

**INSTITUT NATIONAL AGRONOMIQUE  
PARIS - GRIGNON**

**INTRODUCTION  
A LA THÉORIE DE LA DÉCISION  
ET  
A LA STATISTIQUE BAYÉSIENNE**

**1<sup>er</sup> Année**

*par*

**G. LEFORT**

PREMIERE PARTIE

ELEMENTS DE THEORIE DE LA DECISION

METHODES BAYESIENNES

Cette partie définit les principes de base et les méthodes de la théorie de la décision et de la statistique bayésienne. Les règles d'utilisation générale qui y sont données seront appliquées dans la deuxième partie à des cas particuliers importants ; ces cas particuliers peuvent donc servir d'exemples dans l'étude de la première partie.



Un atelier fournit chaque jour  $N$  pièces utilisées chacune dans le montage d'un appareil. Un lot contient une proportion inconnue  $\theta$  de défectueux et deux actions sont possibles :

- $a_1$  contrôler chaque pièce et éliminer les défectueuses
- $a_2$  procéder directement au montage et en cas de non fonctionnement de l'appareil démonter la pièce défectueuse qui sera remplacée le lendemain.

Les coûts  $c_1$  de contrôle d'une pièce et  $c_2$  de remplacement d'une pièce défectueuse après montage sont connus.

Que la décision prise soit  $a_1$  ou  $a_2$ , le nombre total d'appareils montés est  $N(1 - \theta)$  ; les coûts dus aux pièces défectueuses sont :

$$\begin{aligned} \text{pour } a_1 \quad c(a_1, \theta) &= Nc_1 \\ \text{pour } a_2 \quad c(a_2, \theta) &= N\theta c_2 \end{aligned}$$

Si  $\theta$  était connu, on choisirait l'action de coût minimum et donc on déciderait  $a_2$  dans le cas où :

$$\theta \leq c_1 / c_2$$

Mais  $\theta$  est inconnu ; pour choisir entre  $a_1$  et  $a_2$ , il faut utiliser l'information dont on peut disposer, par exemple les valeurs obtenues pour ce paramètre dans les lots précédents (nous considérons le cas où un échantillon de  $n$  pièces du lot a été contrôlé au § III).

#### 1ère méthode : critère minimax

Si on sait que  $0 \leq \theta \leq \theta_0$ , on étudie les coûts maximums :

$$\sup_{\theta} c(a_1, \theta) = Nc_1 \qquad \sup_{\theta} c(a_2, \theta) = N\theta_0 c_2$$

et on choisit l'action minimisant le coût maximum ; on décidera donc  $a_2$  si

$$\theta_0 \leq c_1 / c_2$$

En utilisant ce critère, on se place dans les conditions les plus défavorables (on donne à la proportion  $\theta$  de défectueux sa valeur la plus grande  $\theta_0$ ) ; si  $\theta_0$  dépasse  $c_1/c_2$ , l'action retenue est la plus coûteuse pour toutes les valeurs de  $\theta$  inférieures à  $c_1/c_2$ .

### 2ème méthode : utilisation de la distribution empirique

La distribution empirique de la proportion  $\theta$  dans les lots précédents définit une loi de probabilité sur  $\theta = [0,1]$  des valeurs possibles de  $\theta$  : accepter cette loi de probabilité pour l'étude du lot considéré revient à considérer qu'une valeur de  $\theta$  est d'autant plus probable qu'elle est apparue plus fréquemment dans les lots précédents.

Dans ces conditions, le coût de l'action  $a_2$ , fonction réelle de  $\theta$  :

$$\theta \rightarrow c(a_2, \theta) = N\theta c_2$$

devient une aléatoire réelle  $C_{a_2}$  et nous déciderons  $a_2$  si l'espérance de cette aléatoire est inférieure au coût  $Nc_1$  de  $a_1$ , c'est-à-dire,  $\theta_m$  désignant l'espérance de la loi de probabilité définie sur  $\theta$ , si

$$N\theta_m c_2 \leq Nc_1 \Leftrightarrow \theta_m \leq c_1/c_2$$

Le remplacement du coût aléatoire  $C_{a_2}$  par son espérance se justifie parce que nous avons ici une opération répétitive ; on sait (loi des grands nombres) que la moyenne des coûts sur un assez grand nombre de jours tend en probabilité vers l'espérance de  $C_{a_2}$  ; c'est donc le coût moyen que nous avons minimisé.



