (43)$$

Où  $\text{long}(k, M^k)$  est la longueur du chemin pris par l'algorithme.

2-  $\rho^{k+1}$  est calculé tel que:

$$\rho^{k+1} = \begin{cases} \alpha_1 \rho_k & \text{si } \varphi^1(\bar{X}^k, u^k) \leq \alpha_2 \\ \rho_k & \text{sinon} \end{cases} \quad (44)$$

Les valeurs recommandées pour  $\alpha_1$  et  $\alpha_2$  sont:

$\alpha_1 \in [0.5, 0.9]$  et  $\alpha_2 \in [0.005, 0.1]$ .

b.2: Utilisation de l'estimation du gradient:

On prend  $\varphi^2 = \phi^k$  au lieu de  $\varphi^1(\bar{X}^k, u^k)$  dans (44).  $\phi^k$  est l'une des estimations du gradient discutée en (§a).

**b.3: Ratio du chemin progressif:**

On choisit une séquence  $M$ .

Soient  $X^k$  le point à l'itération  $k$  et  $X^{k-M}$  le point à l'itération  $k-M$ .

$\|X^{k-M} - X^k\|$  représente la progression de l'algorithme de l'itération  $k-M$  à l'itération  $k$ .

Si le pas est gardé constant, l'algorithme commence à osciller après avoir atteint un voisinage du point optimal. Ainsi, le chemin total parcouru par l'algorithme, entre l'itération  $k-M$  et l'itération  $k$ , commence à augmenter comparé à la distance entre le point  $X^{k-M}$  et le point  $X^k$ . On pourra utiliser alors comme mesure de performance ce qui suit:

$$\varphi^3(\bar{X}^k, u^k) = \frac{\|X^{k-M} - X^k\|}{\sum_{i=k-M}^{k-1} \|X^{i+1} - X^i\|} \quad (45)$$

Dans (44) nous remplaçons alors  $\varphi^1$  par  $\varphi^3$ .

**b.4: ratio des directions:**

Le choix du pas peut aussi être basé sur le produit scalaire des directions de descente adjacentes.

En effet, si  $(\phi^{k-1}, \phi^k) > 0$ , ceci peut impliquer, ou être un signe, que le comportement de la fonction stochastique est régulier. La fonction diminue alors que le pas augmente.

Nous choisissons comme règle pour le changement du pas ce qui suit:

$$\rho = \begin{cases} \rho^k & \text{si } -\alpha_1 \leq (\phi^{k-1}, \phi^k) \leq \alpha_1 \\ \alpha_2 \rho^k & \text{si } (\phi^{k-1}, \phi^k) > \alpha_1 \\ \alpha_3 \rho^k & \text{si } (\phi^{k-1}, \phi^k) < -\alpha_1 \end{cases}$$

Les valeurs recommandées pour  $\alpha_1$ ,  $\alpha_2$  et  $\alpha_3$  sont:

$$0.4 \leq \alpha_1 \leq 0.8$$

$$1 < \alpha_2 \leq 1.3$$

$$0.7 \leq \alpha_3 < 1$$

### 2.2.2: CONCLUSION:

Dans la méthode du quasi-gradient stochastique et la méthode des approximations techniques, qui sont des méthodes actives, les problèmes résolus sont déterministes et la solution obtenue est ponctuelle.

La méthode des approximations techniques présente quelques difficultés, du point de vue précision, car il n'est pas toujours facile de pouvoir trouver les meilleures approximations. Nous pouvons obtenir ces dernières en augmentant le nombre de sous ensembles  $\Omega_c \in \mathcal{E}$ , mais ceci crée des complexités du point de vue numérique.

La méthode du quasi-gradient stochastique présente quant à elle, d'autres difficultés, encore plus grandes.

Les premières sont celles dues au comportement oscillatoire de la méthode, de part sa nature probabiliste. Les autres sont celles rencontrées lors du choix du pas et de la direction de déplacement. Ce choix étant déjà difficile dans le cas déterministe [ZOW 85], la nature probabiliste de la méthode ne fait qu'accentuer sa complexité.

Pour ces deux méthodes, l'effet probabiliste est éliminé aux dépend d'une augmentation de la complexité numérique. Lorsque le coût du calcul numérique s'avère être trop élevé, ne serait-il pas intéressant de présenter au décideur des informations sur le comportement de l'optimum (loi de probabilité, moyenne, variance), plutôt que de lui donner une valeur de l'optimum lui même.

**CHAPITRE 3:**

**METHODE DE RESOLUTION ET RESULTATS**

### 3.1: APPROCHE ETUDIEE:

En nous inspirant des travaux effectués en programmation stochastique linéaire [TIN 55], nous nous proposons de réaliser un travail similaire en programmation stochastique non linéaire.

Dans l'approche passive (wait and see), nous nous proposons d'observer la réalisation des variables aléatoires, puis de résoudre un ensemble de problèmes déterministes afin d'analyser le comportement de l'optimum. Nous nous proposons aussi de rechercher un lien, s'il existe, entre les lois de probabilités des paramètres aléatoires et la loi de probabilités de l'optimum. Notre travail sera basé sur les deux questions suivantes:

1- Existe-t-il un rapport entre les lois de probabilités suivies par les paramètres et celle de l'optimum? Plus encore, les lois des paramètres définissent-elles celle de l'optimum. Le nombre de paramètres aléatoires influe-t-il sur le comportement de l'optimum.

2- La loi de l'optimum est-elle indépendante de celle des paramètres aléatoires? Y a-t-il une loi générale gérant son comportement?

Le champ d'étude étant très vaste, nous nous limiterons au cas suivant:

- Fonction objectif différentiable.
- Problèmes non contraints.
- Paramètres aléatoires indépendants.
- Lois des paramètres aléatoires identiques.

### 3.2: METHODOLOGIE DE TRAVAIL:

Soit le problème de programmation stochastique suivant:

$$\text{Opt } F(X) = E_{\xi} [f(X, \xi)]$$

Nous considérons l'ensemble des observations de  $\xi$ ,  $(\xi^1, \xi^2, \dots, \xi^L)$ .

Nous résoudrons alors  $L$  problèmes de programmation non linéaire déterministe.

Nous considérons le cas de minimisation:  $\text{Min } f(X, \xi^i), i=\overline{1, L}$

Nous résoudrons, pour chaque fonction donnée,  $L$  problèmes,  $L$  variant de 10 à 200 avec un pas de 10.

Au total, pour chaque fonction donnée, nous procéderons à 2100 optimisations.

Nous générerons alors 20 échantillons pour l'optimum dont la taille varierait aussi de 10 à 200 avec un pas de 10.

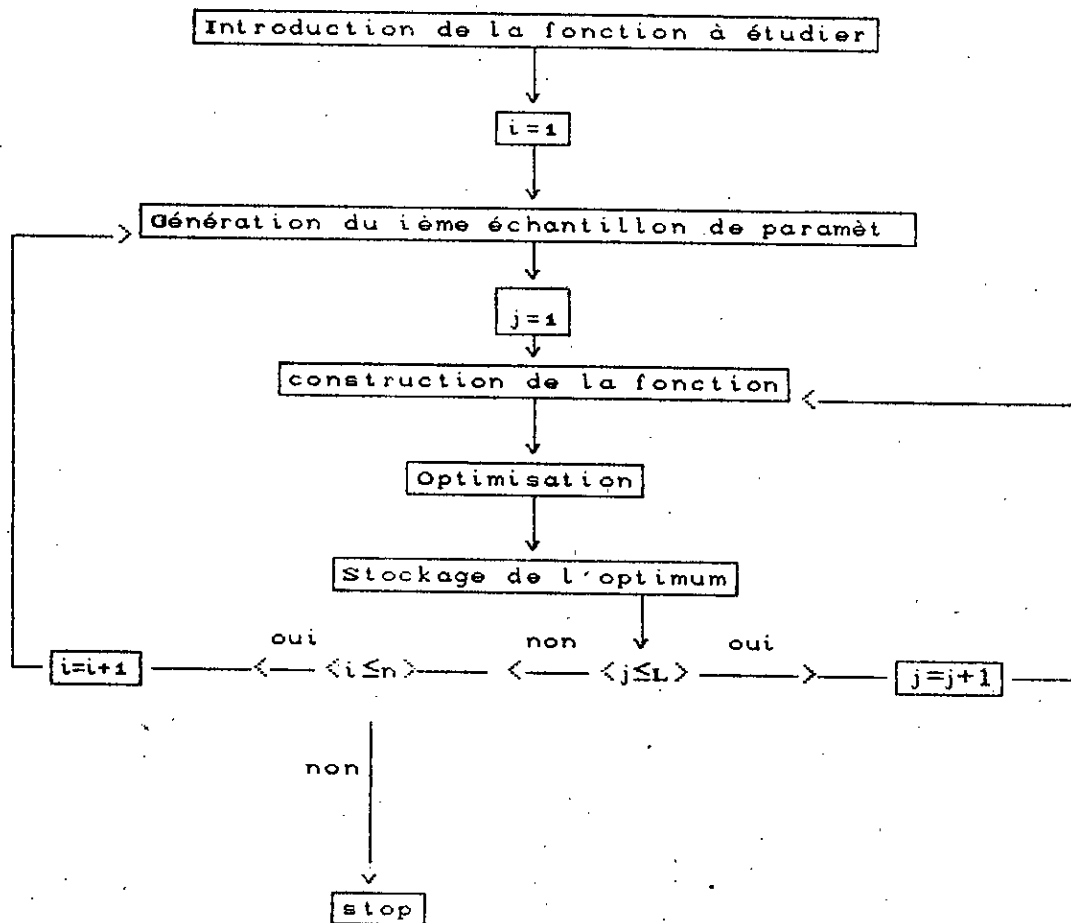
Les fonctions étudiées sont prises dans la littérature [Annexe 1]. Nous avons choisi 15 fonctions tests. Nous faisons remarquer qu'elles sont déterministes et que nous avons nous même apporté les perturbations aléatoires. Nous avons construit 15 problèmes [Annexe 1] à étudier comme suit:

- Prendre, dans la littérature, une fonction  $f(X) \quad X \in \mathbb{R}^n$ .
- Soit la composante  $x_i$  du vecteur  $X$  et soit  $a_i$  le paramètre associé à  $x_i$  dans la fonction  $f(X)$ . Affecter l'aléa au paramètre de  $x_i$  soit alors  $\xi_i$  le nouveau paramètre associé à  $x_i$ .  $i$  variant de 1 à  $n$

Un problème peut contenir, au minimum,  $n$  fonctions à étudier

Notre processus de calcul est représenté par le diagramme suivant:





ce processus comprend deux parties:

-La première relative à la génération des paramètres aléatoires:  
Nous générons des échantillons de paramètres aléatoires, dont la taille varie de 10 à 200 avec un pas de 10, selon les lois de probabilité suivantes:

- Loi Normale
- Loi Gamma
- Loi Exponentielle
- Loi Poisson
- Loi Uniforme
- Loi Binomiale

-la seconde est relative à l'optimisation des fonctions:  
Les échantillons de fonctions, ainsi obtenus par la génération des nombres aléatoires, sont optimisés afin de générer des échantillons de l'optimum dont la taille varie aussi de 10 à 200.

Le langage informatique utilisé est le FORTRAN 77. Nous avons élaboré les programmes principaux des générations des nombres aléatoires et de l'optimisation ainsi que la routine concernant la construction de la fonction. Les autres routines ont été prises dans la littérature [NUM 87]. Ce sont des routines déjà étudiées et testées.

L'exécution des programmes s'est effectuée sur DEC-VAX 11/750 ou DEC-MICROVAX II. Il est cependant possible de les exécuter sur Micro-Ordinateur.

### 3.3: METHODES DE RESOLUTION ET RESULTATS:

Les méthodes de résolution utilisées sont les méthodes d'optimisation non linéaire sans contraintes.

Ces méthodes sont itératives et formulées, généralement, comme suit (cf chap1):

$x^0$  point initial

$\epsilon$  choisi tel que  $0 < \epsilon < 1$

$$x^{k+1} = x^k + s_k d^k$$

L'implémentation de ces méthodes conduit à résoudre un certain nombre de problèmes numériques.

#### 3.3.1: CONSIDERATIONS NUMERIQUES LIEES A LEUR IMPLEMENTATION:

L'implémentation de telles méthodes impliquent la nécessité de tenir compte d'un certain nombre de problèmes numérique car il sont la source d'erreurs principale à savoir:

- Le choix du pas de progression.
- Le critère d'arrêt.
- Le choix du point initial.

**A: LE PAS DE PROGRESSION:**

Soit  $\theta_k$  tel que  $\psi(\theta_k) = \min_{\theta \geq 0} \psi(\theta)$

Il est déterminé numériquement, il est sujet donc à différentes sources d'erreurs numériques:

- Erreurs que peuvent engendrer les méthodes d'optimisation unidimensionnelles choisies (problème de convergence).
- Erreurs dues aux calculs numériques de l'ordinateur.
- l'instabilité numérique.

**B: LE CRITERE D'ARRET:**

Parmi les différents critères d'arrêt, nous citons les suivants:

- $\|\nabla f(X^k)\| \leq \epsilon$
- $\alpha \|X^{k+1} - X^k\| \leq \epsilon$
- $\frac{\|f(X^{k+1}) - f(X^k)\|}{\|f(X^k)\|} \leq \epsilon$

Quel que soit le critère choisi, nous pouvons commettre deux erreurs possibles:

- Le processus peut être arrêté trop tôt, nous risquons de ne pas converger.
- Le processus peut être arrêté trop tard, nous aurions fait des itérations inutiles.

Pour cela, lors des implémentations le nombre d'itérations maximums est toujours fixé.

**C: CHOIX DU POINT INITIAL:**

Le mauvais choix du point initial peut orienter le processus d'optimisation vers un point qui n'est pas l'optimum.

Soit par exemple les fonctions suivantes:

$$f^l(x) = U_1^2 + U_2^2 + U_3^2$$

$$U_1 = a_1 - x_1 \cdot (1 - x_2)$$

$$U_2 = 2.25 - x_1 \cdot (1 - x_2^2)$$

$$U_3 = 2.625 - x_1 \cdot (1 - x_2^3)$$

Point initial		$a_1$	Itérations	Optimum
2	2	1.9319	288	0.031
4283.661	0.999	1.9319	86	1.064
2	2	0.4073	335	0.485
4958.962	0.998	0.4073	4	0.448
2	2	1.3719	6	0.004
4430.537	0.998	1.3719	400	0.365
2	2	3.1033	5	0.013
4764.902	0.998	3.1033	5	4.469
2	2	1.4418	6	0.008
5365.045	0.998	1.4418	6	0.395
2	2	0.5309	134	0.337
4788.980	0.998	0.5309	17	0.337
2	2	1.882	5	0.026
4473.214	0.998	1.882	400	0.976

Nous remarquons que le changement du point initial a deux effets soit:

- l'optimum change (problème de convergence).
- ou bien l'optimum ne change pas et le nombre d'itérations requises augmente (problème de temps de calcul).

**D: ERREUR DE LA SOLUTION:**

En plus de ces erreurs liées aux méthodes d'optimisations il y a aussi l'erreur de la solution. Car le fait d'avoir obtenu une solution informatiquement satisfaisante ne garantit en rien de sa précision; Tout dépend d'un certain domaine d'incertitude, que l'on peut définir [LAP 82] et qui détermine sa précision.

**3.3.2: METHODE DES GRADIENTS CONJUGUES:**

Nous avons choisi comme méthode de résolution, la méthode des gradients conjugués, pour sa bonne convergence.

Cette méthode est intéressante car elle nécessite le stockage de très peu d'informations (trois vecteurs de dimension  $n$ ), contrairement à la méthode de Newton qui nécessite le stockage d'une matrice ( $n \times n$ ).

La convergence globale de cette méthode n'est assurée que si nous procédons à une réinitialisation périodique ( toutes les  $n$  itérations par exemple).

**Définition:**

Soit  $H_{n \times n}$  une matrice carrée symétrique. Soient  $d^1, d^2, \dots, d^k$   $k$  vecteurs. Il seront appelés conjugués par rapport à  $H$ , ou simplement conjugués, si et seulement s'ils sont linéairement indépendants et si

$$(d^i)^T H d^j = 0 \text{ pour } i \neq j$$

**Théorème:**

soit  $\{d^i\}_{i=0}^{n-1}$  un ensemble de vecteurs non nuls et conjugués par rapport à  $H$ . Pour tout point  $X^0 \in \mathbb{R}^n$ , la suite  $\{X^k\}$  générée par:

$$X^{k+1} = X^k - \vartheta_k d^k \quad k \geq 0$$

$$\vartheta_k = \frac{-g^k d^k}{(d^k)^T H d^k}, \quad g^k = \nabla f(X^k)$$

Converge vers l'unique solution  $X^*$  après  $n$  étapes.

Principe:

La méthode est due à Fletcher et Reeves (1964). Elle consiste en la recherche de l'optimum le long de  $n$  directions conjuguées, la direction initiale étant le gradient de la fonction objectif.

**A: ALGORITHME:**

Pas 0: Initialisation

Soient  $\varepsilon$  un scalaire

$X^1$  point initial et  $Y^1 = X^1$

$d^1 = -\nabla f(Y^1)$  la direction initiale

$k=j=1$

Aller au pas 1.

Pas 1:

Si  $\|\nabla f(Y^j)\| < \varepsilon$  STOP (un critère d'arrêt possible).

Sinon soit  $\psi(\vartheta_j) = \min_{\vartheta \geq 0} \psi(\vartheta)$  où  $\psi(\vartheta) = f(Y^j + \vartheta d^j)$

$Y^{j+1} = Y^j + \vartheta_j d^j$  si  $j < n$  aller au pas 2  
sinon aller au pas 3

Pas 2:

$$d^{j+1} = -\nabla f(Y^{j+1}) + \alpha^j d^j \quad \alpha^j = \frac{\|\nabla f(Y^{j+1})\|^2}{\|\nabla f(Y^j)\|^2}$$

faire  $j=j+1$  aller au pas 1.

**Pas 3: Réinitialisation**

$$j=1$$

$$k=k+1$$

$$Y^j = Y^{n+1}$$

$$X^{k+1} = Y^{n+1}$$

$$d^j = -\nabla f(Y^j)$$

Aller au pas 1.

Au pas 3 il est procédé à une réinitialisation, si au bout de  $n$  étapes la méthode n'aboutit pas à l'optimum.

**B: CONSIDERATIONS LIEES A SON IMPLEMENTATION:**

A chaque itération, cette méthode nécessite le calcul de la dérivée de la fonction. Les expressions de nos fonctions ainsi que leurs dérivées sont connues. Nous tenterons cependant, d'établir un processus de calcul automatique. A cet effet, les dérivées seront calculées analytiquement comme suit:

Soit  $f(X)$  la fonction à optimiser, le gradient de  $f(X)$  sera le vecteur  $\nabla f(X)$  tel que:

$$\nabla f^T = (\partial f / \partial x_1, \partial f / \partial x_2, \dots, \partial f / \partial x_n)$$

$$\frac{\partial f}{\partial x_i}(X) = \frac{f(x_1, \dots, x_{i-1}, x_i+h, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_{i-1}, x_i-h, x_{i+1}, \dots, x_n)}{2h}$$

$h$  est le pas de dérivation.

Le problème du choix du pas est très complexe. La question qui est posée est : qu'elle est la valeur de  $h$  qui donne la bonne approximation?

Pour tout nos calculs, nous avons décidé de fixer la valeur de  $h$ . Afin d'analyser les risques de ce choix, nous avons procédé à quelques tests [Annexe 2].

A cet effet, nous avons utilisé 8 fonctions. Certaines ont été prises dans la littérature, d'autres ont été construites et testées préalablement sur le logiciel GINO [GIN 84].

Pour toutes ces fonctions, l'optimum obtenu par la méthode des gradients conjugués, pour un pas  $h$  donné, est comparé à celui donné par la littérature ou obtenu par le logiciel.

Pour chaque fonction, les tests sont réalisés pour différentes valeurs de  $h$  comprises entre  $10^{-2}$  et  $10^{-7}$ .

Nous avons obtenu les résultats suivants:

fonction	pas $h$
$f_1$	$10^{-4}$
$f_2$	$10^{-3}$
$f_3$	$10^{-4}$
$f_4$	$10^{-3} \dots 10^{-7}$

fonction	pas $h$
$f_5$	dépassement
$f_6$	$10^{-6}$
$f_7$	$10^{-7}$
$f_8$	$10^{-3} \dots 10^{-7}$

Si nous fixons le pas  $h$  tel que  $h=10^{-3}$ , il donne une assez bonne approximation [Annexe 2]. Nous avons fixé donc la valeur de  $h$  à  $10^{-3}$  pour nos calculs.

L'inconvénient est que la nature de la fonction peut changer, d'une observation à une autre des paramètres aléatoires. Dans ce cas, le pas fixé peut ne plus être adéquat pour le calcul de la dérivée de cette fonction.

En fait,  $h$  varie avec le changement des paramètres aléatoires. Si nous prenons ceci en considération, nous aurons alors à optimiser le problème suivant:

$$\text{Opt } f(X, \xi, h)$$



Où  $\xi$  est le vecteur des paramètres aléatoires  $\xi \in \mathbb{R}^d$ .

Et  $h$  est le pas de dérivation à déterminer pour chaque observation de  $\xi$ .

Ce cas reviendrait, dans le cadre de notre travail, à étudier l'influence de  $h$  sur le comportement de l'optimum, ce qui ne nous permettra pas de répondre à notre objectif principal à savoir connaître le lien entre les paramètres aléatoires de la fonction et l'optimum.

De plus cette éventualité implique l'introduction d'un nouveau paramètre  $h$  qui est dépendant du vecteur  $\xi$ . Or le cas de dépendance a été écarté de notre étude en raison de sa complexité [RUB 81].

Une autre possibilité serait de déterminer  $h$  statistiquement. Ceci consiste, pour une fonction donnée, à analyser le comportement de  $h$  en fonction de la variation des paramètres aléatoires et à déduire une valeur moyenne. Mais ce cas de figure sera plus complexe, car pour chaque fonction étudiée, il sera nécessaire de déterminer  $h$ , ce qui demandera un temps de calcul extrêmement grand.

Nous préférons alors l'éventualité de fixer  $h$ . C'est la solution la plus facile, néanmoins la plus dangereuse par les erreurs numériques qu'elle engendre. Soit la fonction  $f(X, \xi)$  et soit l'échantillon donné de fonctions  $f(X, \xi^i)$   $i = \overline{1, L}$ .

Soit  $h$  fixé. Dans le pire des cas,  $h$  conviendra à une fonction de l'échantillon. Pour les  $L-1$  fonctions restantes, il sera inadéquat. L'erreur se répartira sur tout l'ensemble des fonctions  $f(X, \xi^i)$   $i = \overline{1, L-1}$ . La nature mathématique de l'échantillon sera alors changée. Cependant nous ignorons l'impact de ces erreurs sur les résultats statistiques, c'est ce que nous nous proposons de vérifier..

#### C: MISE EN OEUVRE ET RESULTATS:

L'analyse statistique est réalisée sur Micro-Ordinateur à l'aide du logiciel UNIFIT [UNI 85].

La première étape de l'analyse a consisté à tracer les histogrammes des échantillons. Elle a permis de sélectionner un ensemble de lois pouvant être ajustées à ces histogrammes. Par exemple, pour les tracés qui présentent une allure de demi cloche, donc un pic et une décroissance, nous sélectionnons des lois pour les tests d'ajustement telles que:

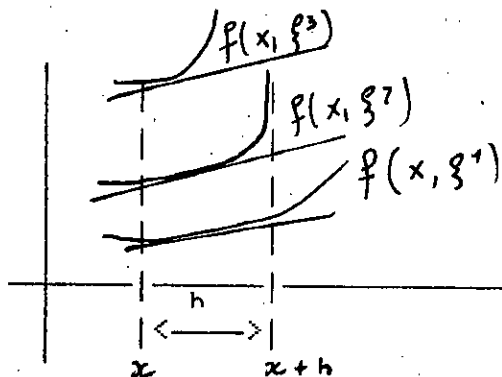
- Loi Exponentielle.
- Loi Weibull.
- Loi Normale.

Comme test d'ajustement, nous avons choisi le test du Chi-deux au niveau  $\alpha=5\%$  [Annexe 3].

Cette analyse a montré qu'il n'existe pas une cohérence dans le comportement des optimums. Ils ne présentent aucune régularité. Cependant quelques échantillons ont eu pour loi ajustée la loi de Weibull et parfois la loi exponentielle. Par contre la loi Normale a été très rarement ajustée [Annexe 3].

