Formation et Analyse d'Images

James L. Crowley

ENSIMAG 3 Premier Semestre 2009/2010

Lesson 9 14 décembre 2009

Analyse et Reconnaissance Statistique

Plan de la Séance :

Notations

$\vec{x}$	Un vecteur
$\vec{X}$	Un vecteur aléatoire (non prévisible).
D	Nombre de dimensions de $\vec{X}$
T_k	La classe k
k	Indice d'une classe
K	Nombre de classes
M_k	Nombre d'exemples de la classe k.
M	Nombre totale d'exemples de toutes les classes
ω_k	L'affirmation que l'Évènement $E \in T_k$
$p(\omega_k)=p(E \in T_k)$	Probabilité que Évènement $E \in T_k$
$\vec{X}$	Une observation (un vecteur aléatoire).
$P(\vec{X})$	Fonction de densité pour la Probabilité d'une observation $\vec{X}$

1. La Reconnaissance

La reconnaissance est une capacité fondamentale de l'intelligence et même de la vie. Pour la survie, il faut savoir reconnaître les amis, les ennemies et la nourriture.

Reconnaissance : Le fait de reconnaître, d'identifier un objet, un être comme tel.

Identifier : Reconnaître une entité comme un individu

Classer : Reconnaître une entité comme un membre d'une catégorie, ou d'une classe.

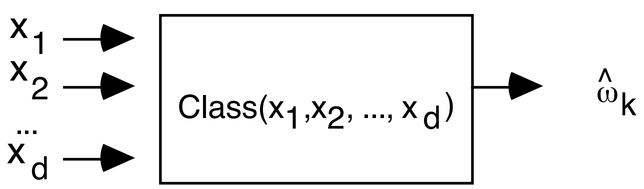
Un ensemble est défini par un test d'appartenance.

La classification est un processus d'association une entité (un événement) à une classe. L'entité est décrite par un vecteur des caractéristiques, produite par une observation. L'affectation de l'entité à une classe est faite par un test, calculer sur le vecteur de caractéristiques.

Caractéristiques : (En anglais : Features) Signes ou ensembles de signes distinctifs. Une ensemble de propriétés. $\{ x_1, x_2 \ldots x_D\}$.

En notation vectorielle : $\vec{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \ldots \\ x_D \end{pmatrix}$

Pour un vecteur de caractéristique, $\vec{X}$, un processus de classification propose une estimation de la classe, $\hat{\omega}_k$



Les techniques de reconnaissance de formes, statistiques fournissent une méthode pour induire des tests d'appartenance à partir d'un ensemble d'échantillons.

Les classes peuvent êtres définis par
extension : une liste complète des membres
intention : une conjonction de prédicats.

par extension : Une comparaison d'une observation avec des membres connus de la classe (des prototypes). Ceci correspond (grosso modo) à des méthode dite "générative" de reconnaissance.

Dans ce cas, la classification peut être fait par comparaison avec le membre de la classe.

$$\hat{\omega}_k = \forall_k \forall_m : \arg\text{-}\max_k \left\{ Sim(\vec{X}, \vec{X}_m^k) \right\}$$

et souvent le « similitude » est mesurée par la distance L2 entre caractéristiques

$$\hat{\omega}_k = \forall_k \forall_m : \arg\text{-}\max_k \left\{ \| \vec{X}, \vec{X}_m^k \| \right\}$$

encore mieux, est de définir une métrique de distance L pour chaque classe

$$\hat{\omega}_k = \forall_k \forall_m : \arg\text{-}\max_k \left\{ (\vec{X} - \vec{X}_m^k)^T \Lambda_k (\vec{X} - \vec{X}_m^k) \right\}$$

où le métrique L_k est l'inverse de la $L_k = C_k^{-1}$

$$C_k = E\{(\vec{X}_m^k - \vec{\mu}_k)^2\}$$

Note que, dans cas, nous pouvons remplacer la liste d'exemple pour la moyenne

$$\vec{\mu}_k = E\{\vec{X}_m^k\} = \frac{1}{M} \sum_{m=1}^{M} \vec{X}_m^k$$

ensuite:

$$\hat{\omega}_k = \forall_k : \arg\text{-}\max_k \left\{ (\vec{X} - \vec{\mu}_k)^T \Lambda_k (\vec{X} - \vec{\mu}_k) \right\}$$

par intention : Conjonction de prédicats définis sur les propriétés observées. Ceci correspond (grosso modo) à des méthode dite "discriminative" de reconnaissance.

Génératives : Les techniques fondaient sur un modèle.
Discriminatives : Les techniques fondaient sur des tests quelconques.

La teste d'appartenance est une forme de partition de l'espace de caractéristiques.

La classification se résume à une division de l'espace de caractéristiques en partition disjoint. Cette division peut-être fait par estimation de fonctions paramétrique ou par une liste exhaustives des frontières.