## II - TRAITEMENT MATHÉMATIQUE D'UN PROBLÈME DE DÉCISION

### 1 - DONNÉES DU PROBLÈME

Pour l'étude mathématique du choix d'une action, il faut définir :

- l'ensemble  $A$  des actions envisagées ;
- l'ensemble  $\Theta$  des paramètres  $\theta$  (le paramètre  $\theta$  est un réel ou, plus généralement un élément  $(\theta_1, \dots, \theta_k)$  de  $\mathbb{R}^k$  ; il définit les états possibles du système étudié) ;
- la fonction de coût  $c$  qui est une application bornée de l'ensemble produit  $A \times \Theta$  dans  $\mathbb{R}$  (le nombre  $c(a, \theta)$  est le coût positif ou négatif de l'action  $a$  pour la valeur  $\theta$  du paramètre).

La notion de coût doit être prise au sens le plus large et non seulement en termes économiques ; ainsi, pour un médecin, le coût d'un traitement traduira les contre-indications éventuelles, les arrière-effets nuisibles et, avec un signe moins puisqu'il s'agit d'un gain, l'amélioration de l'état du malade qu'on peut en attendre.

## 2 - ACTIONS ADMISSIBLES ; ACTION OPTIMALE

La fonction de coût définit un préordre partiel sur l'ensemble A des actions : a' est meilleure que a (ou  $a < a'$ ) si :

$$\forall \theta, c(a', \theta) \leq c(a, \theta)$$

Cette relation est un préordre et non un ordre parce que

$$a < a' \text{ et } a' < a \Leftrightarrow a = a'$$

Une action a strictement moins bonne qu'une action a' c'est-à-dire telle que :

$$\forall \theta, c(a', \theta) < c(a, \theta) \text{ et } \exists \theta_0 : c(a', \theta_0) < c(a, \theta_0)$$

peut être éliminée. Seules sont donc intéressantes les actions qui ne sont strictement inférieures à aucune autre, dites actions admissibles.

S'il existe un plus grand élément  $a_m$  dans A c'est-à-dire si :

$$\forall \theta, \forall a, c(a_m, \theta) \leq c(a, \theta)$$

l'action  $a_m$  est optimale et doit être choisie (s'il existe plusieurs plus grands éléments l'un quelconque d'entre eux peut être choisi). Cette situation se présente en particulier si  $\theta$  contient un seul élément (si le paramètre  $\theta$  est connu) et si l'ensemble des valeurs de la fonction c est un intervalle fermé.



### 3 - CRITERES DE DECISION ; CRITERE MINIMAX

S'il n'existe pas dans  $A$  de plus grand élément , il est nécessaire de compléter la donnée du triplet  $(A, \theta, c)$  par un critère permettant de faire un choix dans l'ensemble des actions admissibles ; nous allons en donner deux exemples simples.

#### 4 - DECISION BAYESIENNE

La donnée du triplet  $(A, \theta, c)$  est complétée par celle d'une mesure de probabilité sur  $\theta$  ; cette mesure traduit la plus ou moins grande vraisemblance attribuée aux éléments  $\theta$  de  $\theta$  en fonction de l'information qu'on a sur le système étudié ; elle sera souvent appelée mesure de vraisemblance.

---



Pratiquement,  $\theta$  est un sous-ensemble de  $\mathbb{R}$  ou  $\mathbb{R}^k$  ; une mesure de probabilité sur  $\theta$  définit donc une aléatoire réelle ou vectorielle que nous noterons  $\underline{\theta}$  pour la distinguer de la valeur  $\theta$  du paramètre qui est un nombre inconnu mais certain.

A toute action  $a$ , on associe alors un coût aléatoire :

$$C_a = c(a, \underline{\theta})$$

(comme pour  $\underline{\theta}$ , la loi de probabilité de cette aléatoire traduit la plus ou moins grande vraisemblance attribuée aux valeurs possibles du coût). Le choix de l'action  $a$  retenue s'effectue en comparant les distributions des aléatoires  $C_a$  ; pour faire cette comparaison, il sera parfois commode de choisir un critère de décision.