Nous pensons que cette incohérence est due aux erreurs induites par le pas de dérivation  $h$ . Nous avons supposé l'homogénéité de l'erreur sur tout l'échantillon. Or une toute petite variation des paramètres peut avoir un grand impact sur le comportement de l'échantillon.

Soit la fonction  $f(X, \xi)$  et soient  $\xi^1$ ,  $\xi^2$  et  $\xi^3$  trois observations du vecteur  $\xi$ . Nous avons alors les fonctions  $f(X, \xi^1)$ ,  $f(X, \xi^2)$  et  $f(X, \xi^3)$ . Le cas de figure suivant peut alors se présenter:



Il est évident alors que si  $h$  est adéquat pour la fonction  $f(X, \xi^1)$  il ne l'est pas forcément pour les deux autres et en particulier pour  $f(X, \xi^3)$ .

#### D: SYNTHÈSE ET CONCLUSION:

Ce genre de situation nous a conduit à intervenir à chaque fois au cours du processus de calcul afin de changer le pas  $h$ .

Pour une fonction donnée et pour une loi donnée des paramètres, le processus procède à 2100 optimisations. Si une optimisation durait une seconde, le total du temps d'exécution serait de 2100 secondes soit 35 minutes. Pour toutes les lois considérées le processus se déroule durant 3 heures. Ceci n'est qu'une approximation car nous avons omis de compter le temps d'exécution pour la génération des paramètres aléatoires et nous avons surestimé le temps d'exécution d'une optimisation, bien que pour certaines fonctions assez compliquées ce temps se trouve justifié.

C'est un processus de calcul relativement long. Les interruptions dues au changement du pas  $h$  augmentent ce temps. De plus il nécessite la présence de l'utilisateur afin qu'il puisse intervenir à tout instant car ces interruptions ne sont pas prévisibles.

Considérons, par exemple, le problème 1 [Annexe 2]:

Le cas où les paramètres suivent la loi Poisson:

-Nous démarrons le processus avec un échantillon de taille 10 et le pas  $h=10^{-3}$ .

-Le processus est interrompu pour un premier changement, il est poursuivi à partir de l'échantillon de taille 70 avec un pas  $h=10^{-2}$ .

-Une deuxième interruption est effectuée pour l'échantillon de taille 80 et pour un pas  $h=10^{-3}$ .

-Le processus est accompli jusqu'à l'échantillon de taille 200 avec un pas  $h=10^{-3}$ .

Pour cette exemple nous avons eu deux interruptions. C'est le cas de figure le plus simple, le nombre d'interruptions pouvant parfois atteindre 4.

L'utilisateur, dans ce cas, essaye empiriquement différentes valeurs de  $h$ . Il ne dispose d'aucun moyen pour apprécier, à l'avance, la convenance d'une valeur donnée.

Lors de nos calculs, nous avons fait un choix empirique de  $h$ . Dès que le pas n'est plus convenable, il est réduit par division de 10. Chaque nouveau pas est calculé ainsi:  $h' = h/10$

Nous aurions tout aussi bien pu choisir une autre manière pour la recherche empirique, par exemple:  $h' = h \cdot \epsilon$  où  $0 < \epsilon < 1$

Le plus important alors est d'automatiser ce processus car les erreurs générées par le mauvais choix du pas ont probablement conduit l'analyse statistique vers des résultats incohérents. Pour cela nous avons penser à deux possibilités, soit :

- automatiser le choix du pas de dérivation.
- opter pour une autre méthode d'optimisation ne calculant pas de dérivées.

Les difficultés numériques rencontrées nous poussé à choisir la seconde possibilité, non pas que ces méthodes soient plus performantes (ce qui n'est pas le cas), mais leur implémentation, à ce stade de notre étude, nous est plus facile, le processus d'optimisation est déjà établi nous n'avons qu'à procéder au changement de l'algorithme.

### 3.3.3: METHODE DE HOOK ET JEEVES:

Nous avons choisi la méthode de HOOK et JEEVES, que nous même avons programmée. Le processus de calcul est alors réalisé avec cette méthode.

Cette méthode est une amélioration de la méthode précédente. Elle peut être appliquée au cas de fonctions non différentiables. Elle se compose de deux phases:

- La phase d'exploration.
- La phase d'optimisation.

#### A: ALGORITHME:

(cas de minimisation):

##### Pas 0: Initialisation

Soient  $\epsilon$  un scalaire

$X^1$  point initial

$Y^1 = X^1$

$k=j=1$

Aller au pas 1.

##### Pas 1: Exploration

Soit  $\psi(\theta_j) = \min_{\theta \geq 0} \psi(\theta)$  où  $\psi(\theta) = f(Y^j + \theta d^j)$

$Y^{j+1} = Y^j + \theta_j d^j$

Si  $j < n$  faire  $j=j+1$  aller au pas 1.

Sinon : si  $j=n$   $k=k+1$   $X^{k+1} = Y^{n+1}$  aller au pas 2

si  $\|X^{k+1} - X^k\| \leq \epsilon$  STOP

##### Pas 2: Minimisation

$j=1$

$d = X^{k+1} - X^k$

$\psi(\theta^*) = \min_{\theta \geq 0} \psi(\theta)$  où  $\psi(\theta) = f(X^{k+1} - \theta d)$

$Y^1 = X^{k+1} + \theta^* d$

Aller au pas 1.

**B: CONSIDERATIONS LIEES A SON IMPLEMENTATION:**

La méthode de Hook et Jeeves pose les problèmes généraux liés à toute méthode d'optimisation non linéaire, à savoir:

- 1- Le choix du pas de progression.
- 2- Le choix du critère d'arrêt.
- 3- Le choix du point initial.

De plus elle est moins performante que la méthode des gradients conjugués. Lors de nos tests nous avons remarquer qu'elle est beaucoup plus lente. A titre indicatif nous présentons le tableau suivant:

	Temps de calcul en milli scd			nombre d'itérations		
	f1	f2	f3	f1	f2	f3
Méthode gradients conjugués	233	119	1	8	2	3
Méthode Hook et Jeeves	234	234	7.8	5	3	2

Tableau comparatif pour une fonction donnée.

La méthode de Hook et Jeeves nécessite le choix d'un point initial. C'est un point que nous avons choisi empiriquement. En fait, le risque que nous encourons, dans ce cas, est d'obtenir des optimums locaux ou d'augmenter le nombre d'itérations (temps d'exécution).

Le nombre d'itérations maximum de notre processus de calcul est de 400. toutes les solutions obtenus sont soit:

- l'optimum exact (lorsque la méthode converge) car nous ne pouvons ignorer l'impact du choix du point initial.
- ou la meilleur solution obtenue au bout de 400 itérations.

**C: MISE EN OUVRE ET RESULTATS:**

En choisissant la méthode de Hook et Jeeves, nous avons ainsi éliminé une source d'erreurs.

Nous avons entrepris ces tests dans le but d'expliquer le comportement aléatoire de l'optimum et éventuellement trouver une loi probabiliste générale gérant son comportement.

Nous avons suivi le même schéma d'analyse statistique, grâce au logiciel UNIFIT:

- Tracé des histogrammes, permettant de sélectionner un certain nombre de lois pouvant être étudiées dans le but d'être ajustées.
- Test du CHI deux qui permet de déterminer la loi ajustée.

D'après l'analyse statistique, nous regroupons les résultats en trois catégories:

- 1: Problèmes dont les tests n'ont donné aucun résultat.
- 2: Problèmes dont les tests ont abouti à un résultat sans lois ajustées.
- 3: Problèmes dont les tests ont donné des lois ajustées.

Parmi les problèmes concernés par la première catégorie nous avons [annexe 1]:

- Problème 4
- Problème 5
- Problème 9
- Problème 12

L'analyse de ces fonctions s'est arrêtée au niveau de l'optimisation. En fait, pour ces fonctions, quelque soit:

- La loi des paramètres.
- La taille des paramètres.
- La taille des échantillons.

L'optimum a toujours atteint la même valeur, par exemple:

- Les problème 4 et 12: l'optimum est toujours égal à zéro.
- Le problème 5: L'optimum a toujours atteint la valeur -2.
- Le problème 9: Numériquement, difficile à résoudre car l'optimum est non borné.

L'analyse statistique, dans ces cas là, est inutile.

En ce qui concerne la seconde catégorie, le tracé des histogrammes ne nous pas permis de sélectionner des lois pouvant être étudiées en tests d'ajustement. Parmi ces problèmes nous citons [Annexe 4]:

- Problème 2
- Problème 3
- Problème 6
- Problème 7
- Problème 11
- Problème 13
- Problème 14
- Problème 15

Dans ce cas, nous n'avons pas ajusté de loi de probabilité connue, mais nous avons fait la constatation suivante:

- pour une loi des paramètres donnée,
- Pour une taille des paramètres donnée, et enfin
- quelle que soit la taille des échantillons.

l'allure des histogrammes reste stable. Chaque problème, pour une loi donnée, a un comportement spécifique et des allures d'histogrammes qui lui sont propres.

La troisième catégorie concerne les problèmes pour lesquels nous avons pu ajuster une loi de probabilité connue. Ces problèmes sont les suivants [Annexe 1]:

- Problème 1
- Problème 8
- Problème 10

Les problèmes 1 et 8 ont eu la loi de Weibull ajustée et ceci quel que soit:

- La loi des paramètres (sauf pour des paramètres suivant la loi Poisson et la loi Binomiale).
- La taille des paramètres.
- La taille des échantillons.

le problème 10 a eu la loi normale ajustée, pour toutes les lois des paramètres.



La question que nous nous posons est de savoir s'il existe de points communs qui caractérisent ce regroupement en catégories.

### 1: Analyse de la première catégorie:

Les problèmes de la première catégorie sont généralement des problèmes à fonctions bornées. Quelle que soient les perturbations aléatoires apportées à la fonction, l'optimum est toujours égal à la borne de la fonction.

### 2: Analyse de la seconde catégorie:

Les problèmes de la seconde catégorie sont essentiellement des problèmes à fonctions quadratiques. Aucune loi de distribution de probabilité connue n'a pu être ajustée dans ces cas là. Cependant, nous pouvons analyser leurs histogrammes.

Soit le problème 2 [Annexe 1]:

$$f(X) = 4(x_1 - 5)^2 + (x_2 - 6)^2$$

a: Soit la fonction  $f^1(X, \xi)$ :

$$f^1(X, \xi) = \xi_1 x_1^2 + x_2^2 - 40x_1 - 12x_2 - 136$$

Afin d'expliquer l'allure des histogrammes, nous considérerons les fonctions génératrices des moments. En effet, soit la variable aléatoire  $Y$ , la fonction génératrice des moments  $M_Y(T) = E(e^{tY})$  de la variable  $Y$ , caractérise parfaitement sa distribution de loi de probabilité.

Soit la variable aléatoire  $Y$  telle que:

$$Y = \xi_1 x_1^2 + x_2^2 - 40x_1 - 12x_2 - 136$$

Il est clair que la variable aléatoire  $Y$  définit aussi l'optimum de la fonction  $f^1(X, \xi)$ :

Posons alors:

$$a = x_1^2$$

$$b = x_2^2 - 40x_1 - 12x_2 - 136$$

D'où:

$$Y = a\xi_1 + b$$

Soit la fonction génératrice des moments de la variable aléatoire Y:

$$M_Y(t) = M(e^{tY}) = M(e^{a\xi_1 t + bt}) = M(e^{a\xi_1 t}) M(e^{bt}) = M_{\xi_1}(ta) M_D(t)$$

b n'est pas aléatoire. b est une constante par rapport à  $\xi_1$ . Nous avons alors:  $M_D(t) = b$

D'où l'expression finale de la fonction génératrice des moments de la variable aléatoire Y:

$$M_Y(t) = b M_{\xi_1}(ta)$$

Si la variable aléatoire  $\xi_1$  suit la loi normale de paramètres  $\mu$  et  $\sigma$ :

$$M_{\xi_1}(t) = e^{t\mu + \frac{1}{2}\sigma^2 t^2}$$

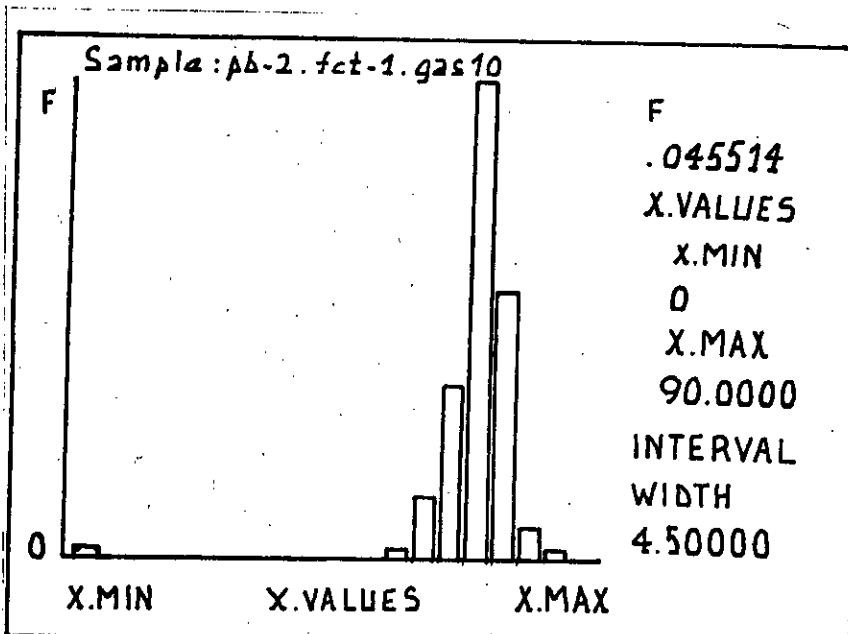
Nous avons alors:

$$M_Y(t) = b M_T(t)$$

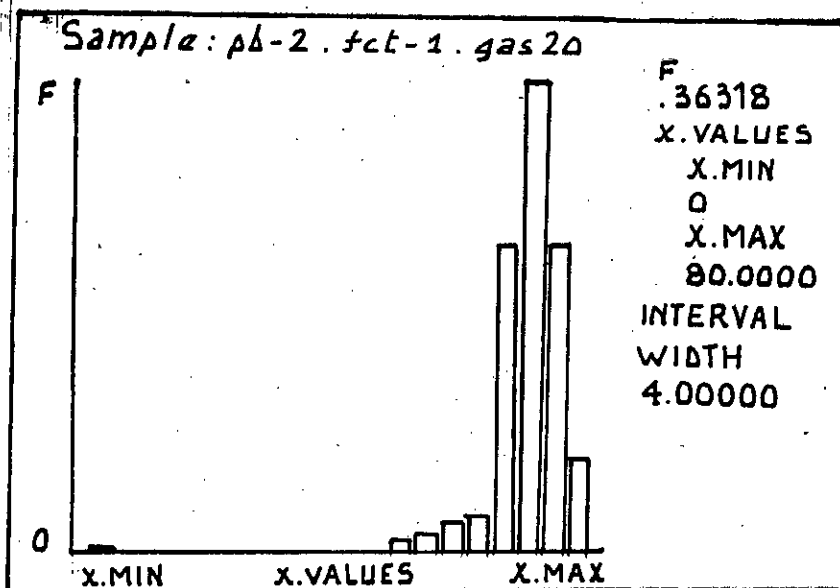
Où T est la variable aléatoire suivant la loi normale de paramètres  $\mu a$  et  $\sigma\sqrt{a}$ .

Moyennant le coefficient b, l'allure de l'histogramme de la variable aléatoire Y, devrait s'approcher de celle de la loi normale. C'est le cas pour cet exemple.

Le même raisonnement est applicable pour une autre loi de probabilité du paramètre  $\xi$  (voir histogrammes [Annexe 4]).



Légende:  
 pb-2: problème 2  
 fct-1: fonction 1  
 gas-10: loi des paramètres normale échantillon de taille 100  
 gas-20: loi des paramètres normale échantillon de taille 200



b: soit la fonction  $P^2(x, \xi)$

$$f^2(x, \xi) = \xi_1 x_1^2 + \xi_2 x_2^2 - 40x_1 - 12x_2 - 136$$

Posons:

$$a = x_1^2$$

$$b = x_2^2$$

$$c = 40x_1 - 12x_2 - 136$$

La variable aléatoire Y s'exprime donc comme suit:

$$Y = a\xi_1 + b\xi_2 + c$$

$\xi_1$  et  $\xi_2$  sont des variables aléatoires indépendantes.

D'où la fonction génératrice des moments de la variable aléatoire Y:

$$M_Y(t) = M_{\xi_1}(ta) M_{\xi_2}(tb) e^{ct}$$

Supposons que  $\xi_1$  suit la loi Poisson de paramètre  $\alpha_1$  et  $\xi_2$  suit la loi Poisson de paramètre  $\alpha_2$ :

$$M_{\xi_1}(t) = e^{\alpha_1(e^t - 1)}$$

$$M_{\xi_2}(t) = e^{\alpha_2(e^t - 1)}$$

Ainsi, la fonction génératrice de Y devient:

$$M_Y(t) = c M_{T_1}(t) M_{T_2}(t)$$

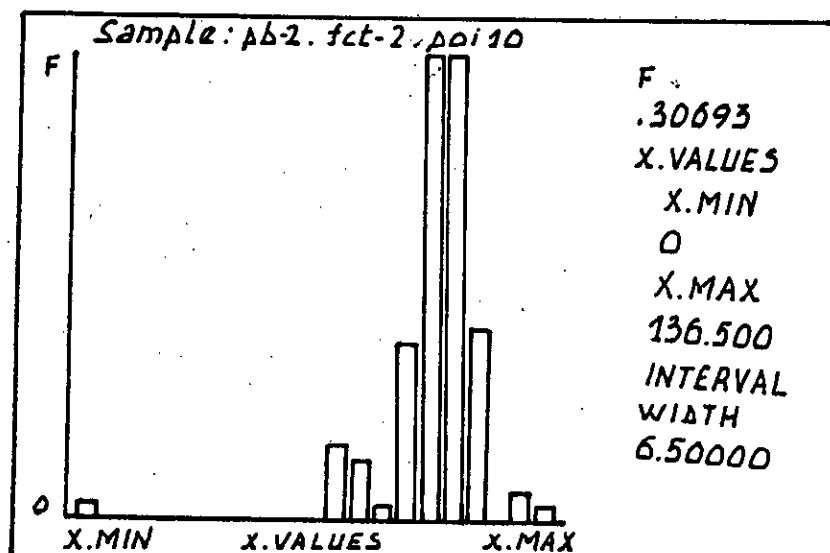
Où  $T_1$  est la variable aléatoire suivant la loi Poisson de paramètre  $\alpha_1$  et  $T_2$  est la variable aléatoire suivant la loi Poisson de paramètre  $\alpha_2$ .

Nous pouvons écrire aussi:

$$M_Y(t) = c M_T(t)$$

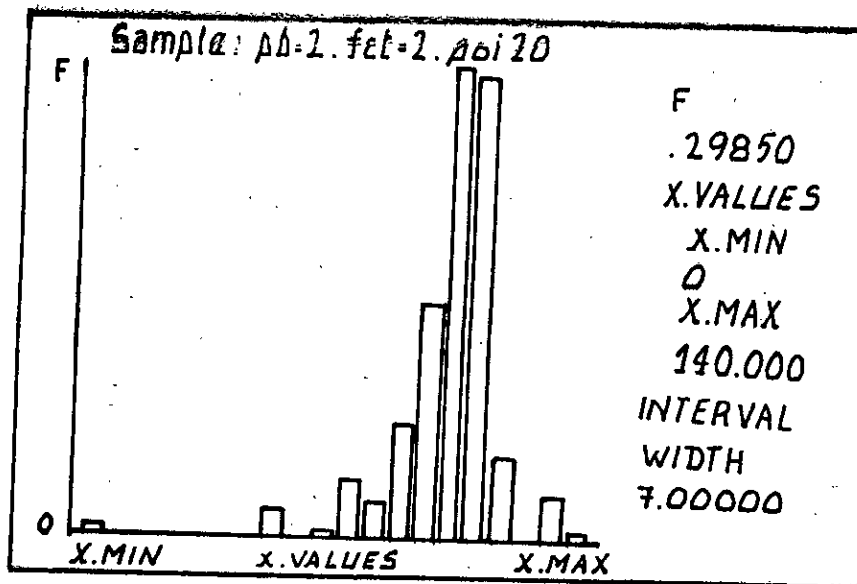
Où T est la variable aléatoire suivant la loi Poisson de paramètre  $\alpha_1 + \alpha_2$ .

Moyennant le coefficient c, l'allure de l'histogramme de la variable aléatoire Y, devrait s'approcher de celle de la loi Poisson. C'est le cas pour cet exemple.



Légende:

pb-2: problème 2  
 fct-2: fonction 2  
 poi-10: loi des  
 paramètres poisson  
 échantillon de  
 taille 100  
 poi-20: loi des  
 paramètres poisson  
 échantillon de  
 taille 200



Le même raisonnement est applicable pour les autres lois de probabilité du paramètre  $\xi$  (voir histogrammes [annexe 4]).

Le problème 2 a été construit de telle manière que le paramètre aléatoire ne soit affecté qu'une seule fois. Pour cela, la fonction  $f(X)$  a été d'abord développée, ensuite les paramètres aléatoires ont été affectés.

Que serait-il arrivé si la fonction  $f(X)$  avait été maintenue sous sa forme originale et que les paramètres aléatoires aient été affectés? C'est ce que nous tentons d'expliquer à travers le cas suivant:

c: Soit la fonction  $f^3(X, \xi)$ :

$$f^3(X, \xi) = 4(\xi_1 x_1 - 5)^2 + (x_2 - 6)^2$$

Nous calculons la fonction génératrice des moments de l'optimum.

Posons:

$$a = x_1$$

$$b = -5$$

$$c = (x_2 - 6)^2$$

D'où la variable aléatoire  $Y_3$ :

$$Y_3 = 4(a\xi_1 + b)^2 + c$$

Soit la variable aléatoire  $Y_3^2 = Y_3 / 4$

La fonction génératrice des moments de  $Y_3^2$  s'exprime comme suit:

$$M_{Y_3^2}(t) = cb^2 M_{\xi_1}(a^2 t) M_{\xi_1}(2abt)$$

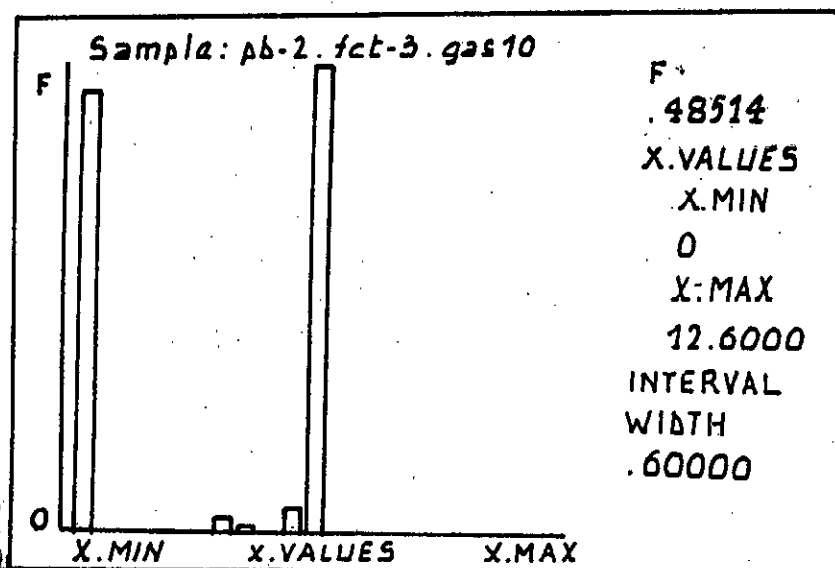
si la variable aléatoire  $\xi_1$  suit la loi normale de paramètres  $\mu$  et  $\sigma$ :

La fonction génératrice des moments de  $Y_3^2$  est le produit de deux fonctions génératrices des moments:

- l'une d'une variable aléatoire suivant une loi normale de paramètre  $\mu$  et  $\sigma$ .
- l'autre d'une variable aléatoire suivant une loi normale de paramètres  $\mu$  et  $\sigma$ , élevée au carré soit d'une chi-deux.