Le critère est la probabilité conditionnelle d'appartenance.

ω_k	Proposition que l'évènement E $\in$ la classe k
$p(\omega_k) = p(E \in T_k)$	Probabilité que E est un membre de la classe k.

Ayant une observation, $\vec{X}$, le critère de partition est la probabilité.

$$p(\omega_k \mid \vec{X}) = Pr(E \in T_k \text{ étant donnée } \vec{X})$$

$$\hat{\omega}_k = \arg\text{-}\max_{\omega_k}\left\{p(\omega_k \mid \vec{X})\right\}$$

Cette probabilité est fournie par la règle de Bayes.

$$p(\omega_k \mid \vec{X}) = \frac{p(\vec{X} \mid \omega_k)p(\omega_k)}{p(\vec{X})}$$

2. La Probabilité d'un Evénement.

La sémantique (ou "sens") de la probabilité d'un événement peut être fourni par sa fréquence d'occurrence ou par un système d'axiomes. L'approche fréquentielle a l'avantage d'être facile à comprendre. Par contre, elle peut entraîner les difficultés dans l'analyse. La définition axiomatique favorise les analyses mathématiques.

Dans le deux cas, la probabilité est une fonction numérique, $Pr() \in [0, 1]$.
Le domaine de la fonction Pr() est les évènements, E.

Définition Fréquentielle.

Une définition "Fréquentielle" de la probabilité sera suffisante pour la plupart des techniques vues dans ce cours.

Soit M observations de l'événement aléatoire dont M_k appartient à la classe A_k.
La Probabilité d'observer un événement E de la classe A_k est

$$p(E \in A_k) \equiv \lim_{M \to \infty} \left\{ \frac{M_k}{M} \right\}$$

Pour le cas pratique ou M est fini, $p(E \in A_k) \approx \frac{M_k}{M}$
La validité (ou précision) de l'approximation dépend du nombre d'échantillons M.

Définition Axiomatique.

Une définition axiomatique permet d'appliquer certaines techniques d'analyse de systèmes probabilistes. Trois postulats sont suffisants :

Postulat 1 : $\forall A_k \in S$: $p(E \in A_k) \geq 0$
Postulat 2 : $p(E \in S) = 1$
Postulat 3 :
$\forall A_i, A_j \in S$ tel que $A_i \cap A_j = \varnothing$: $p(E \in A_i \cup A_j) = p(E \in A_i) + p(E \in A_j)$

La probabilité de la valeur d'une variable aléatoire

Pour x entier, tel que $x \in [x_{min}, x_{max}]$, on peut traiter chacun des valeurs possibles comme une classe d'événement.

Si les valeurs de x sont entières, tel que $x \in [x_{min}, x_{max}]$
on peut estimer la probabilité à partir de M observations de la valeur, $\{X_m\}$.

Pour estimer la probabilité d'une valeur on peut compter le nombre d'observation de chaque valeur, x, dans une table, h(x).

L'existence des ordinateurs avec des centaines de mégaoctets rendre des tables de fréquence très pratique pour la mise en œuvre en temps réel des algorithmes de reconnaissance. Dans certains domaines, comme l'analyse d'images, par abus de langage, une telle table s'appelle un histogramme. Proprement dit, l'histogramme est une représentation graphique de h(x)

Ainsi la probabilité d'une valeur de $X \in [X_{min}, X_{max}]$ est la fréquence d'occurrence de la valeur. Avec M observations de la valeur, X, on peut faire une table, h(x), de fréquence pour chacun des valeurs possibles. On observe M exemples de X, $\{X_m\}$.

Pour chaque observation on ajoute "1" à son entré dans la table.

$$\forall m=1, M \ : \ h(X_m) := h(X_m) + 1; \ \ M := M+1;$$

h(x) est une table de fréquence pour chaque $x \in [x_{min}, x_{max}]$.

Ainsi, on peut définir la probabilité d'une valeur x par sa fréquence :

$$p(X_m=x) \equiv \lim_{M\to\infty} \{\frac{1}{M} \ h(x)\}$$

Quand M est fini, on peut faire appel à l'approximation.

$$P(X=x) \approx \frac{1}{M} h(x)$$

La validité de l'approximation dépend du nombre de valeurs possible et de M.
En règle générale, on dit qu'il faut 10 exemples par cellule de l'histogramme.

3. La Règle de Bayes

Soit un évènement "E". Soit deux tribus d'évènements T_1 et T_2 tel que certains évènements sont commun à T_1 et à T_2.
E peut appartenir à $T_1 \cap T_2$ ou a $\neg T_1 \cap T_2$ ou a $T_1 \cap \neg T_2$ ou a $\neg T_1 \cap \neg T_2$

Soit deux propositions p et q.

$$A \equiv E \in T_1 \text{ et } B \equiv E \in T_2$$

donc $P(A) \equiv Pr\{E \in T_1\}$ et $P(B) \equiv Pr\{E \in T_2\}$.