Si l'action  $a$  retenue doit être répétée, le critère choisi sera en général l'espérance de  $C_a$  : décider à si

$$\forall a, \mathbb{E}[C_{\hat{a}}] \leq \mathbb{E}[C_a]$$

En effet, comme il a été indiqué dans l'exemple, ce choix signifie qu'on cherche l'action  $a$  dont le coût moyen est minimal. Mais il sera bon de compléter l'étude en déterminant la dispersion de  $C_{\hat{a}}$  (définie par la variance ou tout autre indice); de fortes variations du coût peuvent en effet compromettre la trésorerie d'une entreprise, même si en moyenne l'action est optimale.

## 5 - GENERALISATION : FONCTION DE COUT ALEATOIRE

Le coût d'une action  $a$  pour une valeur  $\theta$  du paramètre peut être aléatoire.

L'information relative au paramètre  $\theta$  est en général obtenue à partir de résultats expérimentaux  $x_1, \dots, x_n$  entachés d'erreurs aléatoires.

Ainsi dans l'exemple I, on peut déterminer le nombre  $x$  de défectueux dans un échantillon aléatoire de  $n$  pièces extrait du lot utilisé ; pour étudier l'espérance du rendement d'une variété dans des conditions déterminées, on peut mesurer les rendements  $x_1, \dots, x_n$  de cette variété dans  $n$  parcelles différentes. La prise en compte des résultats expérimentaux dans la définition de la mesure de vraisemblance sur  $\theta$  se fait en utilisant le théorème de Bayes.

#### 1- MODELE PROBABILISTE DE L'EXPERIENCE

L'expérience effectuée est considérée comme une réalisation particulière d'une épreuve aléatoire susceptible de répétitions et le résultat  $x = (x_1, \dots, x_n)$  est la valeur prise dans cette réalisation par une aléatoire  $X = (X_1, \dots, X_n)$  associée à l'épreuve.

L'analyse des conditions expérimentales permet de définir pour toute valeur du paramètre  $\theta$  la loi de probabilité de l'épreuve (ou modèle probabiliste de l'expérience) et donc la loi  $L(\theta)$  de l'aléatoire  $X = (X_1, \dots, X_n)$ . Nous désignerons par  $p(x|\theta)$  la probabilité (ou la densité) au point  $x$  de la loi  $L$  discrète (ou continue).

Dans l'exemple, l'épreuve est le tirage d'un échantillon de  $n$  pièces et l'aléatoire  $X$  associée le nombre de défectueux dans l'échantillon ; sous l'hypothèse que chaque pièce du lot a la même probabilité d'appartenir à l'échantillon,



## 2- MESURE DE VRAISEMBLANCE A PRIORI

La connaissance antérieure du système étudié (ainsi, dans l'exemple, les proportions de défectueux relevées les jours précédents) permet de définir une mesure de vraisemblance sur  $\theta$  préalable à l'expérience dite mesure à priori. Pour des raisons de simplicité, la mesure a priori sera souvent choisie dans une famille de distributions particulières dépendant de 2 paramètres au moins de manière à pouvoir fixer la valeur centrale et la dispersion (par exemple l'espérance et la variance). La dispersion sera d'autant plus grande que l'information préalable sera moins précise et, en passant à la limite, on se placera dans le cas où l'information préalable est négligeable et où seul compte

### 3 - MESURE DE VRAISEMBLANCE A POSTERIORI

Le problème est de modifier la vraisemblance a priori pour tenir compte de l'information supplémentaire fournie par le résultat expérimental ; grâce à la formulation mathématique que nous venons de donner, c'est un problème élémentaire de calcul des probabilités.

La loi du couple  $(X, \underline{\theta})$  est bien déterminée par la loi marginale de  $\underline{\theta}$  (vraisemblance a priori définie en 2) et par la loi conditionnelle de  $X$  sachant  $\underline{\theta} = \theta$  (loi  $L(\theta)$  définie par le modèle probabiliste de l'expérience en 1). La loi cherchée est la loi de  $\underline{\theta}$  après expérience c'est-à-dire la loi conditionnelle de  $\underline{\theta}$  sachant  $X = x$  dite loi a posteriori.