$Y_3^2$  est la somme d'une variable aléatoire suivant la loi normale et d'une variable aléatoire suivant la loi du chi-deux

Si nous comparons les histogrammes des variables aléatoires  $Y_1$  et  $Y_3^2$ , dans le cas où le paramètre  $\xi$  suit une loi normale, nous remarquons que l'histogramme de la variable aléatoire  $Y_1$  conserve l'allure de la loi normale. Par contre l'histogramme de la variable aléatoire  $Y_3^2$  présente une allure inconnue. Ce qui était prévisible.



Légende:

pb-2: problème 2  
 fct-3: fonction 3  
 gas-10: loi des  
 paramètres normale  
 échantillon de  
 taille 100

Ce même résultat peut être établi pour une autre loi de probabilité du paramètre  $\xi$ .

Finalement, nous pouvons dire que le comportement de l'optimum dépend essentiellement:

- De la loi des paramètres aléatoires.
- Du nombre de paramètres aléatoires si ce dernier modifie la nature de la fonction génératrice des moments.
- De la manière dont l'aléa est affecté à la fonction objective.

### 3: Analyse de la troisième catégorie:

Les problèmes de cette troisième catégorie sont des problèmes à fonctions quelconques.

Pour ces problèmes, les lois Weibull et normale ont été ajustées.

Pour réaliser des tests d'ajustement, en utilisant toute l'information contenue dans les échantillons, nous devons prendre le nombre de classes  $k$  égal à la taille  $n$  de l'échantillon. En pratique, ceci n'est pas toujours réalisable, nous avons donc regroupé les données en classes selon la règle de Sturges [Annexe4].

Le nombre, assez faible, de problèmes de cette catégorie ne nous permet de faire une analyse plus poussée.

Cependant, nous pouvons penser que le problème 10 a vérifié le théorème central limite. Tous les échantillons de l'optimum ont été ajustés par la loi normale.

En ce qui concerne les problèmes 1 et 8, le théorème central limite n'a pas été vérifié. Par contre, la loi Weibull a été ajustée (en regroupant les données en classes, suivant la règle de Sturges). Sachant que cette loi à trois paramètres, constitue une enveloppe à un bon nombre de lois (exponentielle, gamma, normale,...), nous pouvons penser qu'avec l'augmentation de la taille des échantillons, le théorème central limite pourrait être plus rapidement vérifié que pour les problèmes de la seconde catégorie.

Cette dernière catégorie peut cependant être analysée comme les problèmes de la seconde catégorie.

D: CONCLUSION DE L'ANALYSE:

Cette analyse nous a montré que le comportement de l'optimum dépend essentiellement

-Du comportement des paramètres aléatoires.

-De leur taille.

-Et surtout de la manière dont les perturbations aléatoires sont affectuées sur la fonction objectif.



CONCLUSION

## CONCLUSION

Dans certaines conjonctures économiques, l'optimisation stochastique est nécessaire. Bien que la programmation stochastique linéaire ait connu une grande évolution, la programmation stochastique non linéaire présente encore certains aspects à explorer.

Les approches de résolution de l'optimisation stochastique, en particulier les approches actives présentées, se heurtent souvent à des problèmes numériques. Lorsque nous connaissons la difficulté du choix du pas de déplacement et du critère d'arrêt dans le cas déterministe, il est aisé d'imaginer la difficulté dans le cas stochastique.

Notre étude a porté sur les problèmes stochastiques non linéaires et sans contraintes, en considérant une approche de résolution passive.

Nous avons pu réaliser nos tests à l'aide de deux méthodes d'optimisation:

- La méthode des gradients conjugués:

les résultats de l'analyse statistique entachés d'erreurs dues à la résolution nous ont amené à utiliser une autre méthode. Notre conception du processus d'analyse a fait que celle-ci soit inadéquate bien qu'elle présente de bonnes performances du point de vue convergence.

En effet, le changement de fonctions durant le processus implique, à tout instant, le choix d'un nouveau pas de dérivation qui est souvent mauvais. L'application de la méthode des gradients conjugués ne donne pas le bon optimum car elle est appliquée dans de mauvaises conditions.

- Enfin nous avons opté pour la méthode de Hook et Jeeves. cette dernière a l'avantage de procurer la flexibilité nécessaire à notre processus d'analyse bien qu'elle soit plus lente par rapport à la méthode des gradients conjugués car elle nécessite un temps de calcul plus grand..

Les problèmes numériques ne sont pas directement apparents. Cependant, nous ne pouvons pas ignorer les erreurs qui peuvent être induites durant la représentation informatique [LAP 82] de notre processus. Toute opération (addition, soustraction, multiplication, division et affectation ) exécutée en machine est entachée d'erreurs. Tout résultat obtenu est accompagné d'une perte de précision. Ainsi les valeurs numériques récoltées en fin de calculs ne sont que des valeurs voisines des valeurs exactes:

Durant l'exécution de notre processus nous cumulons des erreurs pendant:

- la génération des nombres pseudo-aléatoires.
- l'optimisation des fonctions. En plus des erreurs cumulées lors de l'exécution des opérations élémentaires de calcul, il y a une autre source d'erreur due au critère d'arrêt.

Soit la fonction  $f(X)$ , le processus itératif d'optimisation doit s'arrêter à l'itération  $k$  si le gradient de la fonction en ce point s'annule:

$$\frac{\partial f}{\partial X}(X^k) = 0$$

Dans la pratique, la nullité de la dérivée n'est jamais atteinte, étant donné que les calculs sont effectués avec une arithmétique à précision limitée [VIG 82]. Le test d'arrêt pratique optimal est alors le suivant:

$$\frac{\partial f}{\partial X}(X^k) \leq \epsilon$$

Le processus est donc arrêté lorsque l'optimum informatique est atteint. L'inconvénient est que si  $\epsilon$  est choisi:

- trop grand, alors le processus est arrêté avant l'obtention de la solution et le résultat est par conséquent mauvais.

- trop petit, alors le processus n'est pas arrêté assez tôt et de nombreux calculs sont inutilement effectués, ce qui nous amène à fixer le nombre maximal d'itérations que doit réaliser le processus.

Souvent l'optimum considéré n'est autre que la meilleure valeur de la fonction obtenue au bout de 400 itérations.

Si parfois, certaines anomalies sont apparues dans l'analyse statistique, nous pensons qu'elles sont dues essentiellement à l'ensemble de ces erreurs cummulées qui sont quasiment inévitables.

Au terme de ce travail, nous tirons certaines conclusions:

- Nous n'avons pu établir l'existence d'une loi probabiliste générale régissant le comportement de l'optimum.

En fait nous ne pouvons affirmer que ce résultat existe ou n'existe pas car la réalisation de tels tests demande la résolution de beaucoup de problèmes numériques et l'évaluation des erreurs cumulés tout au long du processus de calcul..

Les seules conclusions que nous pouvons tirer sont que le comportement de l'optimum dépend essentiellement:

-De la nature mathématique de la fonction.

-Des distributions des paramètres aléatoires: ces distributions définissent celle de l'optimum.

-Du nombre de paramètres aléatoires contenus dans la fonction objectif car il est en mesure de modifier la distribution de l'optimum.

-De la manière dont ces paramètres sont répartis dans la fonction objectif.

En fait notre étude devient particulièrement intéressante lorsque :

- Les distributions de probabilités des paramètres aléatoires de la fonction objectif ne sont pas connues. Il nous suffira de disposer d'observations de ces paramètres. Ceci nous permettra de générer la distribution de l'optimum.
- Les distributions de probabilités des paramètres aléatoires de la fonction objectif sont connues mais sont complexes. La distribution de l'optimum est obtenue grâce aux observations disponibles des paramètres aléatoires.
- La fonction objectif est mathématiquement complexe, ce qui rend les méthodes actives difficiles à appliquer.

En résumé, nous pouvons dire que quelle que soit la perturbation aléatoire apportée à la fonction, l'optimum présente une certaine stabilité.

Grâce à ce travail, nous avons pu établir un schéma de base général pour l'étude du comportement de l'optimum pour une fonction donnée. Ce schéma pouvant être amélioré de la manière suivante :

- 1- Reconsidérer la méthode des gradients conjugués (pour ses performances de convergence), en introduisant un module pour le traitement du choix du pas de dérivation. Ce cas impliquera nécessairement des difficultés numériques, qui une fois levées feront de cette méthode un bon instrument d'optimisation pour notre processus d'analyse.
- 2- Inclure un module d'analyse statistique dans notre processus, afin d'éviter l'utilisation de logiciels qui parfois ne répondent pas à nos besoins. De cette façon, nous obtiendrons un processus expérimental complet qui va de la génération des paramètres à l'analyse statistique.

3- Considérer le cas où chaque paramètre aléatoire suit une loi de probabilité différente des lois de probabilité suivies par les autres paramètres. En fait, nous ne pensons pas que les résultats qui pourraient être obtenus dans ce cas soient différents de ceux déjà obtenus car la stabilité du comportement de l'optimum a déjà été vérifiée par la méthode des approximations techniques [WET 85] et par notre étude expérimentale. Mais cet aspect mériterait d'être analysé car le cas considéré dans notre travail n'est rien d'autre qu'un cas particulier de ce dernier.

4- Nous pourrions aussi penser au cas où les paramètres aléatoires seraient dépendants, mais les méthodes de génération de nombres pseudo-aléatoires sont d'une grande complexité numérique [RUB 81]. Ce dernier cas pourrait à lui seul faire l'objet d'une étude.

**BIBLIOGRAPHIE**

[BAA 77] "*Linear programming and network flows*"

BAZARAA M.S & JARVIS C.M

John Wiley & sons 1977.

[BAA 77] "*Non linear programming*"

BAZARAA M.S & JARVIS C.M

John Wiley & sons 1977.

[BER 68] "*On some stochastic investment problem*"

BEREANU.B

Conference on operation research, Mathematiches  
forschungsinstitut, Oberwolfach. pp 18-24. August 1968.

[BIR 89] "*Sublinear upper bound for stochastic programs with  
recourse*"

BIRGE J.R & WETS R.J.B

Mathematical programming volume 43 1989, pp 131-149

[BRA 77] "*Applied mathematical programming*"

BRADLEY, Hax & MAGNANTI

Addison Wesley 1977

[CHA 58] "*Cost horizons and certainty equivalents: an approach to  
stochastic programming of heating oil*"

CHARNES A; COOPER W.W, SYMOND G,H

Managment science, volume.4,pp 235-263,1958.

[CHE 82] "*Dissertation*"

CHEPURNOI.N

V.Glushkov Institute of Cybernetics, Kiev 1982

[ERM 73] "*Limit extremum problems*"

ERMELIOV Y.U

Kibernetika volume 4, pp 130-132,1973

[ERM 83] "*Stochastic quasigradient methods and their applications to  
systems optimization*"

ERMELIOV Y.U

Stochastics volume 9, pp 1-36, 1983



- [GAI 88] "*Stochastic subgradient methods and their implementation*"  
GAIVORONSKI A.  
EMERLIOV, WETS R.J.B eds Springer Verlag 1988, pp 313-351.
- [GAS 84] "*A tight upper bound for the expectation of a convex function of a multivariate random variable*"  
GASSMAN.H & ZIEMBA.W  
Mathematical Programming study 27 1986, North Holland, pp 39-53.
- [GIN 84] "*Non linear optimization*"  
LASDON L & WARON A  
Lindo System Inc, 22 August 1984
- [KAL 88] "*Approximation techniques in stochastic programming*"  
KALL P, RUSZCZYNSKI A and FRAUNENDORFER M.  
Y.U EMERLIOV, R.J.B WETS eds Springer Verlag 1988
- [LAP 82] "*Algorithmes numérique analyse et mise en oeuvre tome 2: Equations et systèmes non linéaires*"  
LAPORTE M et VIGNES J  
eds Technip 1982
- [MAN 62] "*Methods of solutions of linear programs under uncertainty*"  
MANDANSKY  
Operations Research volume 10 1962, pp 165-176
- [MIN 83] "*Programmation mathématique: théorie et algorithmes*"  
MINOUX M.  
Tome 1 Editions DUNOD 1983.
- [NIE 86] "*Multidimensional numerical integration using pseudo numbers*"  
NIEDERREITER H.  
Mathematical Programming study 27 1986, North Holland, pp 17-38.
- [NUM 87] "*Numerical Recipes: The art of scientific computing*"  
PRESS W.H, FLANNERY B.P, TEUKOLSKY S.A and VETTERLING W.T.  
Cambridge University Press 1987

- [POL 66] "*A general method for solving extremum problems*"  
POLYAK B.T  
Soviet Mathematics volume 8, 1966, pp 593-596
- [QI 86] "*An alternative method for stochastic linear programming with simple recourse*"  
QI Liqun  
Mathematical Programming study 27 1986, North Holland, pp 183-190.
- [RUB 81] "*Simulation and Monte carlo method*"  
RUBENSTEIN  
John Wiley & sons eds New York 1981.
- [SZA 86] "*Evaluation of special multivariate gamma distribution function*"  
SZANTAI T.  
Mathematical Programming study 27 1986, North Holland, pp 1-16.
- [TIN 55] "*Stochastic linear programming with application to agricultural economics*"  
TINTNER G.  
In Proc 2nd sym L.P eds by H.A Antosiewicz 1955.
- [UNI 85] "*An interactive computer package for fitting probability distributions to observed data- version 2.11*"  
LAW A.M and VINCENT S.G.  
Marketing by simulation modeling and analysis compagny 1985.
- [VIG 84] "*Algorithmes numérique analyse et mise en oeuvre tome 1: Arithmétique des ordinateurs systèmes linéaires*"  
VIGNES J  
eds Technip 1984
- [WET 72] "*Characterization theorems for stochastic programs*"  
WETS R.J.B  
Mathematical programming, volume 2-1972, pp 166-175.

- [WET 74] "*Stochastic programs with fixed recourse: the equivalent deterministic problem*"  
WETS R.J.B.  
SIAM review volume 16-1974, pp 309-339.
- [WET 83] "*stochastic programming : Solution techniques and approximation schemes*"  
WETS R.J.B.  
Mathematical Programming: The State of the Art eds. A.BACHEM  
M.GRÖSCHEL and B.KORTE, Springer Verlag, Berlin 1983,  
pp 556-603
- [WET 84] "*Modeling and solution strategies for unconstrained stochastic optimization problems*"  
WETS R.J.B.  
Annals of operation research 1984, pp 3-22.
- [WET 85] "*Algorithmic procedures for stochastic optimization*"  
WETS R.J.B.  
NATO ASI series volume F15. Computational mathematical programming. Edited by K.SHITTOWSKI. Springer Verlag Berlin Heidelberg 1985.
- [WIL 66] "*On programming under uncertainty*"  
WILSON R.  
Operation Research volume 14 1966, pp 652-657
- [ZIE 74] "*Stochastic programs with simple recourse*"  
ZIEMBA W.T.  
Mathematical Programming Theory and Practice. P.L HAMMER and G.ZOUTENDJIK EDS. North Holland publishing compagny 1974.
- [ZOW 85] "*Non differentiable optimization*"  
ZOWE J.  
NATO ASI series volume F15. Computational mathematical programming. Edited by K.SHITTOWSKI. Springer Verlag Berlin Heidelberg 1985.

A N N E X E 1

**Problème 1:**

Source: Beale 1958

N° de variables indépendantes: 2

$$\text{Min } f(X) = U_1^2 + U_2^2 + U_3^2$$

$$U_i = C_i + x_i (1 - x_2^i)$$

$$C_1 = 1.5 \quad C_2 = 2.25 \quad C_3 = 6.25$$

$$X^0 = [1, 0.8]^T$$

$$X^* = [3, 0.5]^T \quad f(X^*) = 0$$

$$\text{Min } f^1(X, \xi) = (\xi_1 + x_1 (1 - x_2))^2 + U_2^2 + U_3^2$$

$$\text{Min } f^2(X, \xi) = (\xi_1 + x_1 (1 - x_2))^2 + ((2.25 - x_1 (1 - \xi_2 x_2^2))^2 + U_3^2$$

**Problème 2:**

Source: Himmelblau, Applied nonlinear programming, Mc Graw Hill 1972

N° de variables indépendantes: 2

$$\text{Min } f(X) = 4(x_1 - 5)^2 + (x_2 - 6)^2$$

$$X^0 = [8, 9]^T$$

$$X^* = [5, 6]^T \quad f(X^*) = 0$$

$$\text{Min } f^1(X, \xi) = \xi_1 x_1^2 + x_2^2 - 40x_1 - 12x_2 + 136$$

$$\text{Min } f^2(X, \xi) = \xi_1 x_1^2 + \xi_2 x_2^2 - 40x_1 - 12x_2 + 136$$

**Problème 3:**

Source: Engvall 1966

N° de variables indépendantes: 2

$$\text{Min } f(X) = x_1^4 + x_2^4 + 2x_1^2 x_2^2 - 4x_1 + 3$$

$$X^0 = [0.5, 2]^T$$

$$\text{Min } f^1(X, \xi) = \xi_1 x_1^4 + x_2^4 + 2x_1^2 x_2^2 - 4x_1 + 3$$

$$\text{Min } f^2(X, \xi) = \xi_1 x_1^4 + x_2^4 + \xi_2 x_1^2 x_2^2 - 4x_1 + 3$$


---

**Problème 4:**

Source: Himmeblau, Applied nonlinear programming, Mc Graw Hill 1972

N° de variables indépendantes: 2

$$\text{Min } f(X) = e^{-x_1 - x_2} (2x_1^2 + 3x_2^2)$$

$$X^0 = [2.5, 2.5]^T$$

$$X^* = [0, 1]^T \quad f(X^*) = 1.1036$$

$$\text{Min } f^1(X, \xi) = e^{-x_1 - x_2} (\xi_1 x_1^2 + 3x_2^2)$$

$$\text{Min } f^2(X, \xi) = e^{-x_1 - x_2} (\xi_1 x_1^2 + \xi_2 x_2^2)$$


---

**Problème 5:**

Source: Himmeblau, Applied nonlinear programming, Mc Graw Hill 1972

N° de variables indépendantes: 2

$$\text{Min } f(X) = \sin(3.6032 x_1 / 12) + \cos(1.8254 x_2 / 16)$$

$$X^0 = [0, 0]^T$$

$$X^* = [1, 1]^T \quad f(X^*) = -2$$

$$\text{Min } f^1(X, \xi) = \sin(\xi_1 x_1 / 12) + \cos(1.8254 x_2 / 16)$$

$$\text{Min } f^2(X, \xi) = \sin(\xi_1 x_1 / 12) + \cos(\xi_2 x_2 / 16)$$


---

**Problème 6:**

Source: Himmeblau, Applied nonlinear programming, Mc Graw Hill 1972

N° de variables indépendantes: 2

$$\text{Min } f(X) = (x_1^2 + x_2 - 11)^2 + (x_1^2 + x_2 - 7)^2$$

$$X^0 = [1, 1]^T$$

$$X^* = [3, 2]^T \quad f(X^*) = 0$$

$$\text{Min } f^1(X, \xi) = (\xi_1 x_1^2 + x_2 - 11)^2 + (x_1^2 + x_2 - 7)^2$$

$$\text{Min } f^2(X, \xi) = (\xi_1 x_1^2 + \xi_2 x_2 - 11)^2 + (x_1^2 + x_2 - 7)^2$$


---

**Problème 7:**

Source: Himmeblau, Applied nonlinear programming, Mc Graw Hill 1972

N° de variables indépendantes: 2

$$\text{Min } f(X) = (x_1^2 + 12x_2 - 1)^2 + (49x_1^2 + 49x_2^2 + 84x_1 + 2324x_2 - 681)^2$$

$$X^0 = [1, 1]$$

$$X^* = [-21.026653, -36.760090]^T \quad f(X^*) = 0$$

$$\text{Min } f^1(X, \xi) = (\xi_1 x_1^2 + 12x_2 - 1)^2 + (49x_1^2 + 49x_2^2 + 84x_1 + 2324x_2 - 681)^2$$

$$\text{Min } f^2(X, \xi) = (\xi_1 x_1^2 + 12x_2 - 1)^2 + (49x_1^2 + \xi_2 x_2^2 + 84x_1 + 2324x_2 - 681)^2$$


---

**Problème 8:**

Source: J. Bracken and G.P. McCormick, selected applications of non linear programming. John Wiley &amp; sons Inc New York.

N° de variables indépendantes: 2

$$\text{Min } f(X) = (x_1 - 2)^2 + (x_2 - 1)^2 + 0.04/g(x) + h(x) / 0.2$$

$$x^0 = [0, 0]^T$$

$$x^* = [1.7954, 1.3779]^T \quad f(x^*) = 0.16904$$

$$\text{Min } f^1(x, \xi) = (x_1 - \xi_1)^2 + (x_2 - 1)^2 + 0.04/g(x) + h(x) / 0.2$$

$$\text{Min } f^2(x, \xi) = (x_1 - \xi_1)^2 + (x_2 - \xi_2)^2 + 0.04/g(x) + h(x) / 0.2$$

### Problème 9:

Source: Himmelblau, Applied nonlinear programming, Mc Graw Hill 1972

N° de variables indépendantes: 2

$$\text{Min } f(x) = x_1^3 + x_2^2 - 3x_1 - 2x_2 + 2$$

$$x^0 = [0, 2]^T$$

$$x^* = [1, 1]^T \quad f(x^*) = -1 \text{ ou } \infty$$

$$\text{Min } f^1(x, \xi) = x_1^3 + \xi_1 x_2^2 - 3x_1 - 2x_2 + 2$$

$$\text{Min } f^2(x, \xi) = x_1^3 + \xi_1 x_2^2 - \xi_2 x_1 - 2x_2 + 2$$

### Problème 10:

Source: Himmelblau, Applied nonlinear programming, Mc Graw Hill 1972

N° de variables indépendantes: 3

$$\text{Min } f(x) = 100 \left( x_3 - \left( \frac{x_1 + x_2}{2} \right)^2 \right)^2 + (1 - x_1)^2 + (1 - x_2)^2$$

$$x^0 = [-1.2, 2, 0]^T$$

$$x^* = [1, 1, 1]^T \quad f(x^*) = 0$$

$$\text{Min } f^1(x, \xi) = 100 \left( x_3 - \left( \frac{x_1 + x_2}{2} \right)^2 \right)^2 + 1 + \xi_1 x_1^2 - 2\xi_1 x_1 + (1 - x_2)^2$$

$$\text{Min } f^2(x, \xi) = 100 \left( x_3 - \left( \frac{x_1 + x_2}{2} \right)^2 \right)^2 + 1 + \xi_1 x_1^2 - 2\xi_1 x_1 + (1 - \xi_2 x_2)^2$$

$$\text{Min } f^3(x, \xi) = 100 \left( \xi_3 x_3 - \left( \frac{x_1 + x_2}{2} \right)^2 \right)^2 + 1 + \xi_1 x_1^2 - 2\xi_1 x_1 + (1 - \xi_2 x_2)^2$$



Problème 11:

Source: Engwall 1966

N° de variables indépendantes: 3

$$\text{Min } f(X) = \sum_{i=1}^5 f_i(X)$$

$$f_1(X) = x_1^2 + x_2^2 + x_3^2 - 1$$

$$f_2(X) = x_1^2 + x_2^2 + (x_3 - 2)^2 - 1$$

$$f_3(X) = x_1 + x_2 + x_3 - 1$$

$$f_4(X) = x_1 + x_2 - x_3 - 1$$

$$f_5(X) = x_1^3 + 3x_2^2 + (5x_3 - x_1 + 1)^2 - 36$$

$$X^0 = [1, 2, 2, 0]^T$$

$$X^* = [0, 0, 1]^T \quad f(X^*) = 0$$

$$\text{Min } f^g(X, \xi) = \sum_{i=1}^5 f_i(X, \xi)$$

$$f_1(X, \xi) = \xi_1 x_1^2 + \xi_2 x_2^2 + \xi_3 x_3^2 - 1$$

$$f_2(X, \xi) = x_1^2 + \xi_2 x_2^2 + (x_3 - 2)^2 - 1$$

$$f_3(X, \xi) = \xi_1 x_1 + \xi_2 x_2 + \xi_3 x_3 - 1$$

$$f_4(X, \xi) = x_1 + x_2 - \xi_3 x_3 - 1$$

$$f_5(X, \xi) = \xi_1 x_1^3 + \xi_2 x_2^2 + (5x_3 - x_1 + 1)^2 - 36$$


---

## Problème 12:

Source: Himmeblau, Applied nonlinear programming, Mc Graw Hill 1972

N° de variables indépendantes: 3

$$\text{Min } f(X) = (x_1 - x_2 + x_3)^2 + (-x_1 + x_2 + x_3)^2 + (x_1 + x_2 - x_3)^2$$

$$X^0 = [100, -1, 2.5]^T$$

$$X^* = [0, 0, 0]^T \quad f(X^*) = 0$$

$$\text{Min } f^1(X, \xi) = (x_1 - x_2 + x_3)^2 + (-\xi_1 x_1 + x_2 + x_3)^2 + (x_1 + x_2 - x_3)^2$$

$$\text{Min } f^2(X, \xi) = (x_1 - \xi_2 x_2 + x_3)^2 + (-\xi_1 x_1 + x_2 + x_3)^2 + (x_1 + x_2 - x_3)^2$$

$$\text{Min } f^3(X, \xi) = (x_1 - \xi_2 x_2 + x_3)^2 + (-\xi_1 x_1 + x_2 + x_3)^2 + (x_1 + x_2 - \xi_3 x_3)^2$$