Par axiome 2 de la définition des systèmes de probabilités :

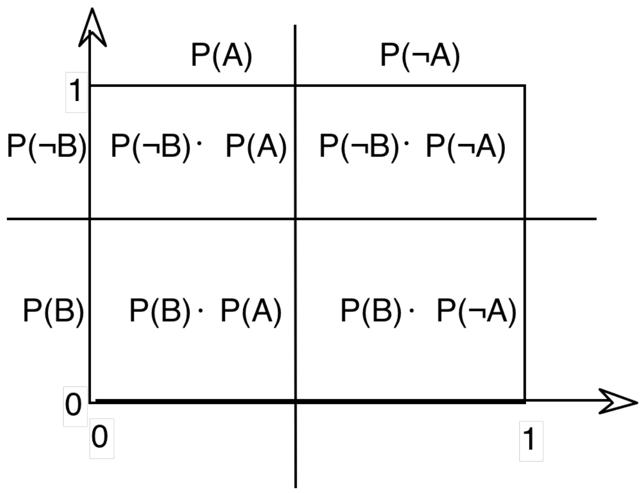
$$P(B) + P(\neg B) = 1.$$

$P(A \wedge B)$ est la probabilité "conjointe" de A et B.

Si A et B sont indépendantes

$$P(A \wedge B) = P(A) \cdot P(B),$$
$$P(A \vee B) = P(A) + P(B).$$

On peut voir ça d'une manière graphique :



$$P(A \wedge B) + P(A \wedge \neg B) + P(\neg A \wedge B) + P(\neg A \wedge \neg B) = 1$$

Dans ce cas, les probabilités marginales sont

$$P(A) = P(A \wedge B) + P(A \wedge \neg B)$$
$$P(B) = P(A \wedge B) + P(\neg A \wedge B)$$

La probabilité conditionnelle de B étant donnée A s'écrit $P(B \mid A)$

$$P(B \mid A) = \frac{P(A \wedge B)}{P(A)} = \frac{P(A \wedge B)}{P(A \wedge B) + P(A \wedge \neg B)}$$

de la même manière :

$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)} = \frac{P(A \wedge B)}{P(A \wedge B) + P(\neg A \wedge B)}$$

Par algèbre on déduire :

$$P(B \mid A)\, P(A) = P(A \wedge B) = P(A \mid B)\, P(B)$$

d'ou

$$P(B \mid A)\, P(A) = P(A \mid B)\, P(B)$$

Ceci est une forme de <u>règle de Bayes</u>. On peut écrire :

$$P(B \mid A)\, P(A) = P(A \mid B)\, P(B)$$

ou

$$P(B \mid A) = \frac{P(A \mid B) P(B)}{P(A)}$$

$P(B \mid A)$ est la probabilité "conditionnelle" ou "postérieur"

4. Classification des Pixels par Ratio d'Histogramme

Un histogramme est une table de fréquence. Il peut fournir une estimation d'une densité de probabilités. Pour x entier, tel que $x \in [x_{min}, x_{max}]$, on peut traiter chacun des valeurs possibles comme une classe d'événement.

Si les valeurs de x sont entières, tel que $x \in [x_{min}, x_{max}]$
on peut estimer la probabilité à partir de M observations de la valeur, $\{X_m\}$.

Pour estimer la probabilité d'une valeur on peut compter le nombre d'observation de chaque valeur, x, dans une table, h(x).

L'existence des ordinateurs avec des centaines de mégaoctets rendre des tables de fréquence très pratique pour la mise en œuvre en temps réel des algorithmes de reconnaissance. Dans certains domaines, comme l'analyse d'images, par abus de langage, une telle table s'appelle un histogramme. Proprement dit, l'histogramme est une représentation graphique de h(x)

Ainsi la probabilité d'une valeur de $X \in [X_{min}, X_{max}]$ est la fréquence d'occurrence de la valeur. Avec M observations de la valeur, X, on peut faire une table, h(x), de fréquence pour chacun des valeurs possibles. On observe M exemples de X, $\{X_m\}$.

Pour chaque observation on ajoute "1" à son entré dans la table.

$$\forall m=1, M \ : \ h(X_m) := h(X_m) + 1; \ \ M := M+1;$$

h(x) est une table de fréquence pour chaque $x \in [x_{min}, x_{max}]$.
Ainsi, on peut définir la probabilité d'une valeur x par sa fréquence :

$$p(X_m=x) \equiv \lim_{M \to \infty} \{\frac{1}{M} \ h(x)\}$$

Quand M est fini, on peut faire appel à l'approximation.

$$P(X=x) \approx \frac{1}{M} \quad h(x)$$

La validité de l'approximation dépend du ratio entre le nombre de Cellules, $Q = N^d$, de h(x) et le nombre d'échantillons, M.