La probabilité (ou densité) a posteriori notée  $p_1(\theta|x)$  est donnée par le théorème de Bayes :

$$p_1(\theta|x) \propto p_0(\theta) p(x|\theta)$$



#### 4 - REPRESENTATION DE LA DISTRIBUTION A POSTERIORI

Nous considérerons le seul cas où la loi a posteriori sur  $\theta \in \mathbb{R}^k$  est définie par une densité  $p_1(\theta|x)$ .

A cette distribution, comme à toute distribution définie par une densité, on peut associer :

des ensembles de probabilité fixée  $\Pi$  (intervalles si  $k = 1$ , régions du plan ou de l'espace si  $k = 2$  ou  $3$ ) ; si la valeur choisie pour  $\Pi$  est assez grande, il est très vraisemblable que  $\theta$  appartienne à l'ensemble correspondant ; par exemple, en admettant que le paramètre inconnu  $\theta$  appartient à un ensemble de probabilité 0,95, on accepte une probabilité d'erreur, ou **risque**, de 0,05 (en moyenne on se trompe une fois sur 20) ;

des ensembles de densité constante  $S_p = \{\theta \mid p_1(\theta|x) = p\}$  (courbes si  $k = 2$ , surfaces si  $k = 3$ ) ; en traçant plusieurs de ces "courbes de niveau", avec pour chacune la valeur de  $p$ , on a une bonne représentation de la densité.

#### Région minimale de vraisemblance donnée

L'ensemble de vraisemblance  $\Pi$  et de grandeur minimum (longueur si  $k = 1$ , aire si  $k = 2$ , volume si  $k = 3, \dots$ ) est

$$V_{\Pi} = \{\theta \mid p_1(\theta|x) \geq p\}$$



## 5 - ESTIMATION PONCTUELLE

Il est parfois demandé de fournir pour le paramètre  $\theta$  une valeur unique appelée estimation ponctuelle de  $\theta$ , c'est-à-dire de résumer en un nombre unique l'information fournie par la distribution de  $\underline{\theta}$  ; il faut donc choisir une valeur centrale de cette distribution ; mais il est indispensable d'accompagner cette valeur d'indications sur la dispersion de la distribution (si  $\theta$  est réel, on donnera la longueur d'un intervalle  $I_{\Pi}$  ou  $V_{\Pi}$ , ou l'écart-type ou la variance de  $\underline{\theta}$ .)

Une valeur naturelle est le mode  $\hat{\theta}$  de la distribution a posteriori de  $\underline{\theta}$  (la valeur  $\hat{\theta}$  maximisant la probabilité ou la densité) c'est-à-dire la valeur la plus vraisemblable (on suppose évidemment la distribution unimodale, ce qui est le cas général) ; mais on peut aussi choisir l'espérance de  $\underline{\theta}$  .

Dans le cas où la densité a priori est sensiblement constante, le mode  $\hat{\theta}$  est la valeur qui maximise la vraisemblance expérimentale  $p(x|\theta)$  ; cette valeur particulière est appelée estimation du maximum de vraisemblance (sous entendu expérimentale).

## 6 - DISTRIBUTION APPROCHEE DE $\theta$ DANS LE CAS DES GRANDS ECHANTILLONS

L'étude de la distribution a post $\acute{e}$ riori de  $\theta$  peut  $\hat{e}$ tre malcommode, par exemple si les tables num $\acute{e}$ riques font d $\acute{e}$ faut ; d'o $\grave{u}$  l'int $\acute{e}$ r $\hat{e}$ t de lois approch $\acute{e}$ es simples.

Si le nombre  $n$  des observations est grand, on admet que les r $\acute{e}$ sultats asymptotiques (obtenus  $\grave{a}$  la limite si  $n$  tend vers  $+\infty$ ) donnent des approximations acceptables ; en particulier, si les lois a priori et a post $\acute{e}$ riori de  $\theta$  sont d $\acute{e}$ finies par des densit $\acute{e}$ s  $p_0$  et  $p_1$ , on peut donner les r $\acute{e}$ sultats suivants :



7 - DISCUSSION D'HYPOTHESES SPECIFIANT LA VALEUR DU PARAMETRE  $\theta$  OU DE FONCTIONS REELLES DE CE PARAMETRE (Tests d'hypothèse)

1er problème :

L'hypothèse faite est que le paramètre  $\theta$  a une valeur donnée  $\theta_0$ .