## Problème 13:

Source: Himmeblau, Applied nonlinear programming, Mc Graw Hill 1972

N° de variables indépendantes: 4

$$\text{Min } f(X) = 10 \sum_{i=1}^7 \left[ \frac{x_1^2 + x_2^2 a_i + x_3^2 a_i^2}{1 + x_4^2 a_i} - b_i \right]$$

$a_i$	$b_i$
0.0	1.391
0.000428	11.18
0.00100	16.44
0.00209	16.20
0.00848	22.20
0.00525	24.3132

$$X^0 = [2.7, 90, 1500, 10]^T \quad f(X^0) = 2.905 \cdot 10^4$$

$$X^* = [2.714, 140.4, 1707, 31.51]^T \quad f(X^*) = 318.5721$$

$$\text{Min } f(X, \xi) = 10 \sum_{i=1}^7 \left[ \frac{x_1^2 + x_2^2 a_i + x_3^2 a_i^2}{1 + x_4^2 a_i} - \xi_i \right]$$

## Problème 14:

Source: M.J.D Powell, Computer J.5, 147(1962)

N° de variables indépendantes: 4

$$\text{Min } f(X) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$$

$$X^0 = [1, 1, 1, 1]$$

$$X^* = [0, 0, 0, 0]^T \quad f(X^*) = 0$$

$$\text{Min } f^1(X) = (\xi_1 x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(\xi_1 x_1 - x_4)^4$$

$$\text{Min } f^2(X) = (\xi_1 x_1 + \xi_2 x_2)^2 + 5(x_3 - x_4)^2 + (\xi_2 x_2 - 2x_3)^4 + 10(\xi_1 x_1 - x_4)^4$$

$$\text{Min } f^3(X) = (\xi_1 x_1 + \xi_2 x_2)^2 + 5(\xi_3 x_3 - x_4)^2 + (\xi_2 x_2 - \xi_3 x_3)^4 + 10(\xi_1 x_1 - x_4)^4$$

$$\text{Min } f^4(X, \xi) = (\xi_1 x_1 + \xi_2 x_2)^2 + (\xi_3 x_3 - \xi_4 x_4)^2 + (\xi_2 x_2 - \xi_3 x_3)^4 + 10(\xi_1 x_1 - \xi_4 x_4)^4$$

## Problème 15:

Source: Levy 1969

N° de variables indépendantes: 4

$$\text{Min } f(X) = \exp(x_1 - x_2)^4 + 100(x_2 - x_3)^4 + \tan^4(x_3 - x_4) + x_1^2 + (x_4 - 1)^2$$

$$X^0 = [1, 2, 2, 2]$$

$$X^* = [0, 1, 1, 1]^T \quad f(X^*) = 0$$

$$\text{Min } f^1(X, \xi) = \exp(\xi_1 x_1 - x_2)^4 + 100(x_2 - x_3)^4 + \tan^4(x_3 - x_4) + x_1^2 + (x_4 - 1)^2$$

$$\text{Min } f^2(X, \xi) = \exp(\xi_1 x_1 - x_2)^4 + 100(\xi_2 x_2 - x_3)^4 + \tan^4(x_3 - x_4) + x_1^2 + (x_4 - 1)^2$$

$$\text{Min } f^3(X, \xi) = \exp(\xi_1 x_1 - x_2)^4 + 100(\xi_2 x_2 - \xi_3 x_3)^4 + \tan^4(x_3 - x_4) + x_1^2 + (x_4 - 1)^2$$

$$\text{Min } f^4(X, \xi) = \exp(\xi_1 x_1 - x_2)^4 + 100(\xi_2 x_2 - \xi_3 x_3)^4 + \tan^4(x_3 - x_4) + x_1^2 + (\xi_4 x_4 - 1)^2$$

A N N E X E 2

Les fonctions suivantes ont été optimisées (dans ce cas minimisées) par notre algorithme utilisant la méthode des gradients conjugués et comparée avec les résultats obtenus par le logiciel d'optimisation GINO [GIN 84].

$$f_1(X) = (x_1 - 1)^2 + (x_2 - 1)^2$$

$$f_2(X) = 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2$$

$$f_3(X) = \sin(\pi x_1 / 12) \cos(\pi x_2 / 16)$$

Les fonctions suivantes ont été tirées de la littérature et optimisées par notre algorithme.

$$f_4(X) = 10(x_1 - x_4) + (x_1 + 10x_2)^2 + 5(x_3 - x_4) + (x_2 - 2x_3)^4$$

Source: MJD Powell, Computer J.5, p 147 (1962)

$$f_5(X) = (x_1 x_2)^2 (1 - x_1)^2 [1 - x_1 - x_2 (1 - x_1)^5]^2$$

Source: Himmeblau, Applied nonlinear programming, Mc Graw Hill 1972

$$f_6(X) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

Source: Himmeblau, Applied nonlinear programming, Mc Graw Hill 1972  
solution d'un ensemble d'équations.

$$f_7(X) = 100[x_3 ((x_1 + x_2) / 2)^2]^2 + (1 - x_1)^2 + (1 - x_2)^2$$

Source: Himmeblau, Applied nonlinear programming, Mc Graw Hill 1972

$$f_8(X) = e^{-x_1 - x_2} (2x_1^2 + 3x_2^2)$$

Source: Himmeblau, Applied nonlinear programming, Mc Graw Hill 1972

	h	Gradients conjugués		GINO	
		$X^*$	$f(X^*)$	$X^*$	$f(X^*)$
$f_1$	$10^{-3}$	(.99,.99)	$2.84 \cdot 10^{-14}$	(1,1)	0
	$10^{-4}$	(1.00,1.00)	$4.98 \cdot 10^{-9}$		
$f_2$	$10^{-3}$	(2.332,2.666)	-12.666	(2.333,2.666)	-12.666
	$10^{-4}$	(2.332,2.664)	-12.666		
	$10^{-5}$	(2.334,2.668)	-12.666		
	$10^{-6}$	(2.344,2.647)	-12.665		
$f_3$	$10^{-2}$	(-6.03,1.68)	$-9.99 \cdot 10^{-1}$	(-6.00,-0.00)	-1.000
	$10^{-3}$	(-6.02,0.00)	$-9.99 \cdot 10^{-1}$		
	$10^{-4}$	(-6.01,0.00)	$-9.99 \cdot 10^{-1}$		
	$10^{-5}$	(.99,.99)	$2.11 \cdot 10^{-9}$		
	$10^{-6}$	(1.5,2.25)	$2.51 \cdot 10^{-1}$		
	$10^{-7}$				

Pour les fonctions  $f_1$  et  $f_3$  le pas  $h = 10^{-4}$  est convenable, parcontre pour la fonction  $f_2$  c'est le pas  $h = 10^{-3}$  qui l'est.

	h	Gradients conjugués		$x^*$	$f(x^*)$
		$x^*$	$f(x^*)$		
$f_4$	$10^{-2}$ ⋮ $10^{-7}$	(0,0,0,0)	0	(0,0,0,0)	0
$f_5$	$10^{-2}$ ⋮ $10^{-7}$	Dépassement de mémoire pour l'ordinateur		(1,00)	0
$f_6$	$10^{-3}$ $10^{-4}$ $10^{-5}$ $10^{-6}$	(2.999,1.999) (2.999,2.000) (2.999,1.999) (3.000,1.999)	$9.9271 \cdot 10^{-6}$ $4.7671 \cdot 10^{-8}$ $1.0311 \cdot 10^{-9}$ $2.082 \cdot 10^{-10}$	(3,2)	0
$f_7$	$10^{-3}$ $10^{-4}$ $10^{-5}$ $10^{-6}$ $10^{-7}$	(.9151,.9151,.981) (.991,.991,.98) (.999,.999,.999) (.999,.999,.999) (.999,.999,.999)	$1.442 \cdot 10^{-2}$ $1.868 \cdot 10^{-4}$ $1.680 \cdot 10^{-6}$ $1.019 \cdot 10^{-8}$ $5.513 \cdot 10^{-12}$	(1,1,1)	0
$f_8$	$10^{-3}$ ⋮ $10^{-7}$	(0,0)	0	(0,1)	1.1036

Les fonctions  $f_1$  et  $f_8$  donnent les mêmes résultats quelque soit le pas  $h$  pris entre  $10^{-3}$  et  $10^{-7}$ .

Les fonctions  $f_6$  et  $f_7$  elles donnent de bons résultats quelque soit le pas  $h$  pris entre  $10^{-3}$  et  $10^{-7}$ . Toutefois, nous pouvons considérer que pour la fonction  $f_6$   $h=10^{-6}$  est meilleur et pour la fonction  $f_7$  le pas  $h=10^{-7}$  l'est aussi.

A N N E X E 3



### Test du $\chi^2$ :

Soit  $Z$  la variable aléatoire étudiée, nous ne connaissons pas sa distribution de probabilité. Nous voulons tester l'hypothèse que  $Z$  suit une distribution de loi particulière.

La procédure consiste à obtenir un échantillon aléatoire de taille  $n$  de la variable  $Z$  dont la densité de probabilité est inconnue. Ces observations sont contenues dans un histogramme de fréquences ayant  $k$  classes d'intervalles.

Soit  $f_i$  la  $i$ ème fréquence observée dans la  $i$ ème classe, nous calculons  $E_i$  la fréquence théorique. Le test statistique est:

$$\chi_0^2 = \sum_{i=1}^k \frac{(f_i - E_i)^2}{E_i}$$

Nous pouvons montrer que  $\chi_0^2$  suit approximativement la distribution du chi-deux avec  $k-p-1$  degrés de liberté où  $p$  représente le nombre des paramètres de la distribution estimée par l'échantillon statistique.

Nous rejetons l'hypothèse que  $Z$  est conforme à la distribution considérée si :

$$\chi_0^2 > \chi_{\alpha, k-p-1}^2$$

Le nombre de classes  $k$  est choisi parfois, selon des règles empiriques, parmi lesquelles, la règle de STURGE qui donne le nombre minimal de classes à prendre:

$$k = 1 + (10 \times \text{Log}_{10}(n))/3$$

Il est, cependant, préférable de prendre le nombre de classe aussi grand que possible, afin de ne pas perdre des informations. Lorsque  $k=n$ , où  $n$  est la taille de l'échantillon, toute l'information contenue dans l'échantillon sera alors prise en considération lors de l'analyse statistique.

## Problème 1:

Source: Beale 1958

N° de variables indépendantes: 2

$$\text{Min } f(X) = U_1^2 + U_2^2 + U_3^2$$

$$U_i = C_i + x_i (1 - x_2^i)$$

$$C_1 = 1.5 \quad C_2 = 2.25 \quad C_3 = 6.25$$

$$X^0 = [1, 0.8]^T$$

$$X^* = [3, 0.5]^T \quad f(X^*) = 0$$

$$\text{Min } f(X, \xi) = U_1^2 + U_2^2 + U_3^2$$

$$U_i = C_i + \xi_i x_i (1 - x_2^i)$$

taille échant	Pas de dérivation h correspondant					
	Norma	Lois des paramètres Pois	Expon	Gamma	Unifo	Binom
10	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
20	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>
30	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
40	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
50	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
60	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
70	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
80	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
90	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
100	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
110	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
120	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
130	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
140	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
150	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
160	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
170	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
180	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
190	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
200	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>

Ce tableau représente la valeur du pas fixée pour chaque échantillon durant le processus d'analyse.

Problème 4:

Source: Himmeblau

N° de variables indépendantes: 2

$$\text{Min } f(X) = e^{-x_1 - x_2} (2x_1^2 + 3x_2^2)$$

$$X^0 = [ \quad ]^T$$

$$X^* = [0, 1]^T \quad f(X^*) = 1.1036$$

$$\text{Min } f(X, \xi) = e^{-\xi_1 x_1 - \xi_2 x_2} (\xi_1 x_1^2 + \xi_2 x_2^2)$$

taille échant	Pas de dérivation h correspondant					
	Norma	Poiss	Expon	Gamma	Unifo	Binom
10						
20						
30						
40						
50						
60						
70						
80						
90						
100						
110						
120						
130						
140						
150						
160						
170						
180						
190						
200						

quel que soit le pas h fixé  
 $10^{-2} \leq h \leq 10^{-6}$   
 l'optimum obtenu tend toujours vers  
 et ceci quel que soit la loi et la  
 valeur des paramètres

## Problème 5:

Source: Himmeblau

N° de variables indépendantes: 4

$$\text{Min } f(X) = 10 \sum_{i=1}^7 \left[ \frac{x_1^2 + x_2^2 a_i + x_3^2 a_i^2}{1 + x_4^2 a_i} - b_i \right]$$

$a_i$	$b_i$
0.0	1.391
0.000428	11.18
0.00100	16.44
0.00209	16.20
0.00348	22.20
0.00525	24.9192

$$X^* = [2.714, 140.4, 1707, 31.51]^T \quad f(X^*) = 318.5721$$

$$\text{Min } f(X, \xi) = 10 \sum_{i=1}^7 \left[ \frac{x_1^2 + x_2^2 a_i + x_3^2 a_i^2}{1 + x_4^2 \xi_i} - b_i \right]$$

taille échant	Pas de dérivation h correspondant					
	Norma	Poiss	Expon	Gamma	Unifo	Binom
10	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
20	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>
30	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
40	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
50	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
60	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
70	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
80	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
90	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
100	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
110	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
120	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
130	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
140	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
150	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
160	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
170	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
180	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
190	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>
200	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>	10 <sup>-3</sup>



**Problème 7:**

Source: Himmeblau

N° de variables indépendantes: 2

$$\text{Min } f(X) = \text{Sin}(3.6032 x_1/12) + \text{Cos}(1.8254 x_2/16)$$

$$X^0 = [ \quad ]^T$$

$$X^* = [0, 1]^T \quad f(X^*) = 1.1036$$

$$\text{Min } f(X, \xi) = \text{Sin}(\xi_1 x_1/12) + \text{Cos}(\xi_2 x_2/16)$$

taille échant	Pas de dérivation h correspondant					
	Norma	Poiss	Expon	Gamma	Unifo	Binom
10	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>
20	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-3</sup>
30	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
40	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
50	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
60	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
70	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
80	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
90	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
100	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
110	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
120	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
130	10 <sup>-1</sup>	10 <sup>-1</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
140	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
150	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
160	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
170	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
180	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
190	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>
200	10 <sup>-1</sup>	10 <sup>-2</sup>		10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-3</sup>









**Problème 14:**

Source: J.Bracken and G.P Cormick, "selected applications of non linear programming "

John Wiley & sons Inc New York

N° de variables indépendantes: 2

$$\text{Min } f(X) = (x_1 - 2)^2 + (x_2 - 1)^2 + 0.04/g(X) + h^2(X)/0.2$$

$$g(X) = -x_1^2/4 - x_2^2 + 1$$

$$h(X) = x_1 - 2x_2 + 1$$

$$\text{Min } f(X, \xi) = (\xi_1 x_1 - 2)^2 + (x_2 - 1)^2 + 0.04/g(X) + h^2(X)/0.2$$

taille échant	Pas de dérivation h correspondant					
	Norma	Poiss	Expon	Gamma	Unifo	Binom
10	10 <sup>-9</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
20	10 <sup>-9</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
30	10 <sup>-9</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
40	10 <sup>-9</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
50	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
60	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
70	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
80	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
90	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
100	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
110	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
120	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
130	10 <sup>-9</sup>	10 <sup>-1</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
140	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
150	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
160	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
170	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
180	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
190	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>
200	10 <sup>-9</sup>	10 <sup>-2</sup>	10 <sup>-2</sup>	10 <sup>-1</sup>	10 <sup>-2</sup>	10 <sup>-9</sup>

## Problème 1:

Loi des paramètres: Normale

n° échan	loi ajustée	$\beta$ calculé
10	Exp, Weib, Norm	1.121899
20	-	0.66321
30	Weibull	0.63025
40	Weibull	0.70934
50	-	0.64730
60	-	0.63495
70	Weibull	0.90886
80	Weibull	0.88656
90	Exp, Weib	0.95326
100	Weibull	0.82862
110	-	0.97401
120	Exp, Weib	0.94043
130	-	0.89282
140	Weibull	0.82705
150	Exp, Weib	1.02594
160	Exponentielle	0.92652
170	Exponentielle	0.91217
180	Exp, Weib	0.89443
190	Weibull	0.86708
200	Exponentielle	0.84817

-Nous entendons par  $\beta$  calculé, la valeur du paramètre  $\beta$  de la loi Weibull estimé d'après l'échantillon.

-La Loi Weibull pouvant être considérée comme une enveloppe de beaucoup de lois, car elle est définie par trois paramètres.

-Le paramètre  $\beta$  peut être un indicateur sur la forme de la loi qui pourrait être ajustée.

-A ce stade de l'analyse nous l'avons relevé pour vérification, vue que les résultats obtenue ne correspondent pas à nos attentes.

Problème 1:

Loi des paramètres: Gamma

n° échan	loi ajustée	$\beta$ calculé
10	-	0.36491
20	-	0.54770
30	-	0.45742
40	Weibull	0.55230
50	-	0.43176
60	Weibull	0.48681
70	Exponentielle	0.76916
80	Weibull	0.51581
90	-	0.64623
100	-	0.55193
110	Weibull	0.57891
120	-	0.61744
130	-	0.60000
140	-	0.5660
150	-	0.61126
160	Weibull	0.50052
170	-	0.57369
180	-	0.59868
190	-	0.56541
200	-	0.65598

## Problème 1:

Loi des paramètres: Poisson

n° échan	loi ajustée	$\beta$ calc
10	-	0.38702
20	-	0.52465
30	-	0.63446
40	Weibull	0.73156
50	Exp, Weib	0.72670
60	-	0.68674
70	Exponentielle	0.69378
80	-	0.78752
90	Exp, Weib	0.89474
100	Exponentielle	0.85781
110	-	0.75000
120	Exp, Weib	1.00836
130	Weibull	0.86891
140	Exp, Weib	0.97669
150	Exp, Weib	0.97961
160	Exponentielle	0.89409
170	-	0.82196
180	Exponentielle	0.84507
190	Weibull	0.85021
200	Exp, Weib	1.00645

Loi des paramètres: Exponentielle

n° échan	loi ajustée	$\beta$ calc
10	Exp, Weib	0.70465
20	Weibull	0.53410
30	Weibull	0.56017
40	-	0.63837
50	-	0.64072
60	Weibull	0.51401
70	Weibull	0.64457
80	-	0.71113
90	Weibull	0.67070
100	Weibull	0.63486
110	Weibull	0.65503
120	Weibull	0.64978
130	-	0.71198
140	Weibull	0.65341
150	Weibull	0.72311
160	-	0.74673
170	-	0.59699
180	-	0.66762
190	-	0.74429
200	Weibull	0.71854

## Problème 1:

Loi des paramètres: Binomiale

n° échan	loi ajustée	$\beta$ calc
10	-	-
20	-	1.16649
30	Weibull	0.27332
40	-	0.52292
50	Weibull	0.37489
60	Exponentielle	0.49376
70	-	0.47577
80	-	0.36553
90	-	0.45093
100	Weibull	0.40561
110	-	0.50686
120	-	0.39870
130	-	0.33487
140	Weibull	0.43475
150	-	0.44903
160	-	0.43786
170	-	0.38801
180	Weibull	0.42632
190	Exponentielle	0.38396
200	Weibull	0.52017

Loi des paramètres: Uniforme

n° échan	loi ajustée	$\beta$ calc
10	Exp, Weib, Norm	0.69667
20	Exp, Weib	0.29515
30	Weibull	0.82311
40	Exp, Weib	1.04550
50	Exp, Weib	1.13856
60	-	1.05863
70	-	1.43041
80	Exp, Weib	1.04486
90	-	0.59107
100	-	1.5467
110	Exp, Weib	1.11707
120	Weibull	1.28133
130	Exp, Weib	1.49780
140	-	1.29333
150	-	1.09527
160	Weibull	1.10672
170	-	1.20809
180	Weibull	1.23734
190	Weibull	1.25751
200	-	1.21197

## Problème 2:

Loi des paramètres: Normale

n° échan	loi ajustée	$\beta$ calculé
10	Exp, Weib	0.59459
20	Exp, Weib	0.60738
30		0.95285
40	Exp, Weib	0.95216
50	Exponentielle	0.75234
60		0.66835
70		0.97154
80		0.92547
90		0.88039
100	Exp, Weib	0.88741
110		0.93916
120		0.93597
130		1.02938
140		0.82014
150		0.91952
160	Exponentielle	0.83123
170		0.78493
180		0.79895
190		0.94131
200		0.85140

Loi des paramètres: Gamma

n° échan	loi ajustée	$\beta$ calculé
10		0.75265
20	Weibull	0.59633
30		0.70215
40	Weibull	0.69388
50		0.58384
60		0.60216
70		0.63485
80		0.56452
90		0.72548
100		0.79668
110	Weibull	0.67881
120		0.77819
130		0.68359
140		0.61318
150		0.72809
160		0.69605
170		0.65373
180		0.64234
190		0.69592
200		0.66540

## Problème2:

Loi des paramètres: "Poisson

Loi des paramètres: Exponentielle

n° échan	loi ajustée	$\beta$ calculé
10	Weibull	0.47032
20		0.45464
30	Exponentielle	0.65560
40	Weibull	0.66020
50		0.59157
60	Weibull	0.60556
70		0.65195
80	Weibull	0.59994
90		0.66032
100		0.62724
110		0.68331
120		0.59095
130		0.55364
140		0.62589
150		0.63484
160		0.58716
170		0.63063
180		0.60430
190		0.59997
200		0.65678

n° échan	loi ajustée	$\beta$ calculé
10	Exp, Weib	0.63529
20		0.67400
30	Exponentielle	0.51081
40	Exponentielle	0.63992
50	Normal	0.68226
60		0.62699
70	Exponentielle	0.53070
80		0.53921
90	Weibull	0.48905
100		0.43106
110	Exponentielle	0.56553
120		0.42308
130		0.46176
140		0.47790
150		0.48606
160		
170		0.45553
180		0.49949
190		0.49003
200		0.47397



## Problème 2:

Loi des paramètres: Uniforme

Loi des paramètres: Binomiale

n° échan	loi ajustée	$\beta$ calculé
10	Exp, Weib	0.66360
20	Weibull	0.36405
30		0.42309
40		0.52106
50		0.54173
60	Exponentielle	0.41568
70		0.53566
80		0.45447
90		0.51917
100		0.45265
110		0.52809
120		
130		0.43042
140	Exponentielle	0.64523
150		0.58745
160		0.66219
170		0.60992
180	Normal	0.46120
190	Exponentielle	0.48950
200		0.43470

n° échan	loi ajustée	$\beta$ calculé
10		
20		
30		
40		
50		
60		
70		
80	Exponentielle	0.48591
90		0.44944
100		0.64639
110		0.55273
120		0.52581
130		0.46803
140		0.52779
150		0.54128
160		0.58100
170	Weibull	0.46428
180		0.49729
190		0.42942
200		0.48185

## Problème 6:

Loi des paramètres: Normale

Loi des paramètres: Gamma

n° échan	loi ajustée	$\beta$ calculé
10	Weibull	0.12603
20		0.12952
30		0.13656
40		0.13936
50		0.14475
60		0.13747
70		0.14217
80		0.13527
90		0.13411
100		0.13744
110		0.14254
120		0.13802
130		0.13201
140		0.13850
150		0.13885
160		0.14808
170		0.18595
180		0.19412
190		0.19549
200		0.19370

n° échan	loi ajustée	$\beta$ calculé
10		0.60199
20		0.58690
30		0.94772
40		0.60006
50	Exponentielle	0.62841
60		0.69248
70		0.21448
80	Normal	0.62762
90		0.36715
100		1.11028
110		1.01763
120		0.98737
130		1.39584
140		1.20853
150		1.19004
160		0.78687
170	Normal	0.82721
180		1.3334
190		1.03301
200		0.98575

Problème 6:

Loi des paramètres: Poisson

Loi des paramètres: Exponentielle

n° échan	loi ajustée	$\beta$ calculé
10	Norm, Weib, Exp	0.40506
20	Normal	1.09196
30	Normal	0.60513
40		0.19678
50	Normal	0.90369
60	Normal	0.75883
70		0.21385
80		0.25251
90		
100		0.30427
110		0.22366
120		1.10828
130		0.12800
140		0.22817
150		0.28300
160		0.25018
170		
180		0.31612
190		0.23359
200		0.28462

n° échan	loi ajustée	$\beta$ calculé
10		0.15182
20	Normal	0.79122
30		0.17140
40		0.13674
50		0.12870
60		0.15157
70		0.10880
80		0.12364
90		0.14114
100		0.13164
110		0.13965
120		0.13170
130		0.14382
140		0.12694
150		0.14918
160		
170		0.13136
180		0.13809
190		0.14367
200		0.12249

## Problème 6:

Loi des paramètres: Uniforme

Loi des paramètres: Binomiale

n° échan	loi ajustée	$\beta$ calculé
10		0.53311
20		0.55489
30		0.60040
40	Weibull	0.45963
50		0.64885
60		0.43115
70		0.52976
80	Exponentielle	0.59195
90		0.51261
100		0.52029
110		0.78396
120		1.02278
130		1.01830
140		1.00872
150	Weibull	1.05466
160		0.97596
170		0.99399
180		1.15755
190		1.05403
200		1.01424

n° échan	loi ajustée	$\beta$ calculé
10	Norm, Weib, Exp	0.72197
20		1.53458
30		0.71101
40		0.39871
50		0.76238
60		0.95221
70		0.78132
80		1.02255
90	Normal	1.05040
100		0.70482
110		0.76697
120	Weibull	0.80806
130		0.81315
140	Normal	0.81908
150		0.93044
160		0.93296
170		0.90183
180		0.89075
190		1.1280
200		0.81837

## Problème 7:

Loi des paramètres: Normale

Loi des paramètres: Gamma

n° échan	loi ajustée	$\beta$ calculé
10	Normal	1.04384
20		1.63974
30		1.27829
40		1.36812
50		1.26828
60		1.57055
70		1.55775
80	Weibull	1.41845
90		1.55795
100		1.47835
110		1.3609
120		1.52482
130	Normal	1.42203
140	Weibull	1.79184
150		1.54535
160		1.72157
170		1.62105
180	Normal	1.67843
190		1.71906
200		1.66885

n° échan	loi ajustée	$\beta$ calculé
10	Weib, Norm	2.83958
20		1.45735
30	Weibull	1.33562
40		1.25283
50		1.33358
60	Weib, Norm	1.72253
70		1.60835
80		1.48656
90	Weibull	1.47690
100		1.23819
110		1.48489
120		1.46036
130		1.40612
140		1.42481
150		1.36532
160		1.34591
170		1.48289
180	0	1.43049
190		1.43251
200		1.35198

## Problème 7:

Loi des paramètres: Poisson

Loi des paramètres: Exponentielle

n° échan	loi ajustée	$\beta$ calcu
10	Weibull	1.89779
20		1.42609
30	Weib,norm	1.46000
40	Weibull	1.52162
50		1.51665
60	Normal	1.64533
70		1.52733
80		1.63753
90		1.76689
100		1.42860
110		1.86250
120		1.68055
130		1.90426
140		1.92370
150	Weibull	2.05602
160		1.49504
170		1.54421
180		1.59031
190		1.70348
200		1.52374

n° échan	loi ajustée	$\beta$ calc
10		1.13663
20	Weibull	0.53323
30		0.66894
40		0.65041
50		0.57046
60		0.64460
70		0.63423
80	Weibull	0.66010
90		0.66516
100		0.61002
110		0.64478
120		0.65765
130		0.56992
140		0.54614
150		0.69578
160		0.61492
170		0.56852
180		0.63985
190		0.56252
200		0.69246

## Problème 7:

Loi des paramètres: Uniforme

Loi des paramètres: Binomiale

n° échan	loi ajustée	$\beta$ calculé
10		
20	Weib, Norm	8.34284
30	Weib, Norm	13.4534
40		10.1388
50		8.19334
60		7.86521
70		7.26648
80		7.99138
90		10.3494
100		13.6090
110	Weib, Norm	6.49987
120	Normal	11.5436
130		13.3604
140		4.52392
150		7.98093
160		7.11424
170		7.14160
180	Normal	5.93222
190		6.45970
200		8.39718

n° échan	loi ajustée	$\beta$ calculé
10		1.16581
20		0.60092
30		0.53214
40	Weib, Expo	0.57914
50	Weibull	0.64431
60		0.60849
70		0.61683
80		0.58855
90		0.72220
100		0.60810
110		0.60241
120	Weibull	0.63062
130		0.00635
140	Exponentiel	0.64079
150		0.00627
160		0.00627
170		0.00611
180		0.63534
190		0.60515
200		0.00685

## Problème 8:

Loi des paramètres: Normale

n° échan	loi ajustée	$\beta$ calculé
10	Weibull	0.36366
20	Weibull	0.61479
30		
40		
50		
60		
70		
80		
90		
100	Weibull	0.69197
110		
120		
130	Weibull	0.76466
140		
150		
160		
170		
180		
190		
200		

Loi des paramètres: Gamma

n° échan	loi ajustée	$\beta$ calculé
10	Exponentielle	1.19948
20	Exponentielle	1.13833
30		1.23085
40		1.08753
50	Weibull	1.33521
60	Weibull	1.28816
70	Weibull	1.56928
80		1.45395
90	Weibull	1.35394
100	Expo, Weib	1.28632
110		
120		
130	Exponentielle	1.32142
140	Weibull	1.36956
150		1.43550
160	Weibull	1.58819
170	Weibull	1.48870
180	Weibull	1.26521
190	Weibull	1.39972
200		1.34140



## Problème 8:

Loi des paramètres: Poisson

Loi des paramètres: Uniforme

n° échan	loi ajustée	$\beta$ calc
10		0.37341
20	Weibull	0.50456
30	Weibull	0.51328
40		0.30721
50	Weibull	0.53189
60		
70		
80		
90		
100		
110		
120	Weibull	0.75878
130		
140		0.79309
150		0.81033
160		
170	Weibull	0.82157
180		0.86140
190		0.85116
200		

n° échan	loi ajustée	$\beta$ calcu
10		
20		
30	Weib, Norm	55.684
40	Weib, Norm	5.6591
50		
60		2.6808
70	Normal	8.3522
80	Weib, Norm	7.6171
90		
100	Weib, Norm	9.6357
110	Weib, Norm	1.8932
120		1.8871
130		.03029
140		.16417
150	Normal	.21968
160	Normal	5.9523
170	Weibull	9.1478
180		5.1516
190		
200	Normal	9.2546

Problème 8:

Loi des paramètres: Binomiale

n° échan	loi. ajustée	$\beta$ calculé
10	Weibull	29.7716
20	Exp, Weib, Norm	0.74157
30	Expc, Weib	0.82948
40	Weibull	0.96166
50	Weibull	1.01756
60		
70		0.92501
80		
90	Normal	1.01588
100		
110		
120		
130		
140		
150		
160		
170		
180		
190		
200		

ANNEXE 4

Nous présentons ici différents histogrammes de certains problèmes.

Légende des histogrammes:

exemples:

Sample: pb-2, fct-1, gas-20

pb-2: problème 2

fct-1: fonction  $f^1(X, \xi)$

gas-20: loi normal pour les paramètres et l'échantillon considéré est de taille 200

Sample: pb-6, fct-2, gam-10

pb-2: problème 6

fct-1: fonction  $f^2(X, \xi)$

gam-10: loi gamma pour les paramètres et l'échantillon considéré est de taille 100

les différentes lois considérées sont indiquées par les symboles suivants:

Gas : Normal

Gam : Gamma

Poi : Poisson

Exp : Exponentielle

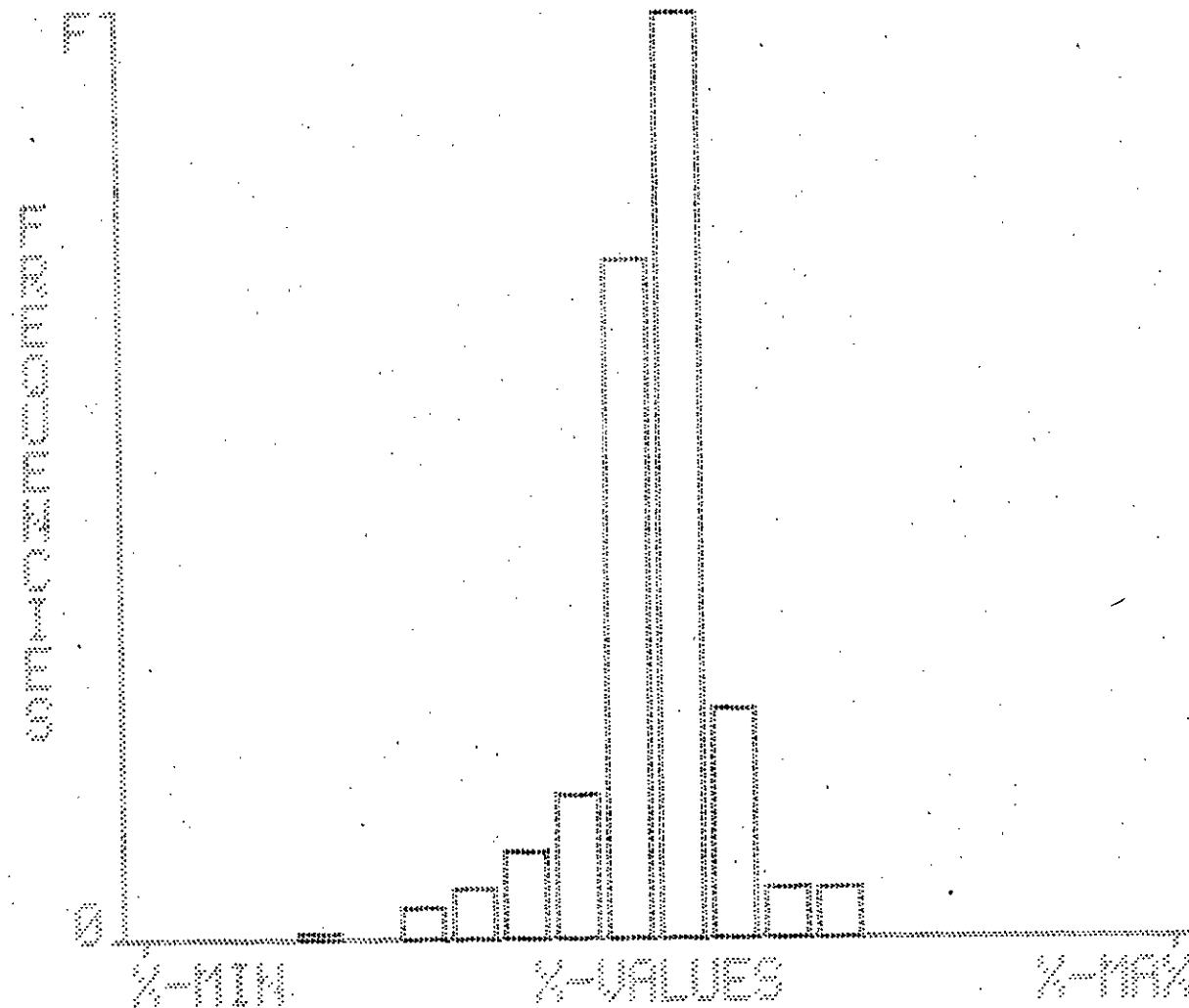
Bnl : Binomiale

Ran : Uniforme

Uni : Uniforme

HISTOGRAM OF

SAMPLE : nb-2, fact - 1, m=120



FREQUENCIES

BELOW PLOT

4.97512E- 3

ABOVE PLOT

0.

F

.40258

%-VALUES

%-MIN

-24.3333

%-MAX

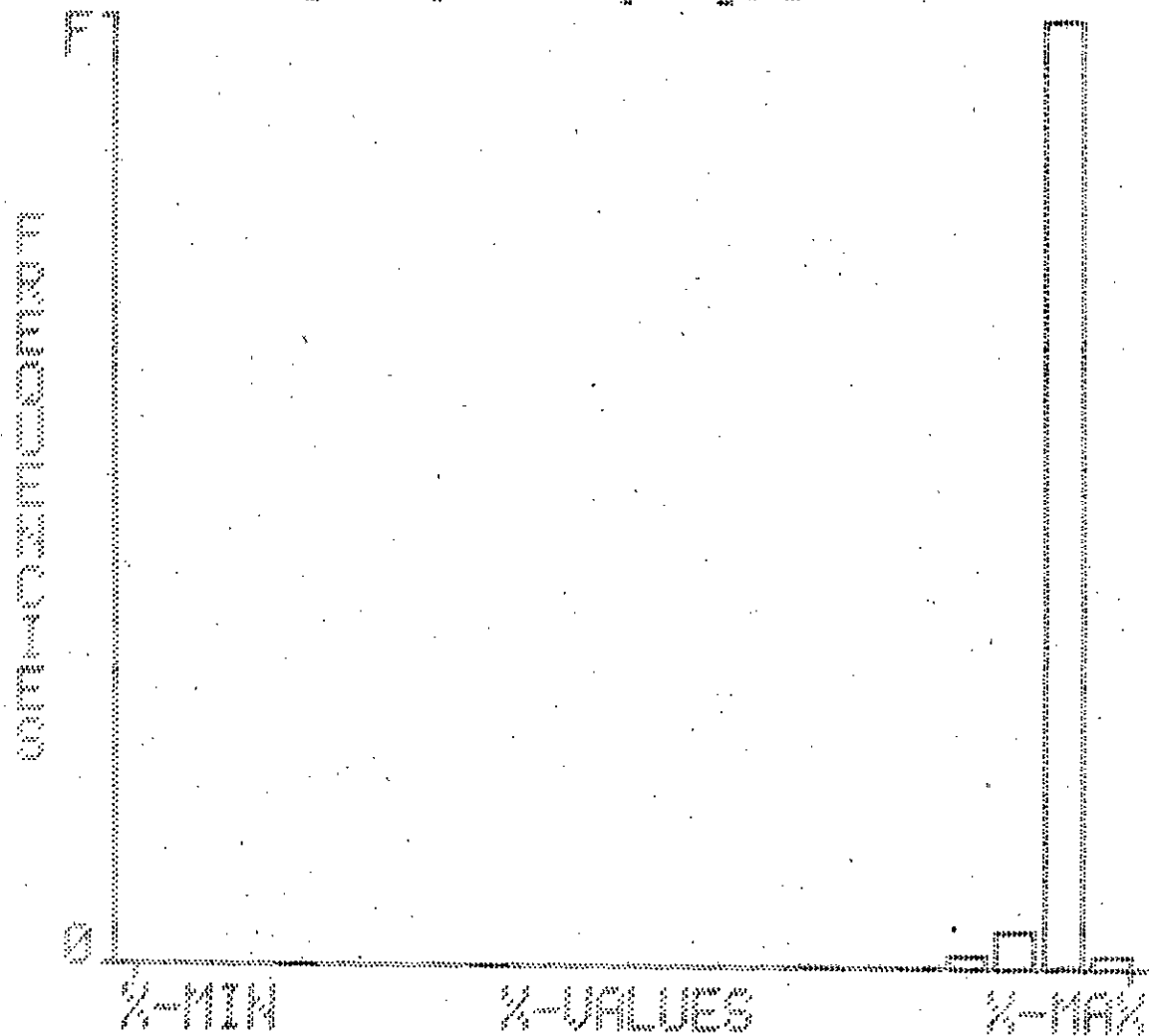
166.667.

INTERVAL WIDTH

10.0000

12-MAY-1992 10:37 AM

HISTOGRAM OF  
 SAMPLE : pb-2, fct-1, exp20



FREQUENCIES  
 BELOW PLOT  
 4.97512E- 3  
 ABOVE PLOT  
 0.

F  
 .90049

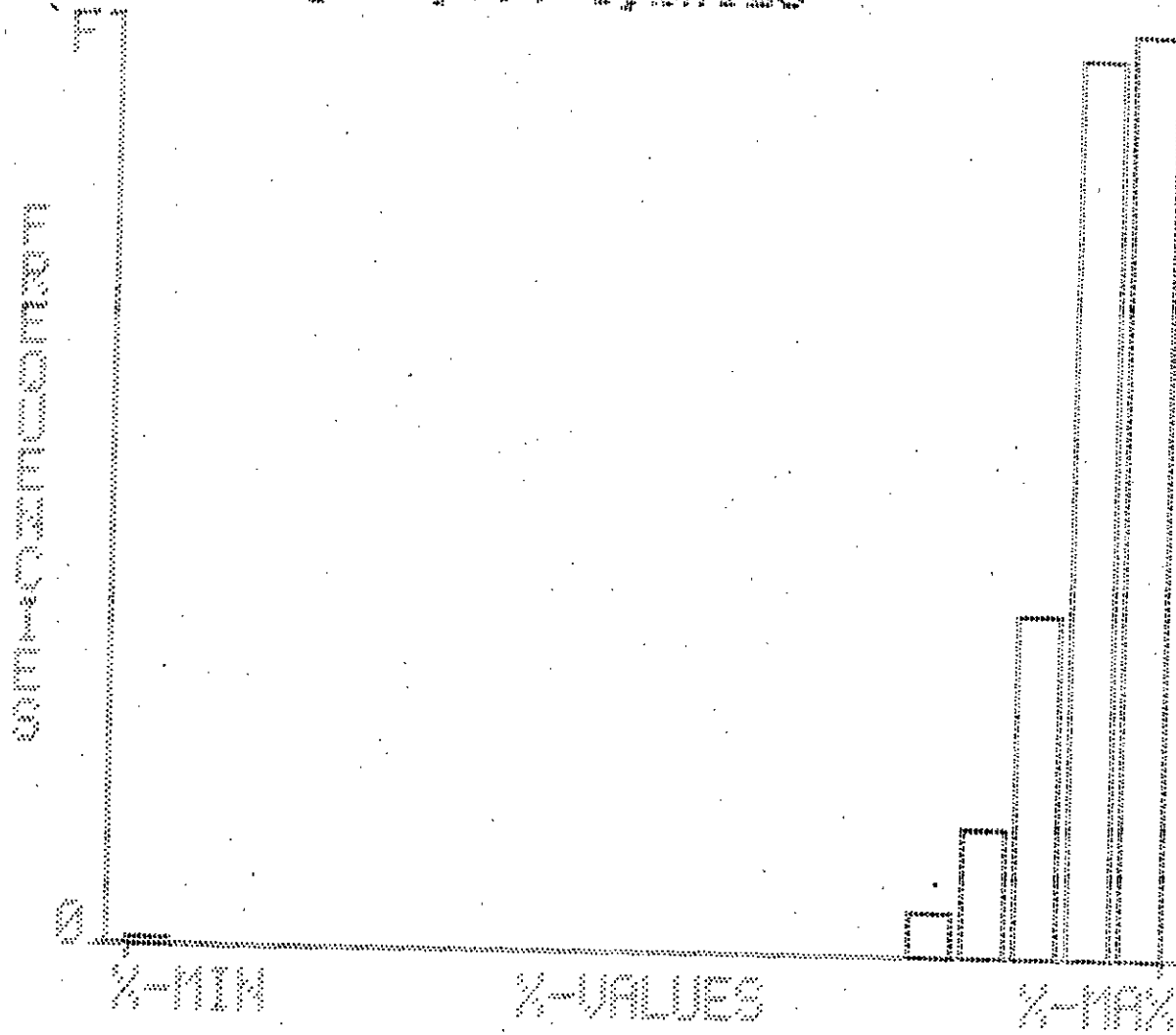
%-VALUES

%-MIN  
 -1.59000E+ 4

%-MAX  
 879.000

INTERVAL WIDTH  
 799.000

HISTOGRAM OF  
 SAMPLE : pt-2, Act-1, Km 120



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

.01990

F

.38308

%-VALUES

%-MIN

0.

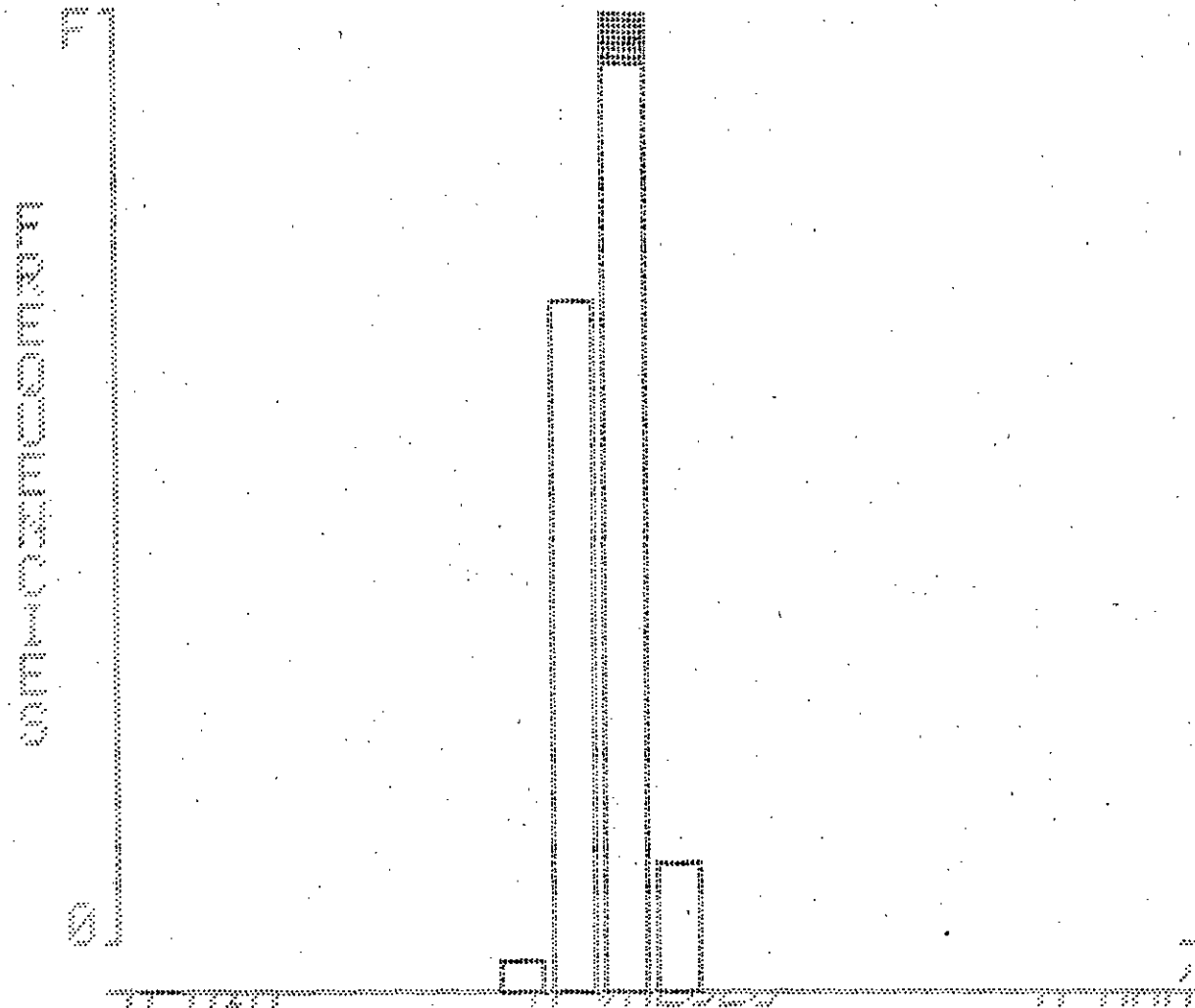
%-MAX

88.0000

INTERVAL WIDTH

4.00000

SAMPLE # 06-2, for -1, ran 26



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

0.

F

.51741

%-VALUES

%-MIN

0.