Si  $W$  est un ensemble de vraisemblance  $\Pi$ , on peut, en prenant un risque d'erreur  $1-\Pi$ , appelé niveau, affirmer que  $\theta$  appartient à  $W$ . On choisit donc le risque  $1-\Pi$ , et on définit l'ensemble  $W$ , de préférence l'ensemble minimal  $V_{\Pi}$  ou parfois dans  $\mathbb{R}$  l'intervalle  $I_{\Pi}$  (ces ensembles ont été définis en 4) ; si le point  $\theta_0$  n'appartient pas à  $W$ , l'hypothèse est rejetée au niveau  $1-\Pi$ .

Remarque : Dire que la valeur  $\theta_0$  est acceptable ne signifie pas que le paramètre  $\theta$  a sûrement cette valeur ; en effet tous les autres points de  $W$  sont aussi des valeurs acceptables. La démarche faite n'a donc d'intérêt que si la valeur  $\theta_0$  est particulièrement intéressante :

- soit pour des raisons de simplicité,
- soit par la connaissance qu'on a du système étudié.

2ème problème :

L'hypothèse faite est que  $l$  fonctions réelles du paramètre  $\theta$  algébriquement indépendantes sont nulles, soit :

$$f_1(\theta) = f_2(\theta) = \dots = f_l(\theta) = 0$$

On peut remplacer le paramètre  $\theta = (\theta_1, \dots, \theta_l)$  par les  $l$  paramètres réels

on rejettera donc l'hypothèse au niveau  $1-\Pi$  si l'origine n'appartient pas à  $W$ . La remarque faite pour le premier problème vaut également ici.

### 8 - CHOIX D'UN ENSEMBLE $\theta_i$ DANS UNE PARTITION FINIE $\theta_1, \dots, \theta_1$ DE $\theta$

Nous considérerons le seul cas où la distribution a postériori est définie par une densité  $p(\theta|x)$  et où chacun des ensembles  $\theta_1, \dots, \theta_1$  contient une boule (nous excluons ainsi le cas où un des ensembles contiendrait un nombre fini de points ou, dans le plan, se réduirait à une courbe).



Trois règles de décision peuvent être envisagées.

Si une décision doit être prise :

Règle 1

Choisir l'ensemble  $\theta_i$  de vraisemblance maximum.

Si la décision peut être différée :

Règle 2

Choisir l'ensemble  $\theta_i$  de vraisemblance maximum si sa vraisemblance dépasse un seuil élevé  $\Pi$  (0,90 , 0,95 ou 0,99 par exemple) et différer la décision dans le cas contraire.

Si le choix d'un des ensembles,  $\theta_1$  par exemple, n'implique pas d'action et peut être révisé sans difficulté à la différence du choix des autres ensembles, on identifie choisir  $\theta_1$  et différer la décision ; d'où, en application de la règle 2 :

Règle 3

Choisir l'ensemble  $\theta_i$  si sa vraisemblance dépasse un seuil  $\Pi$  élevé (largement supérieur à 0,5),  $\theta_1$  si aucun ensemble n'a une vraisemblance supérieure à  $\Pi$ .

On notera que la règle 3 n'a de valeur que si le choix de  $\theta_1$  est provisoire et soumis à révision ; sinon l'ensemble  $\theta_1$  serait abusivement privilégié.

Ces règles peuvent être retrouvées dans le cadre général de la théorie de la

## Discussion de l'intérêt d'une expérience

Dans le cas où on a choisi d'utiliser la règle 2 (ou la règle 3), il faut, avant d'effectuer une expérience calculer la probabilité d'obtenir une décision positive c'est-à-dire le choix d'un ensemble  $\theta_i$  et non le report de la décision (dans le cas de la règle 3, il faut alors dissocier  $\theta_i$  et l'absence de décision). Si plusieurs expériences sont possibles, on préférera celle qui donne la plus grande probabilité d'obtenir une décision positive.

Dans cette étude, on connaît donc la mesure a priori sur  $\theta$  et le modèle probabiliste c'est-à-dire la loi  $L(\theta)$  de  $X$  mais non la valeur  $x$ .