%-MAX

84.0000

INTERVAL WIDTH

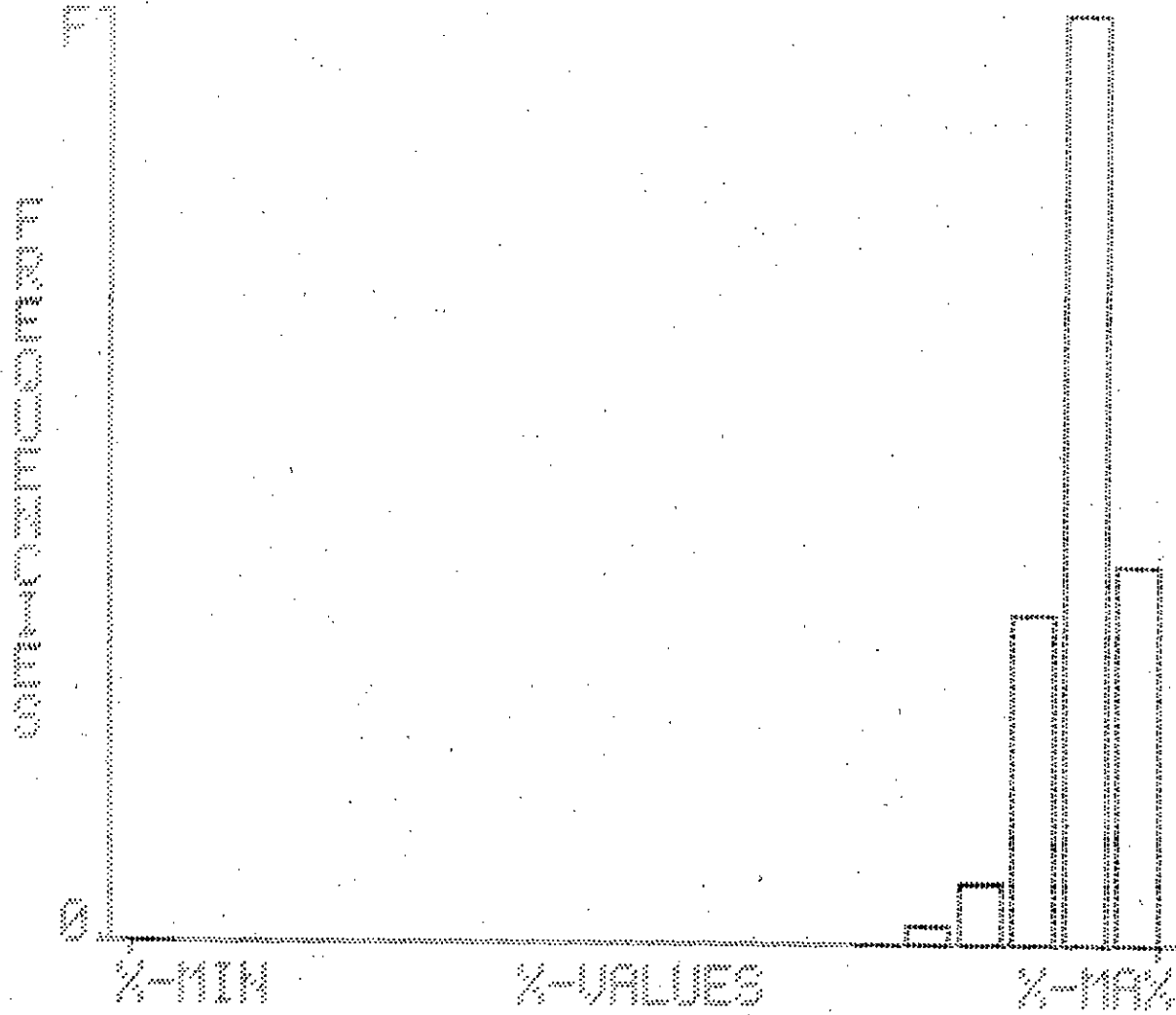
4.20000

Enter F to print, or RETURN to continue : 33333333



HISTOGRAM OF

SAMPLE # ph-2, fot-2, gas20



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

4.97512E-3

F

.52736

%-VALUES

%-MIN

0.

%-MAX

112.000

INTERNAL WIDTH

5.60000

HISTOGRAM OF  
 SAMPLE : pb-2, fct-2, exp20



FREQUENCIES

BELOW PLOT

4.97512E- 3

ABOVE PLOT

0.

F

.79104

%-VALUES

%-MIN

-6657.89

%-MAX

122.313

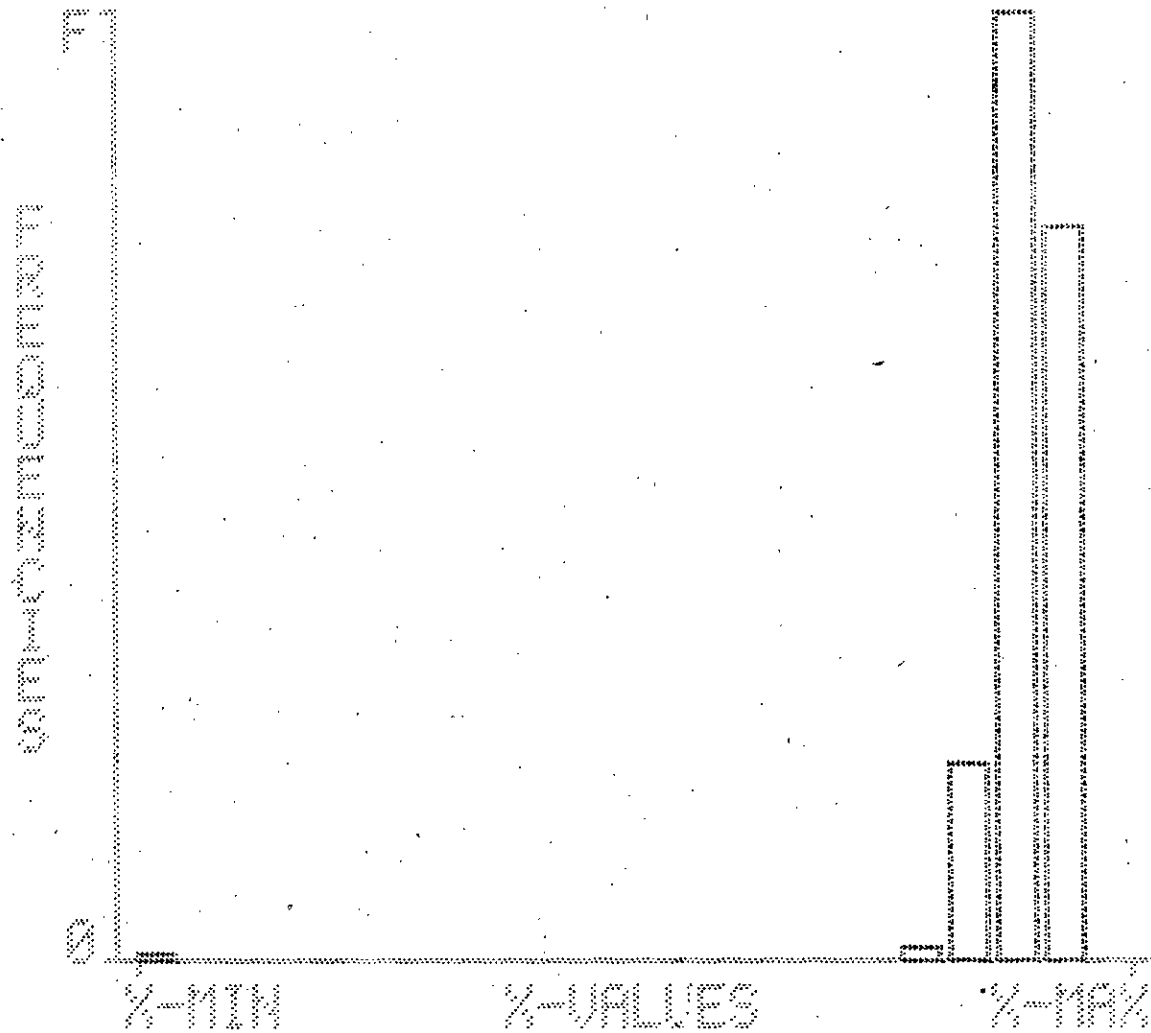
INTERVAL WIDTH

339.000

12-MAY-1993 11:30 AM

HISTOGRAM OF

SAMPLE # 04-2, 04-2, 04-2



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

0.

F

.49751

%-VALUES

%-MIN

0.

%-MAX

119.700

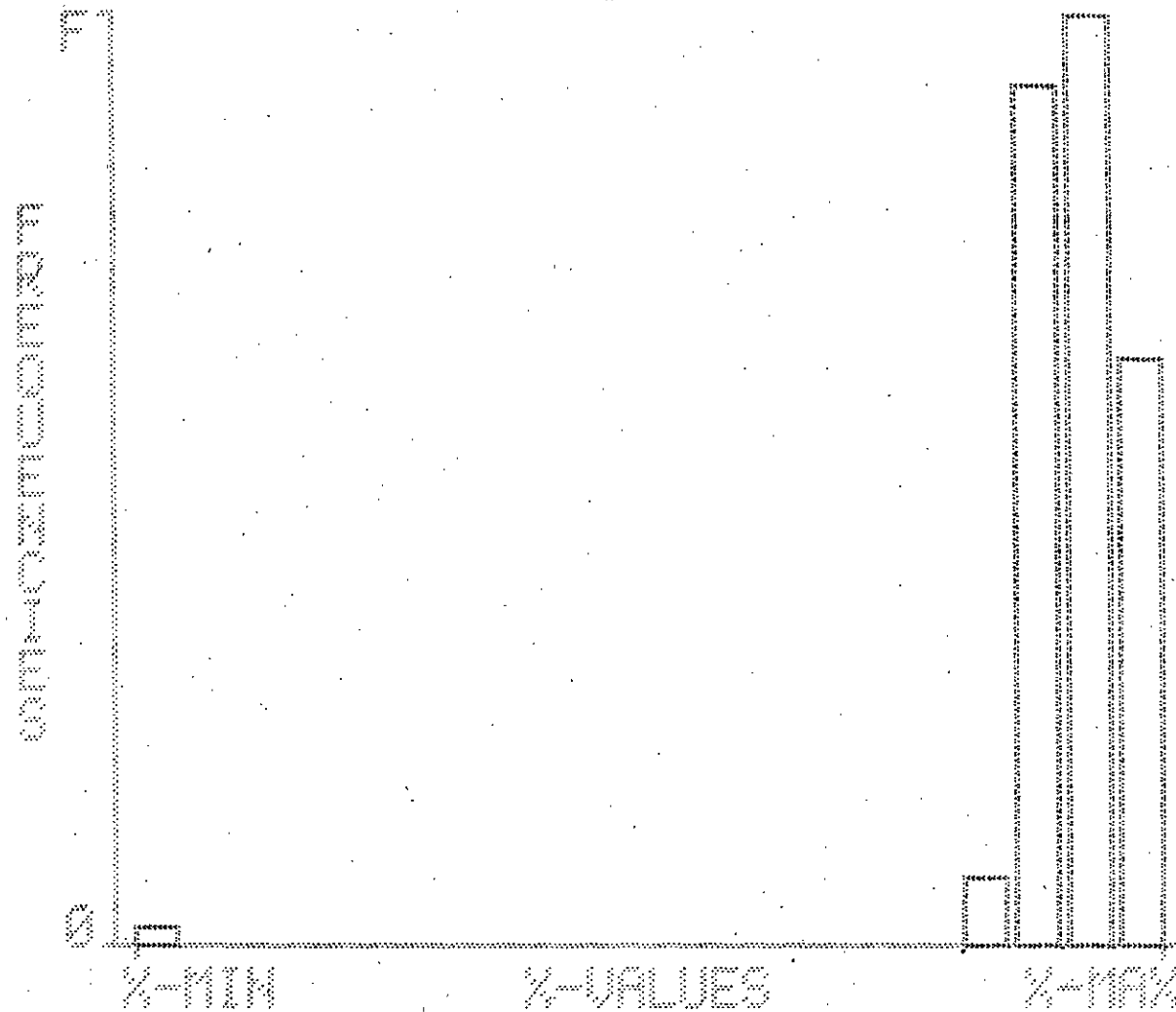
INTERVAL WIDTH

5.70000

12-MAY-1983 11:53 AM

HISTOGRAM OF

SAMPLE : #1-2, tot-2, ran 10



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

0.

F

.37813

%-VALUES

%-MIN

0.

%-MAX

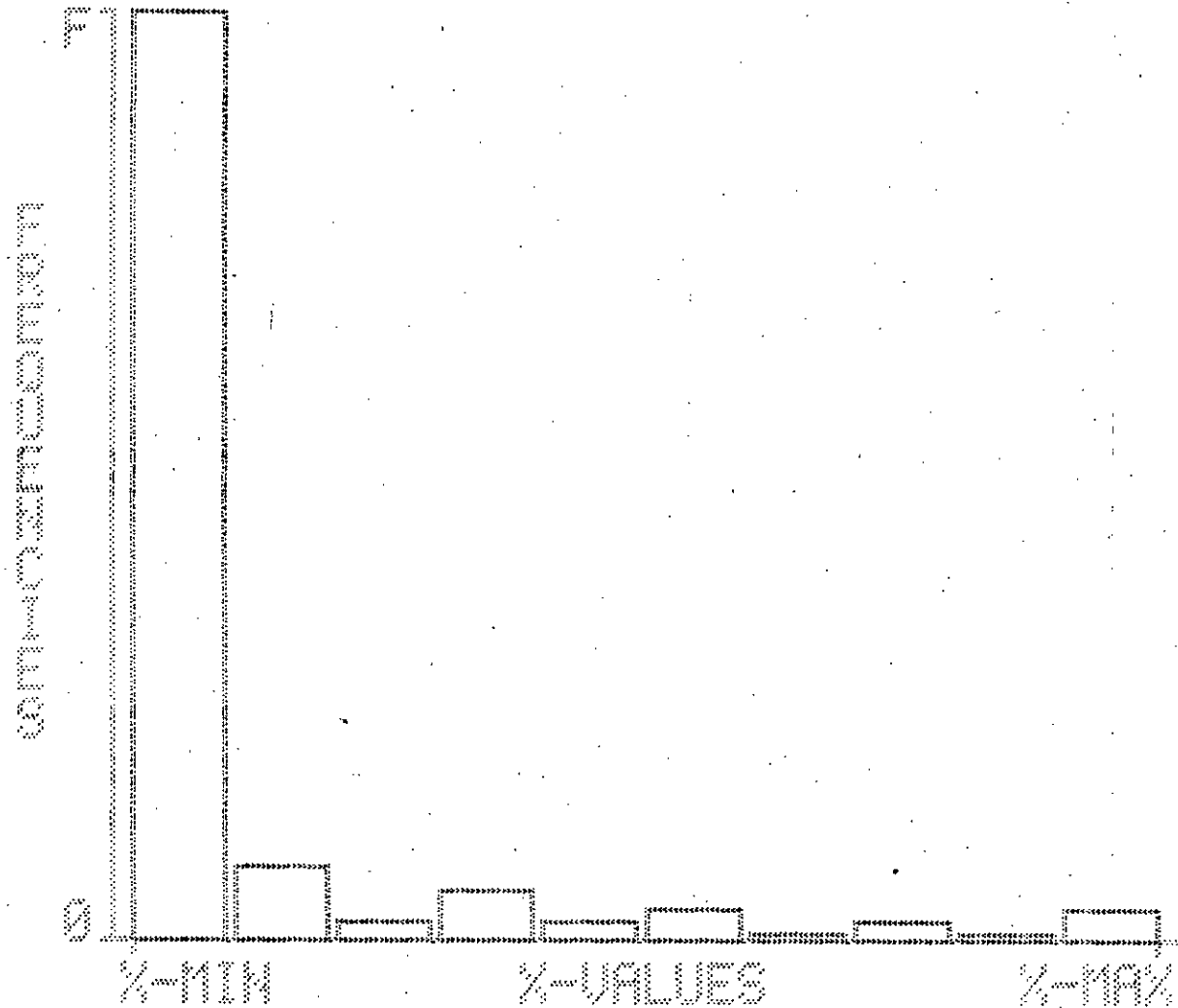
70.0000

INTERVAL WIDTH

3.50000

HISTOGRAM OF

SAMPLE : pt-5, tot-2, exp20



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

0.

F

.75124

%-VALUES

%-MIN

0.

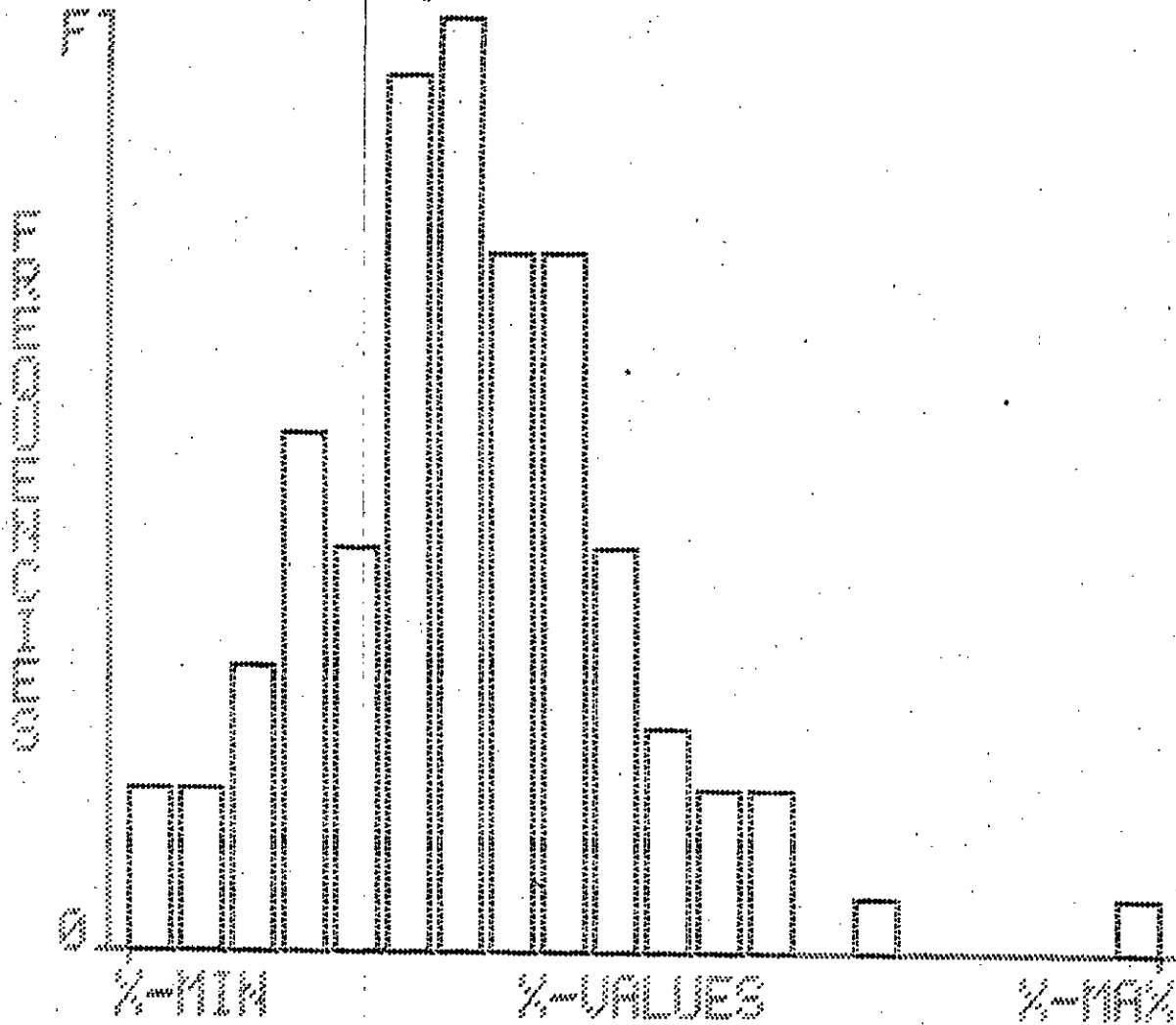
%-MAX

75.2384

INTERVAL WIDTH

7.52384

16.06.11  
 SAMPLE # 10-2, 10-1



FREQUENCIES

BELOW PLOT  
 Q.

ABOVE PLOT  
 Q.

F  
 .15841

%-VALUES

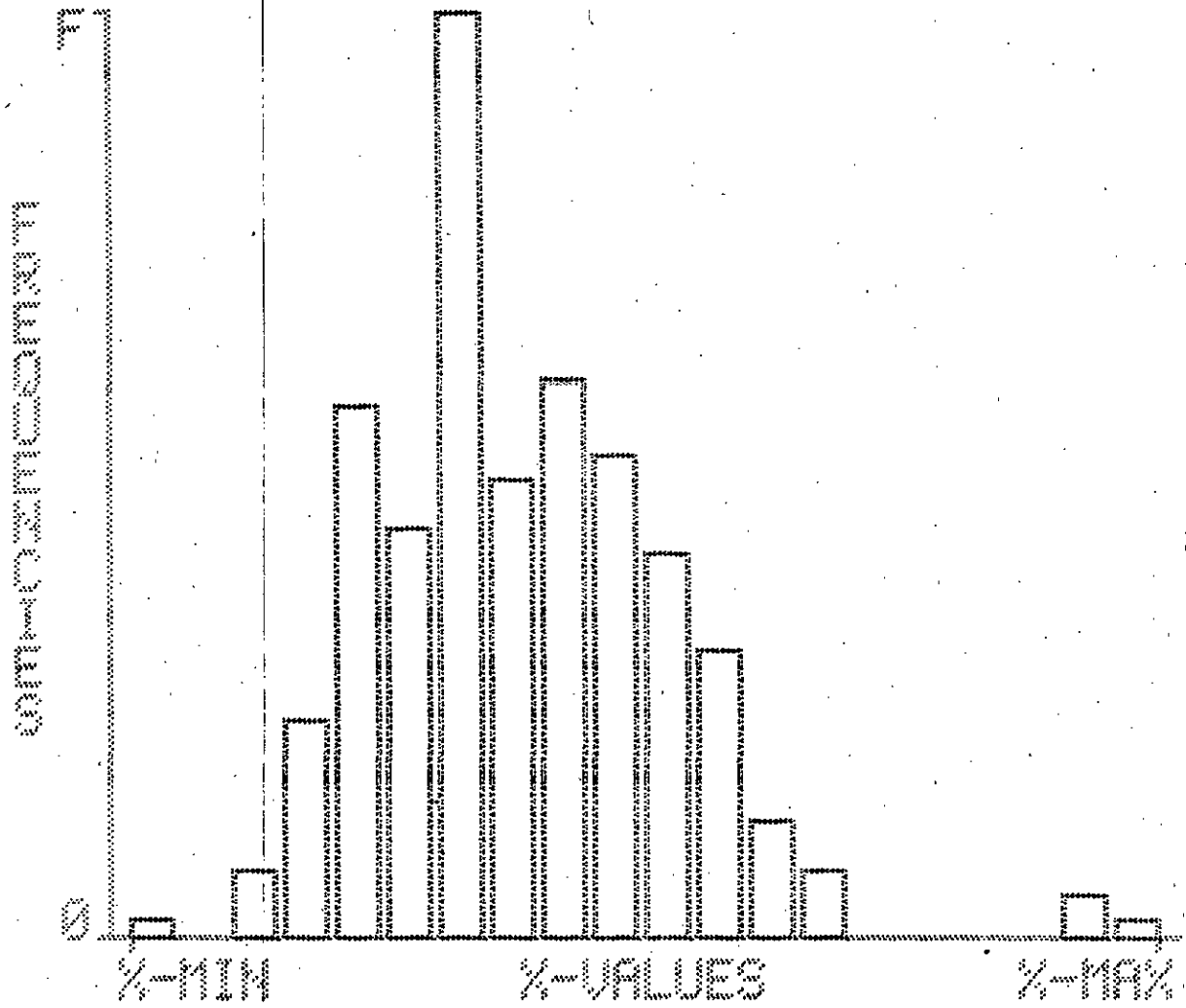
%-MIN  
 -153.093

%-MAX  
 376.937

INTERVAL WIDTH  
 28.5000

16-MAY-1993 11:48 AM

HISTOGRAM OF  
 SAMPLE # 101-7, 101-1, 03M20



FREQUENCIES  
 BELOW PLOT

0.

ABOVE PLOT

0.

F

.19895

%-VALUES

%-MIN

0.

%-MAX

90.0000

INTERVAL WIDTH

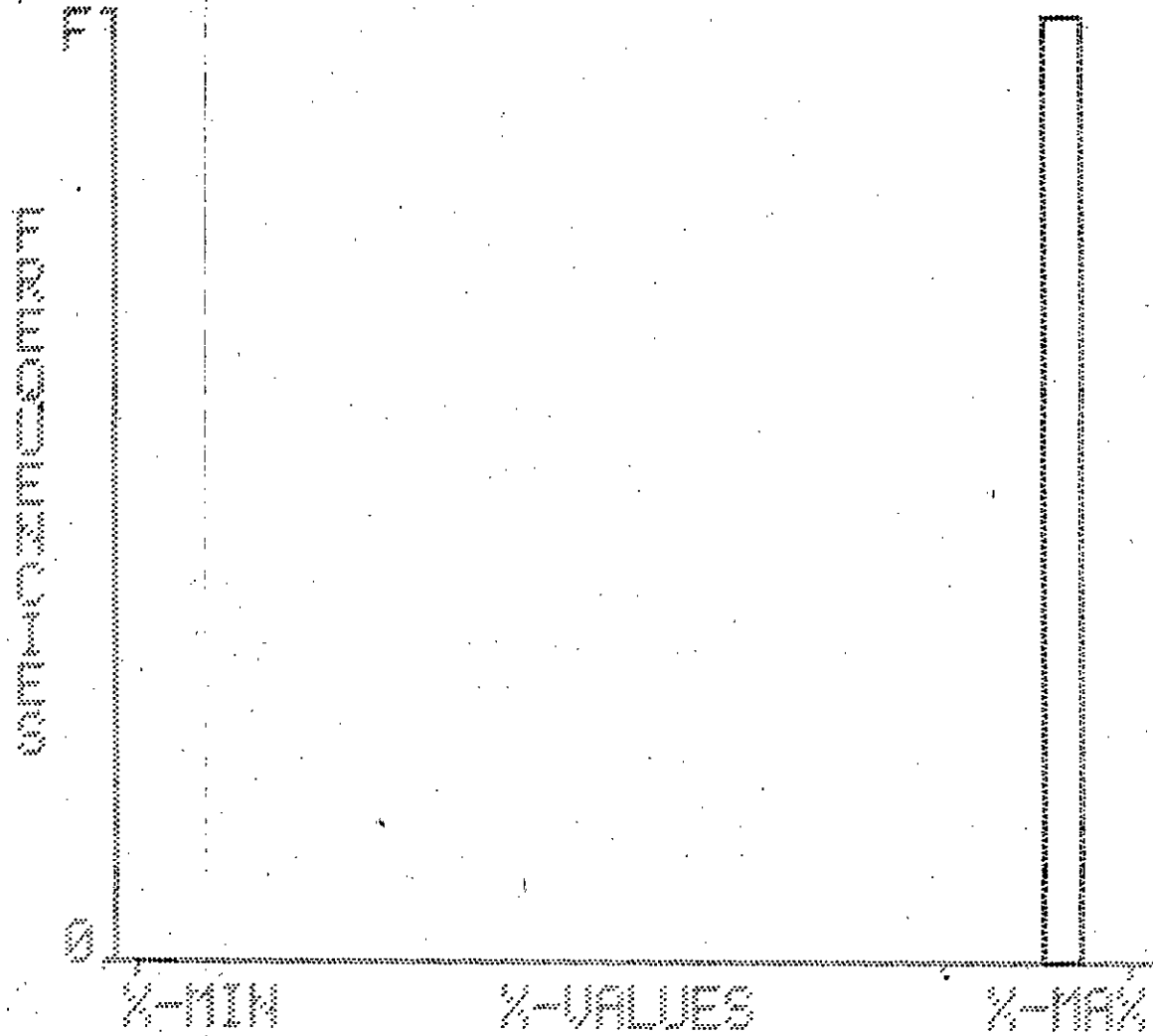
4.50000

10-MAY-1982 11:56 AM





HISTOGRAM OF  
 SAMPLE : 06-7, Act-1, poi 20



FREQUENCIES

BELOW PLOT

G.

ABOVE PLOT

G.

F

.99502

%-VALUES

%-MIN

-8949.19

%-MAX

500.810

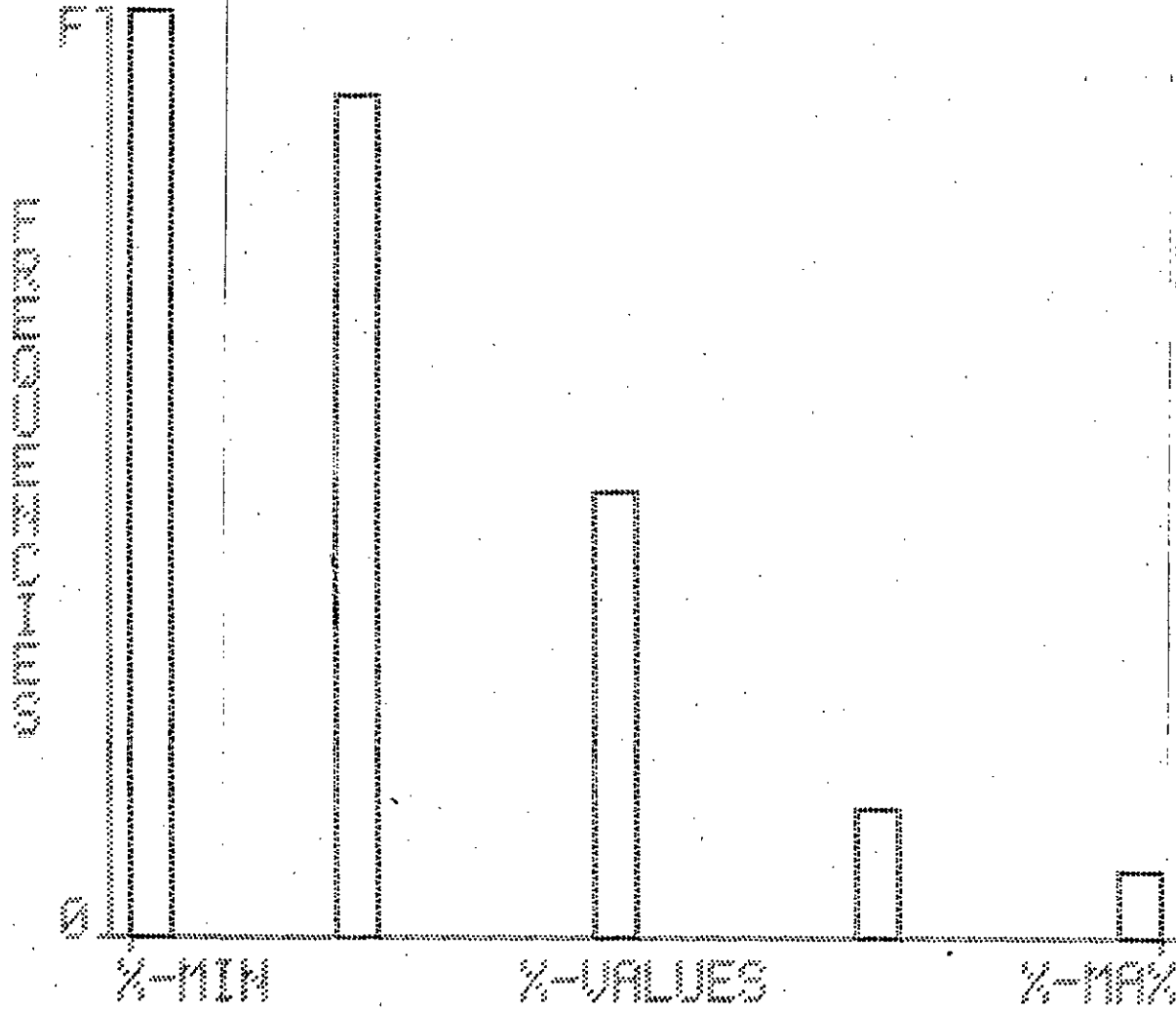
INTERVAL WIDTH

450.000

HISTOGRAM OF

SAMPLE

NO. 1, NO. 2, NO. 3, NO. 4, NO. 5



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

0.

F  
38308

%-VALUES

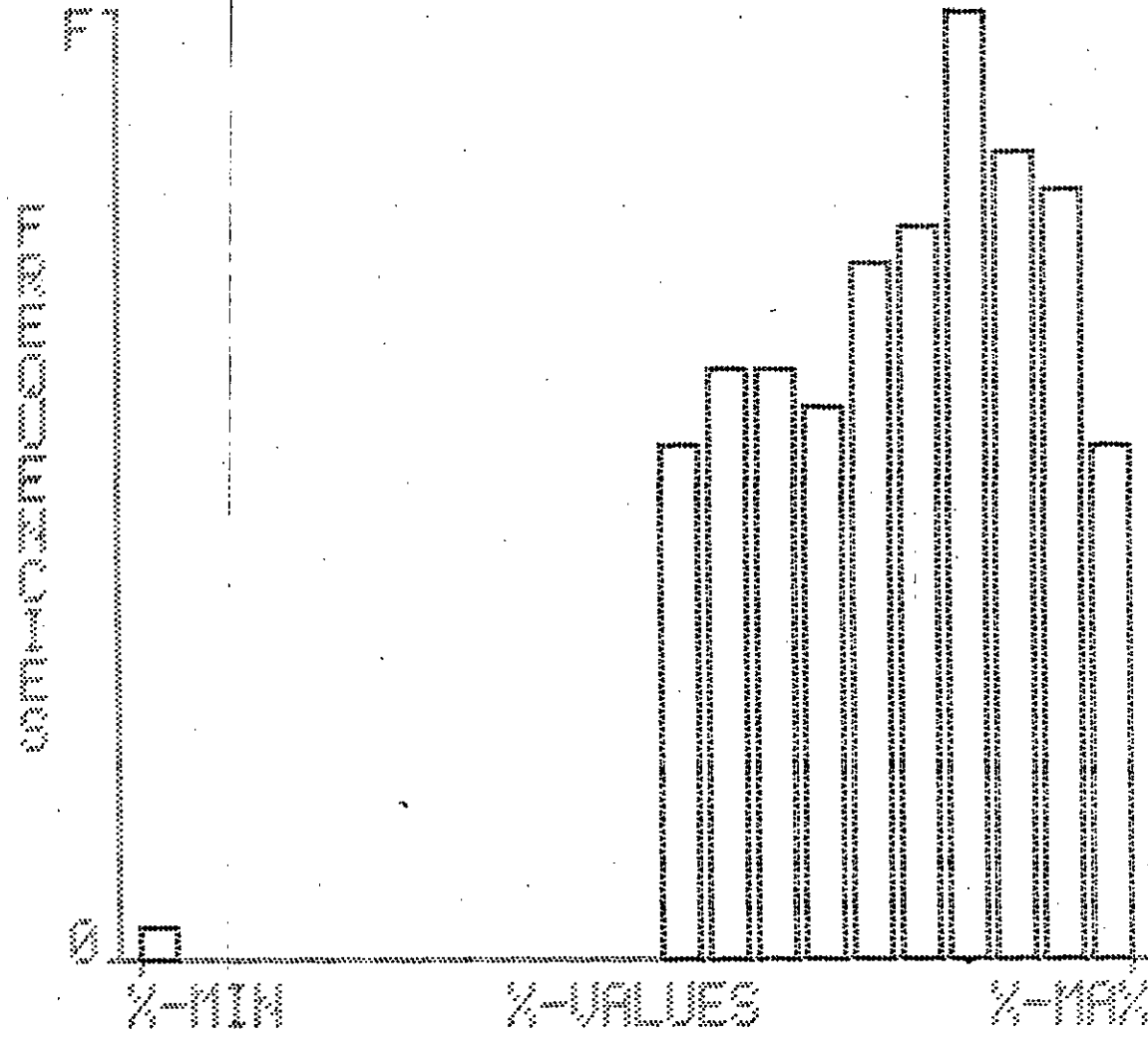
%-MIN  
- .05250

%-MAX  
25.9375

INTERVAL WIDTH

1.30000

HISTOGRAM OF  
 SAMPLE : 00-7, Oct-1, ran 200



FREQUENCIES  
 BELOW PLOT

0.

ABOVE PLOT

.08457

F

.12925

%-VALUES

%-MIN

0.

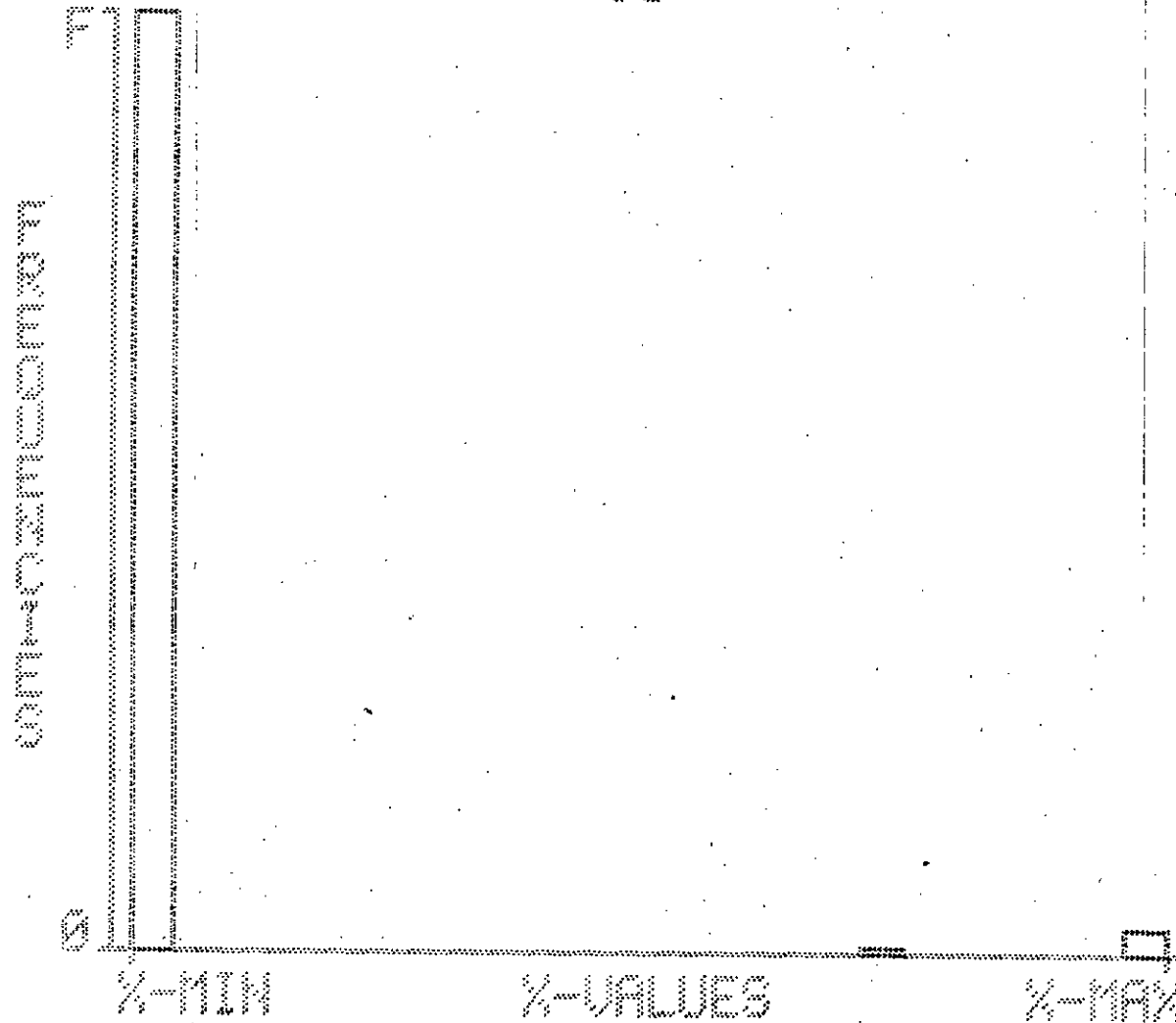
%-MAX

11.7800

INTERVAL WIDTH

.56000

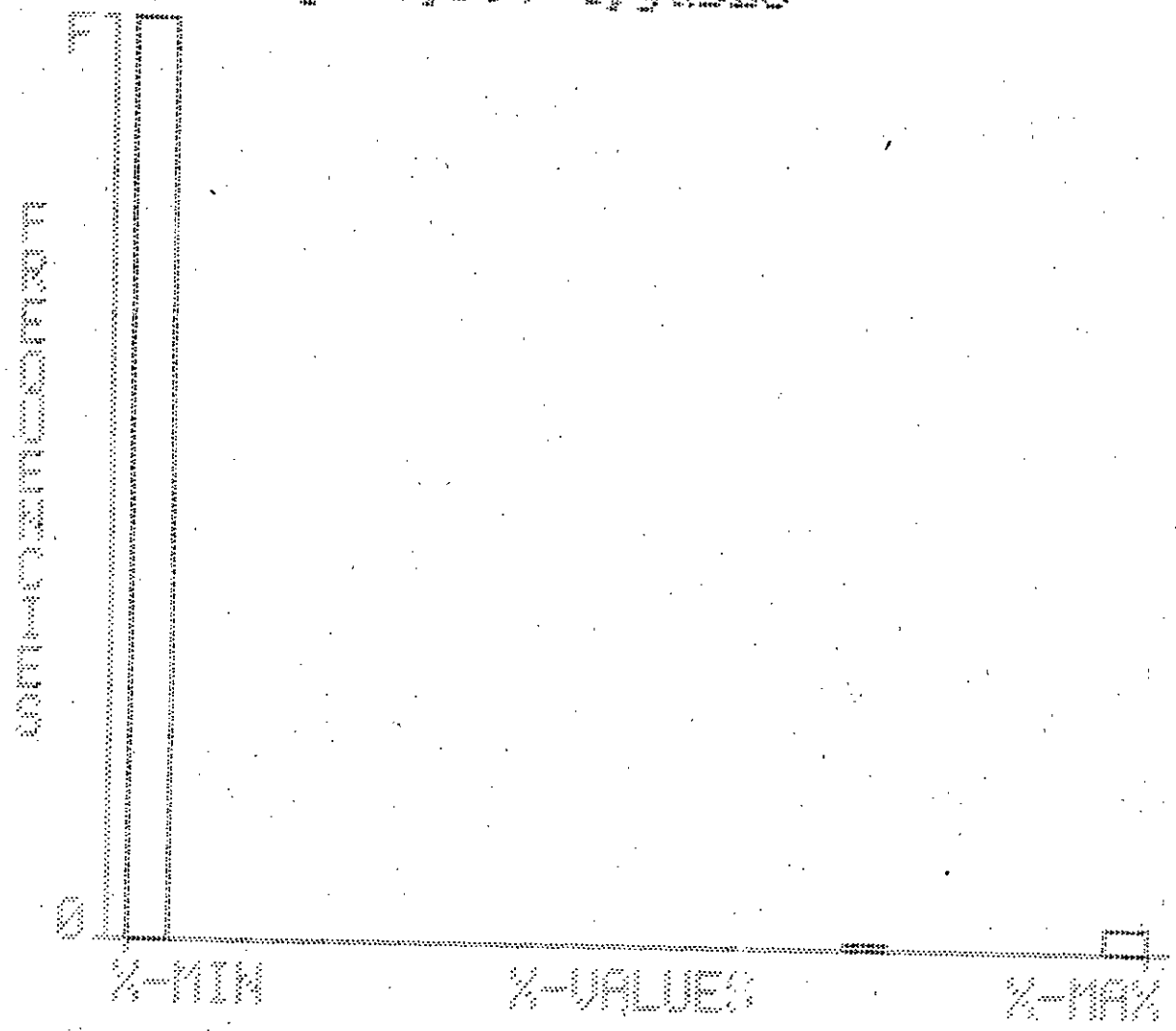
HISTOGRAM OF  
 SAMPLE : pb-7, fct-3, gas20



FREQUENCIES  
 -----  
 BELOW PLOT  
 0.  
 ABOVE PLOT  
 .50746  
 F  
 .47263  
 -----  
 %-VALUES  
 -----  
 %-MIN  
 0.  
 %-MAX  
 6.00000  
 -----  
 INTERVAL WIDTH  
 .20000

HISTOGRAM OF

SAMPLE : nb-7, fct-4, gas20



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

.48258

F

.49751

%-VALUES

%-MIN

0.

%-MAX

8.00000

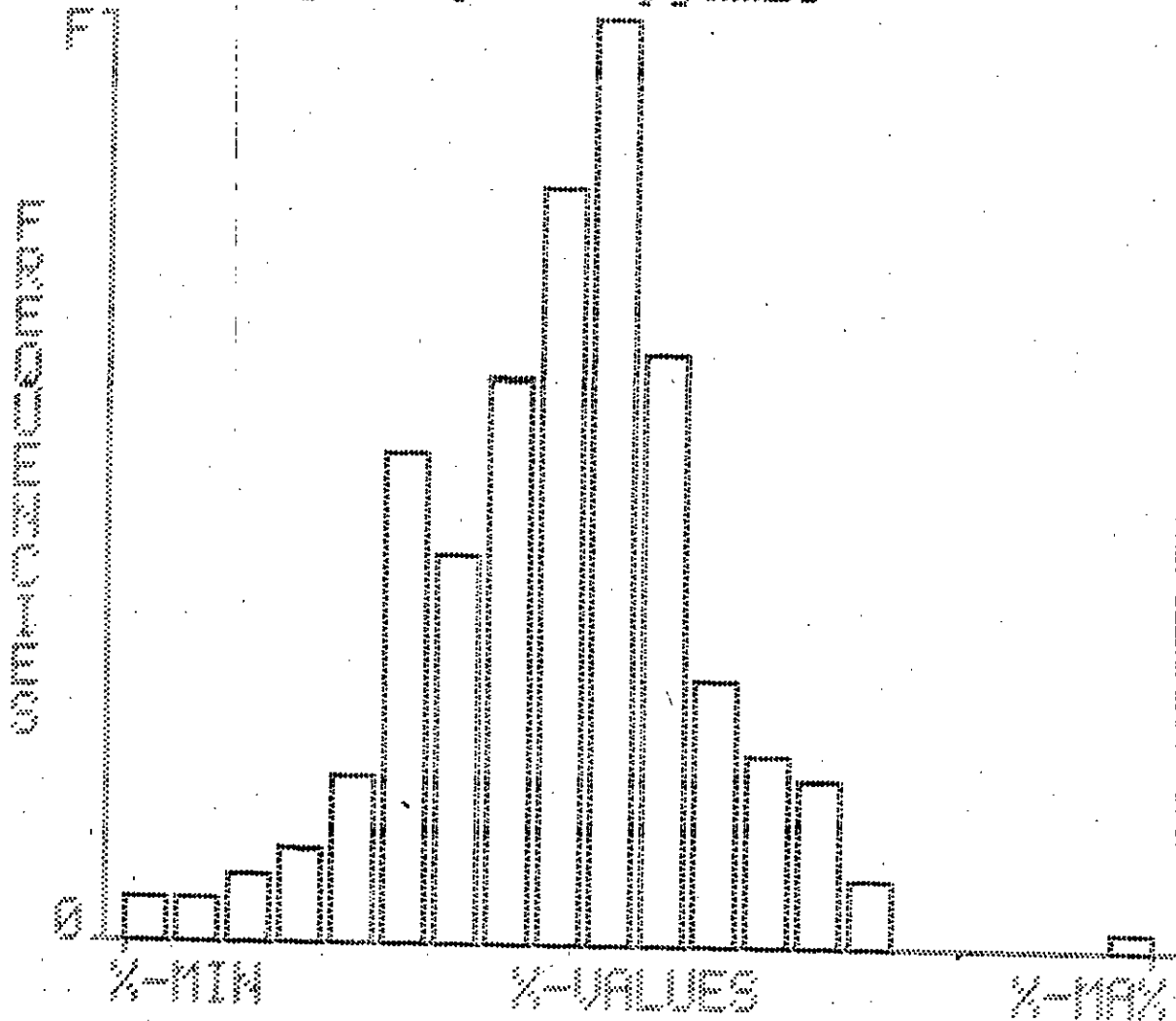
INTERVAL WIDTH

.300000



HISTOGRAM OF

SAMPLE # 101-10, lot 1, CAN 20



FREQUENCIES

BELOW PLOT

4.97512E-3

ABOVE PLOT

0.