La vraisemblance a postérieure de  $\theta_i$  est une fonction de  $x$  et  $\theta$  ; si  $\theta$  est fixé, on peut déterminer l'ensemble  $A_i(\theta)$  des valeurs  $x$  pour lesquelles cette vraisemblance de  $\theta_i$  dépasse le seuil  $\Pi$  et donc  $\theta_i$  est choisi. La probabilité de choisir  $\theta_i$  si le paramètre a la valeur  $\theta$  est :

$$v_i(\theta) = \text{Prob } A_i(\theta)$$

la probabilité étant calculée pour la loi  $L(\theta)$  ; la probabilité d'une décision positive est :

$$v(\theta) = \sum_{i=1}^l v_i(\theta) = 1 - v_0(\theta)$$

où  $v_0$  est la probabilité de l'ensemble  $A_0$  complémentaire de la réunion des  $A_i$  ( $1 \leq i \leq l$ ). Les fonctions  $v_1, \dots, v_l$  sont appelées les composantes de la fonction puissance.

L'étude des fonctions  $v_i$  et de leur somme  $v$  permet de discuter l'intérêt de l'expérience envisagée selon les valeurs possibles de  $\theta$ .

Cette étude sera complétée par le calcul des probabilités a priori  $\bar{v}_i$  de dé-



## DEUXIEME PARTIE

### ETUDE DE MODELES SIMPLES

L'interprétation de données expérimentales nécessite, nous l'avons vu, le choix d'un modèle probabiliste. Ce choix est particulièrement important et ne doit pas être fait pour de simples raisons de commodité de calcul ; c'est l'analyse des conditions expérimentales qui permet de dégager les hypothèses admissibles et de définir ainsi le modèle.

Les modèles étudiés correspondent à des conditions expérimentales fréquemment rencontrées. Pour chacun d'eux, la méthode d'analyse statistique est développée en application des principes de la première partie.

1 - RAPPELS

1) L'aléatoire X a valeurs dans N a la loi Bi(n,θ) si :

$$\text{Prob}(X = x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

2)  $E(X) = n\theta$      $\text{Var}(X) = n\theta(1-\theta)$

3) Quand  $n \rightarrow \infty$

$$\frac{X - n\theta}{\sqrt{n\theta(1-\theta)}} \sim N(0,1)$$

4) Le nombre X de réalisations d'un événement A dans une série de n épreuves indépendantes et identiques a la loi Bi (n, Prob A)

2 - DISTRIBUTIONS A PRIORI ET A POSTERIORI

Modèle probabiliste

Le résultat expérimental est un entier positif x, valeur de l'aléatoire :

$$X \sim \text{Bi}(n, \theta)$$

$$p(x|\theta) \propto \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

Distribution a priori

Elle est choisie dans la famille des lois Beta (r,s) (Voir annexe A3) :

$$p_0(\theta) \propto \theta^{r-1} (1-\theta)^{s-1}$$

$$E(\theta) = \frac{r}{r+s} \quad \text{Var}(\theta) = \frac{r s}{(r+s)^2(r+s+1)}$$

Distribution a postérieure

$$p_1(\theta|x) \propto \theta^{r-1} (1-\theta)^{s-1} \binom{n}{x} \theta^x (1-\theta)^{n-x} \propto \theta^{r+x-1} (1-\theta)^{s+n-x-1}$$

$$\theta \sim \text{Beta}(r+x, s+n-x)$$

que les nombres  $x$  et  $n-x$ .

### Choix de la distribution a priori

Les paramètres  $r$  et  $s$  peuvent être fixés en remarquant que l'information a priori est équivalente à celle fournie par une série de  $r+s$  épreuves dont  $r$  ont donné un résultat positif. Le nombre  $r+s$  définit donc la valeur de l'information a priori.

Le quotient  $r/(r+s)$  est une valeur centrale de la distribution.

L'information est négligeable si  $r+s$  est petit ; nous prendrons donc la limite des résultats quand  $r+s$  tend vers 0. Dans ce cas, a postériori :

$$\underline{\theta} \sim \text{Beta}(x, n-x)$$

### 3 - RESULTATS ASYMPTOTIQUES

Pour obtenir une loi approchée si  $n$  grand, on fait tendre  $n$  vers l'infini ; l'information a priori est alors faible par rapport à l'information expérimentale ; le choix des paramètres  $r$  et  $s$  importe peu et, pour simplifier les formules, nous prendrons  $r = s = 1$ .

Le développement de Taylor à l'ordre 2 de  $\ln p(\theta|x)$  au voisinage du mode