F

.12905

%-VALUES

%-MIN

-40.9552

%-MAX

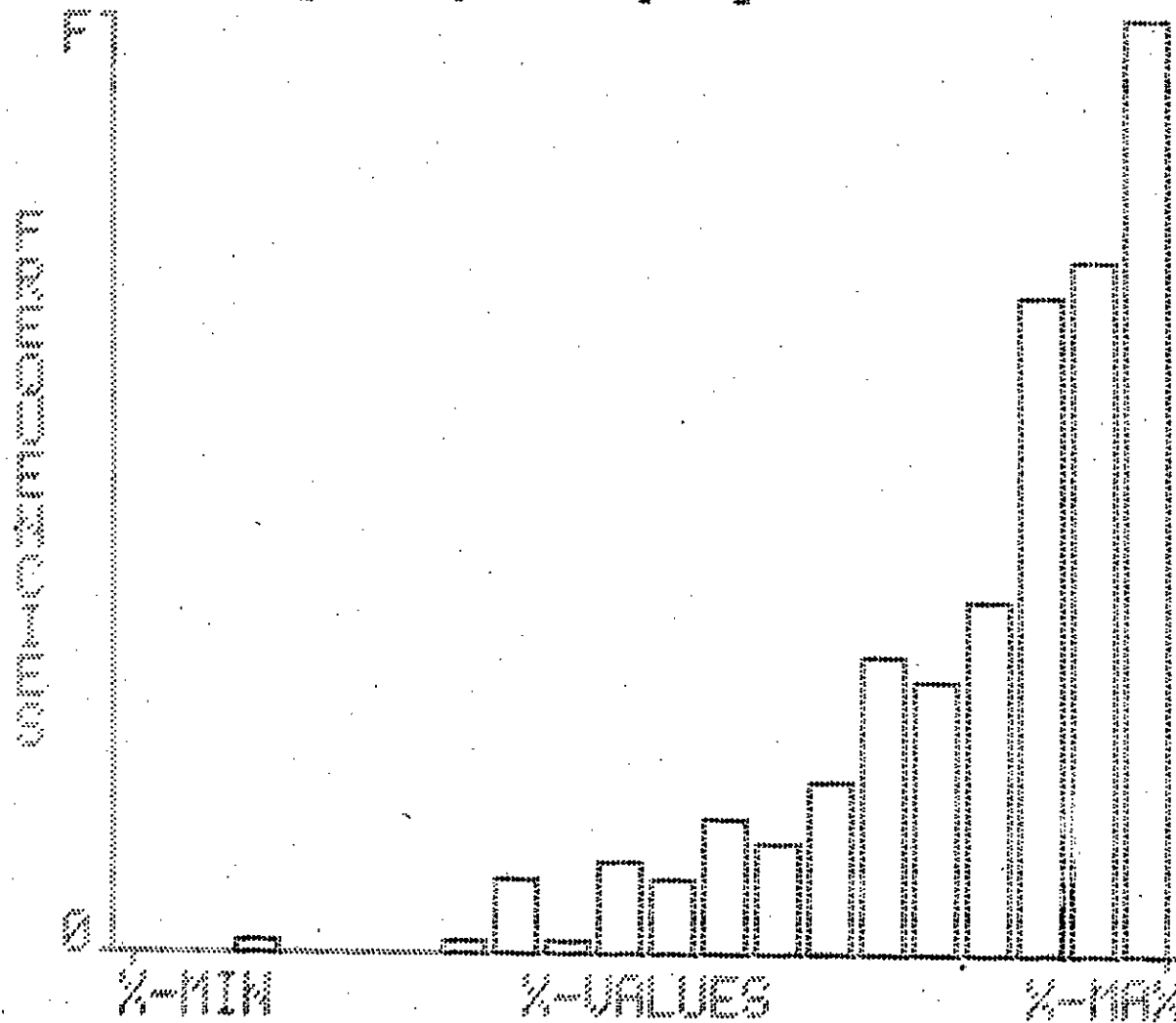
.04479

INTERVAL WIDTH

2.05000

16-MAY-1992 12:41 PM

HISTOGRAM OF  
 SAMPLE : 10-10, lot-1, exp20



FREQUENCIES

BELOW PLOT

4.97512E- 3

ABOVE PLOT

0.

F

.23383

%-VALUES

%-MIN

-94.8786

%-MAX

1.12140

INTERVAL WIDTH

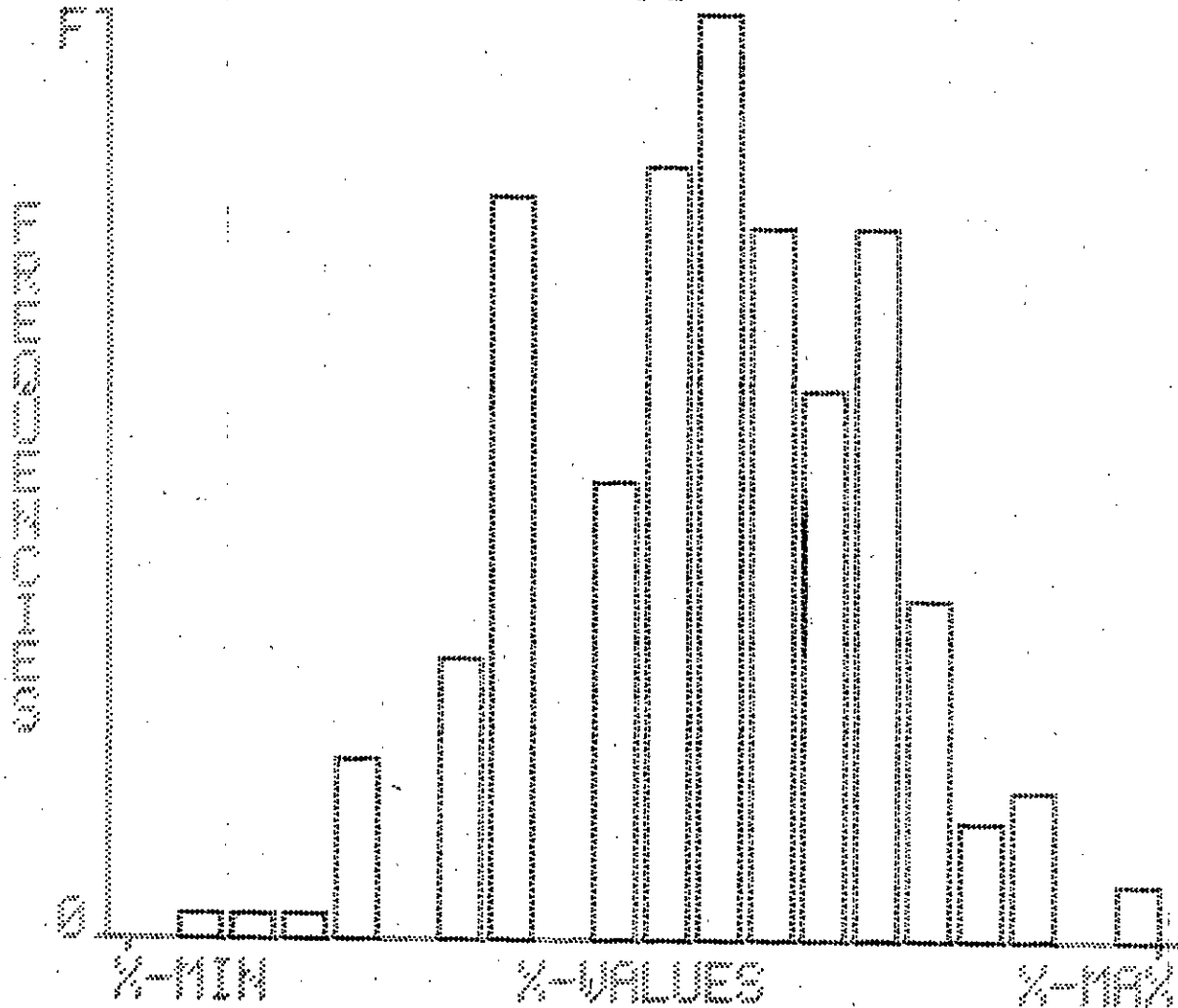
4.80000

16-MAY-1993 12:45 PM



HISTOGRAM OF

SAMPLE # 10-19, 10-1-1, 10-1-20



FREQUENCIES

BELOW PLOT

4.97512E- 3

ABOVE PLOT

9.95025E- 3

F

.14925

%-VALUES

%-MIN

-21.0000

%-MAX

-1.00000

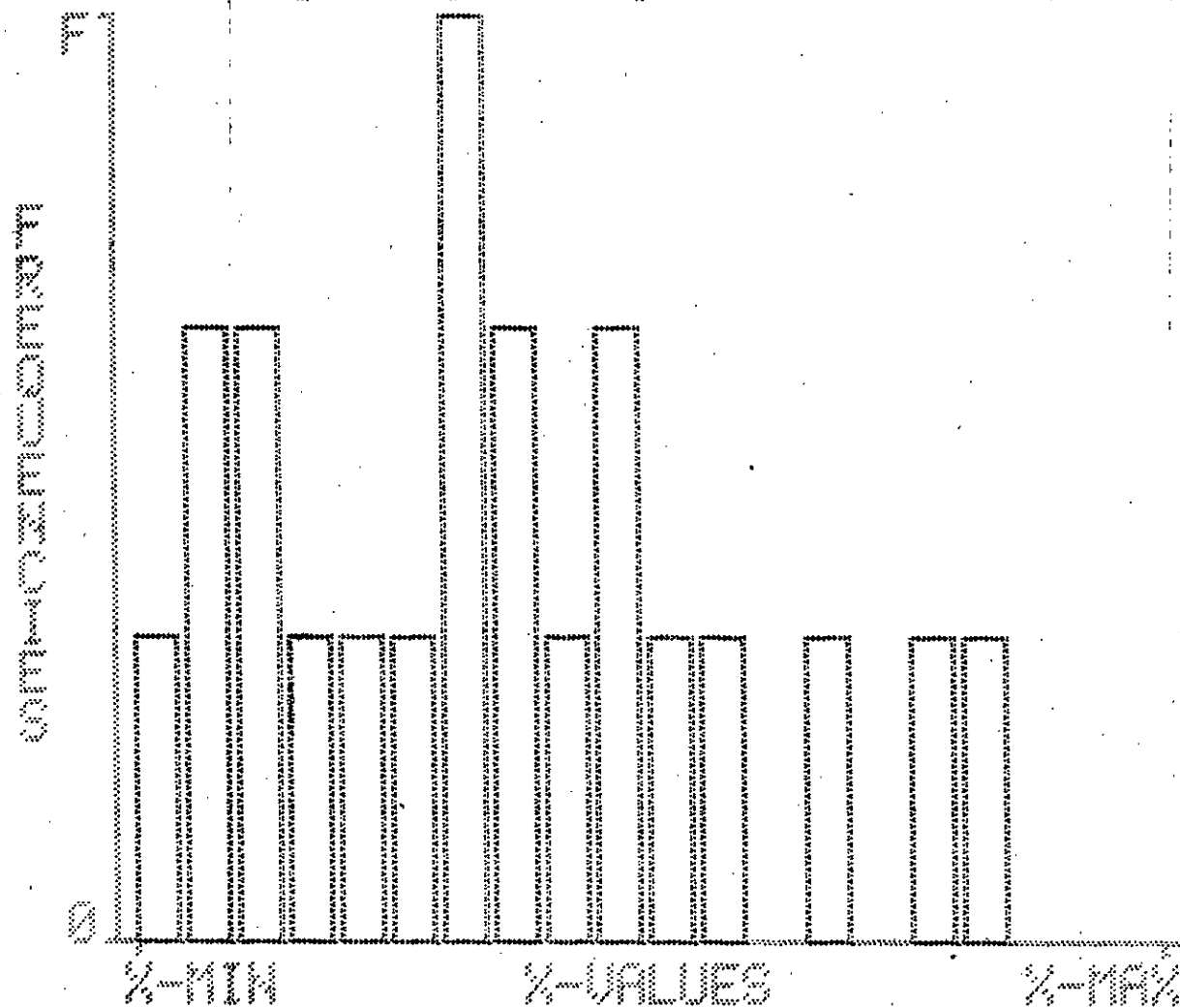
INTERVAL WIDTH

1.00000

10-MAY-1993 12:51 PM

# HISTOGRAM OF

SAMPLE # 10-10, 10-11, 10-12, 10-13



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

.79207

F

.02970

%-VALUES

%-MIN

-1.97000

%-MAX

-1.77000

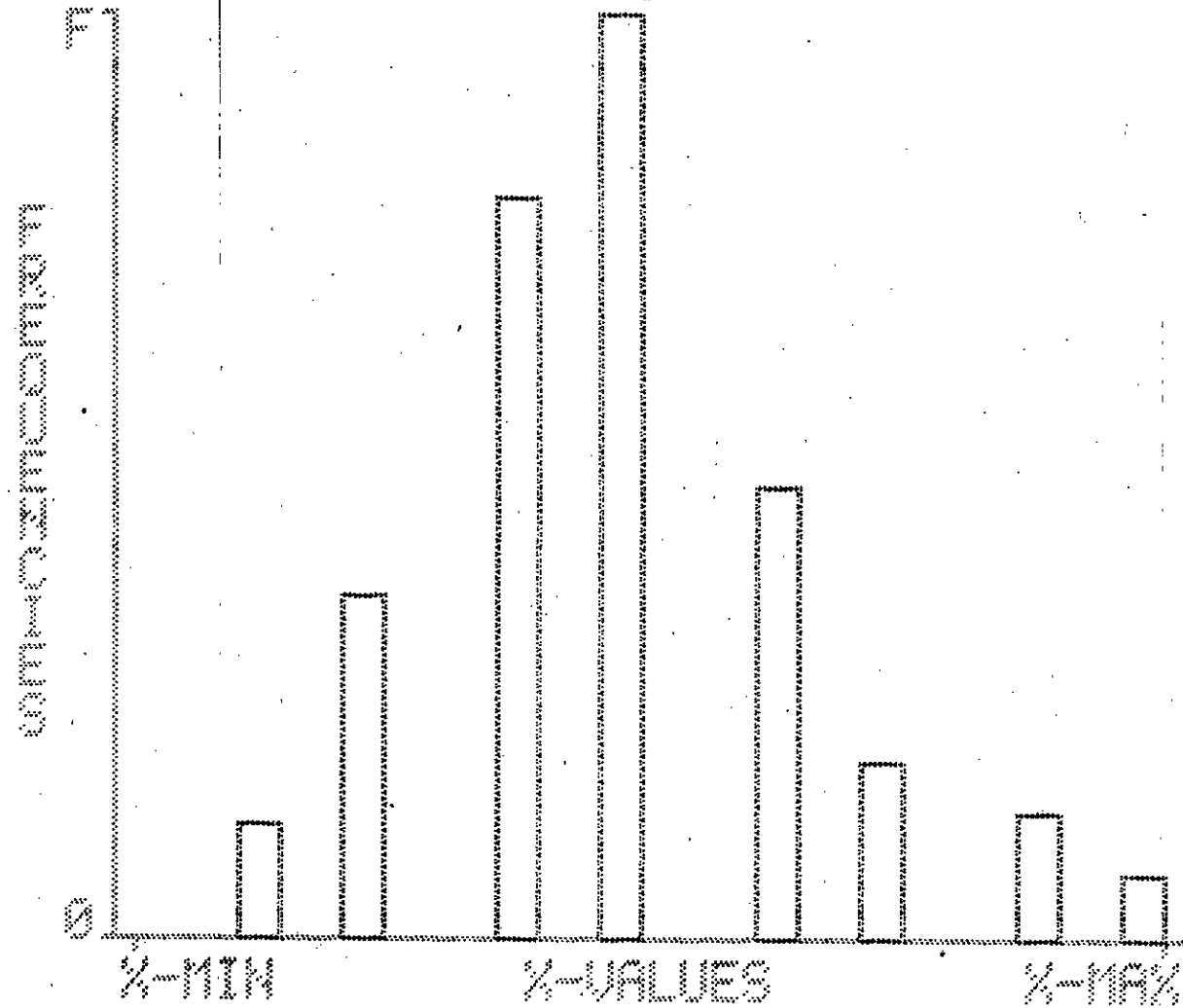
INTERVAL WIDTH

.01000

10-MAY-1993 12:53 PM

HISTOGRAM OF

SAMPLE \* 100 - 100, 100, 100, 100, 100, 100



FREQUENCIES

BELON PLOT

.01999

ABOVE PLOT

6.

F

.38348

%-VALUES

%-MIN

-7.99996

%-MAX

4.01735E- 5

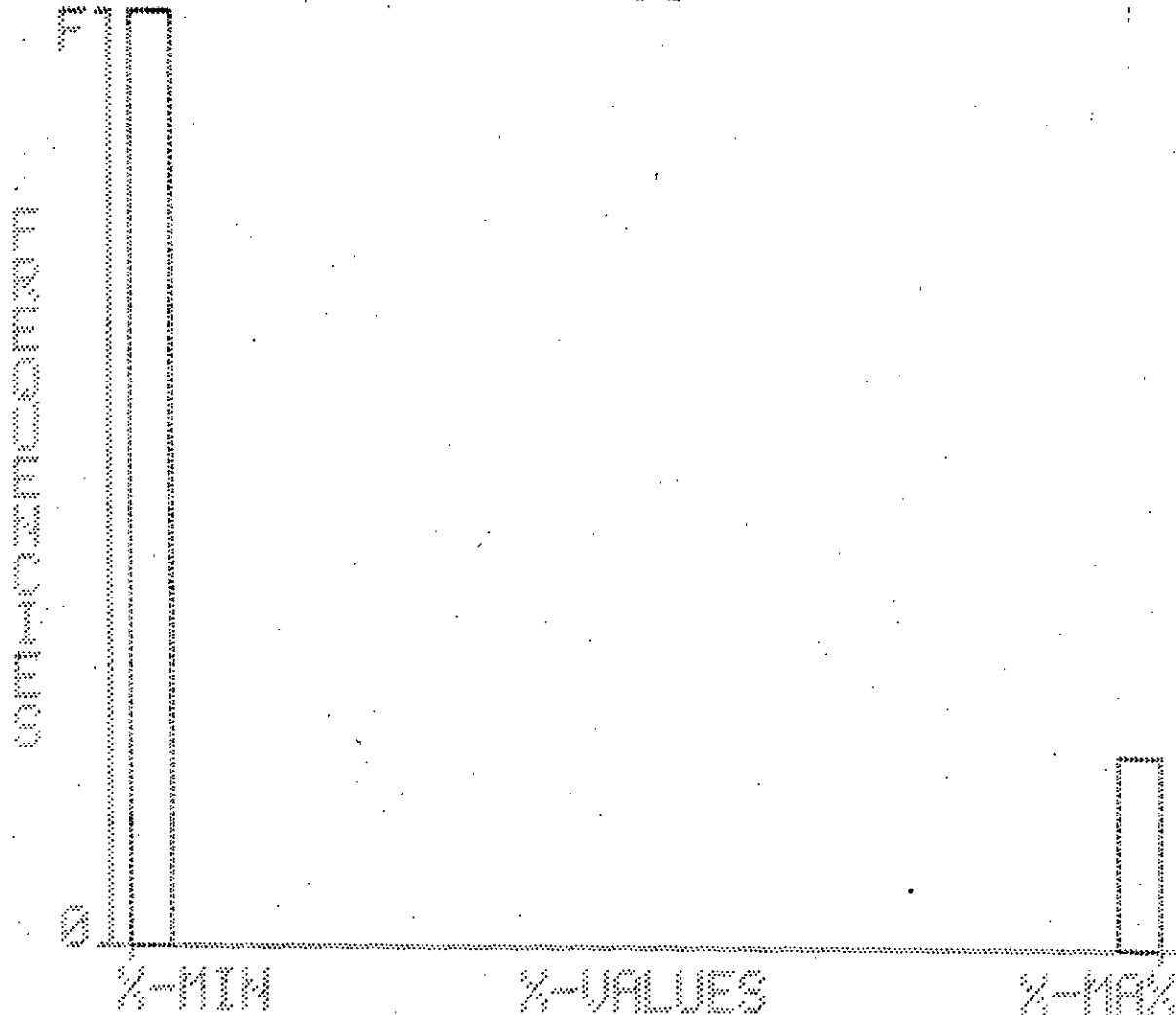
INTERVAL WIDTH

.40000

15-MAY-1982 12:55 PM

HISTOGRAM OF

SAMPLE " 10-11-10-1-1-10-10



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

0.

F

.02178

%-VALUES

%-MIN

0.

%-MAX

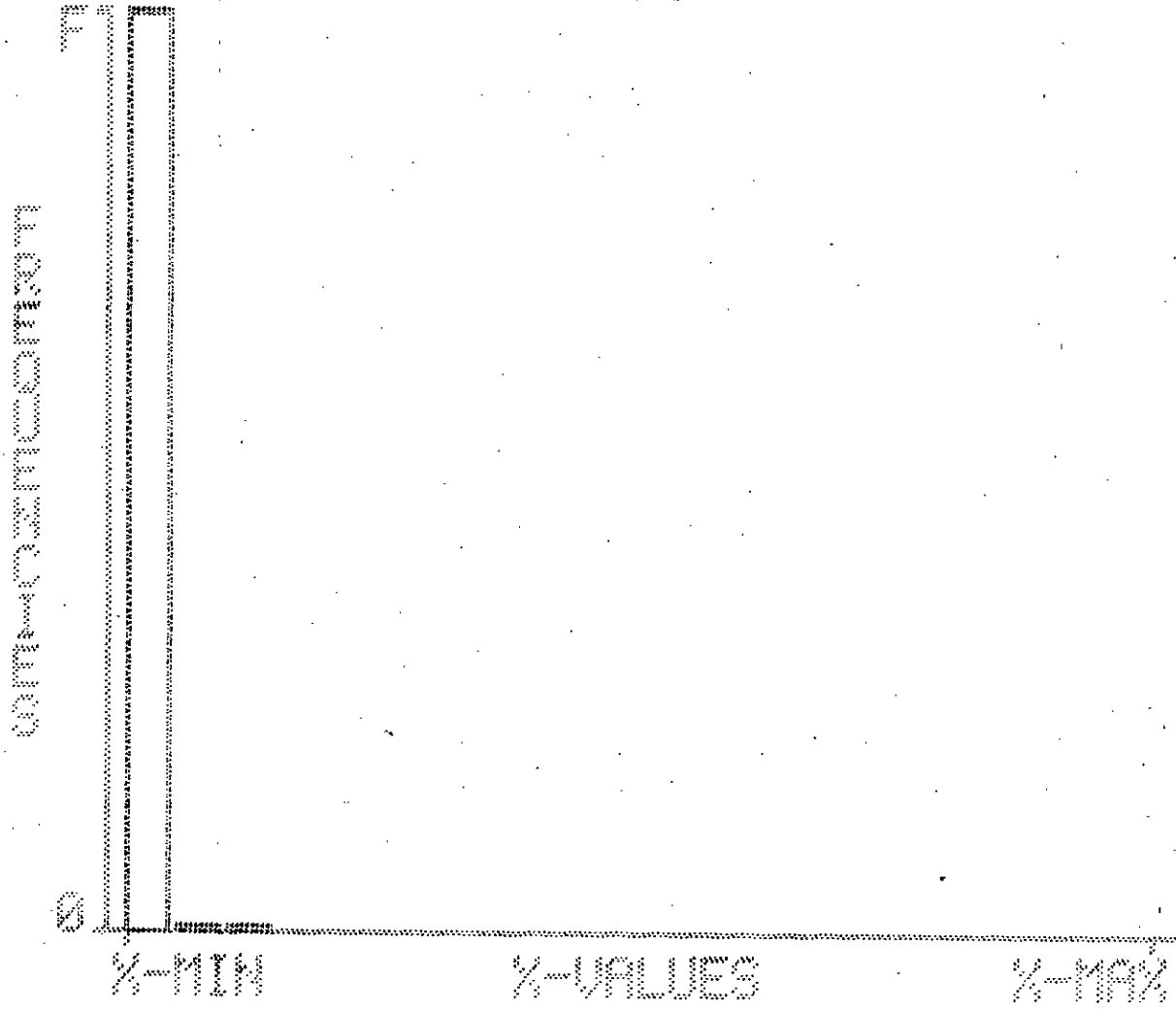
118.000

INTERVAL WIDTH

5.80000

HISTOGRAM OF

SAMPLE # 01-11, Act-2, Gas 18



FREQUENCIES

BELOW PLOT

0.

ABOVE PLOT

0.

F

1.40000

%-VALUES

%-MIN

0.

%-MAX

1.40000

INTERVAL WIDTH

.07000