L'erreur moyenne entre $\frac{1}{M}$ $h(\vec{C})$ et $P(\vec{C})$ est $E_{ms} \sim O(\frac{Q}{M})$

Pour que l'estimation soit "raisonnable", il faut assuré que $M >> Q = N^d$
En règle générale, on dit qu'il faut 10 exemples par cellule de l'histogramme.

Détection d'objet par ratio d'histogramme de couleur

On peut utiliser les histogrammes avec la règle de Bayes pour détecter les objets.

Par exemple, construisons un histogramme pour le vecteur de chrominance (r,v).

La chrominance $C = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ est une signature pour l'objet.

La chrominance peut être définit par plusieurs codages. Par exemple, pour la détection du peau, il est fréquent de voir

$$c_1 = r = \frac{R}{R+V+B} \qquad c_2 = v = \frac{V}{R+V+B}$$

Supposons qu'on code c_1 et c_2 avec les entiers entre 0 et $N - 1$

$$c_1 = \text{Round}\left((N-1) \cdot \frac{R}{R+G+B} \right) \quad c_2 = \text{Round}\left((N-1) \cdot \frac{G}{R+G+B} \right)$$

On alloue un tableau 2D, $h(c_1, c_2)$, de taille N x N cellules.
(exemple Q = 32 x 32 = 1024 cellules)

Pour chaque pixel $\vec{C} = C(i, j)$ dans l'image, on incrémente la cellule de l'histogramme qui correspond à $\vec{C}$: $h(\vec{C}) := h(\vec{C}) +1$

c'a-dire $h(c_1, c_2) := h(c_1, c_2) +1$

Soit M Pixels dans l'image. Un histogramme des chrominance, $h(\vec{C})$, des M pixels dans une l'image donne leurs fréquences d'occurrence.

$$P(\vec{C}) \approx \frac{1}{M} \ h(\vec{C})$$

Considère une région W de M_o pixels du même image correspondance à l'objet O.

$$\forall (i,j) \in W \ : \ h_o(\vec{C}(i,j)) := h_o(\vec{C}(i,j)) + 1$$

Ensuite: pour tout pixel $\vec{C}(i, j) = \binom{r}{v}(i, j)$: $p(\vec{C} \mid \text{objet}) \approx \frac{1}{M_o} \ h_o(\vec{C})$

Parce que W est dans l'image, la probabilité de rencontrer un pixel de W,

$$P(W) = \frac{M_o}{M}$$

L 'histogramme permet d'utiliser la règle de Bayes afin de calculer
la probabilité qu'un pixel corresponde à un objet.

Pour chaque pixel $\vec{C}(i, j)$ $p(\text{objet} \mid C) = p(C \mid \text{objet}) \frac{p(\text{objet})}{p(C)}$

Soit M images de I x J pixels. Ceci fait N = I x J x M Pixels.
Soit h(r, v), l'histogramme de tous les N pixels.
Soit $h_o(r, v)$, l'histogramme des N_o pixels de l'objet "o".

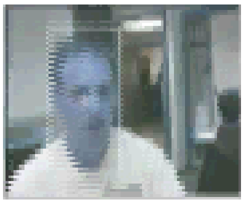
$$p(\text{objet}) = \frac{M_o}{M}$$

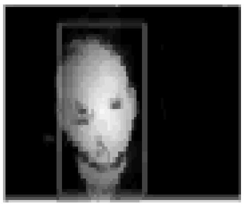
$$p(\vec{C}) = \frac{1}{M} \ h(\vec{C})$$

$$p(\vec{C} \mid \text{objet}) = \frac{1}{M_o} \ h_o(\vec{C})$$

$$p(\text{objet} \mid \vec{C}) = p(\vec{C} \mid \text{objet}) \frac{p(\text{objet})}{p((\vec{C}))} = \frac{1}{M_o} \ h_o(\vec{C}) \frac{\frac{M_o}{M}}{\frac{1}{M} h((\vec{C}))} = \frac{h_o(\vec{C})}{h(\vec{C})}$$

Par exemple, voici une image de la probabilité de peau fait par ratio d'histogramme de r, v





Caractérisation d'un région par moments

Les ensemble connexes de pixels s'appelles les "blobs".

On peut décrire une blob par une vecteur de caractéristiques "invariantes" à l'orientation grâce aux "moments”

Les moments sont invariants aux transformations affines.

Pour une fenêtre (imagette) w(i, j) de taille N x M

Somme des Pixels : $$S = \sum_{i=1}^{M} \sum_{j=1}^{N} w(i, j)$$

Premiers moments :

$$\mu_i = \frac{1}{S} \sum_{i=1}^{M} \sum_{j=1}^{N} w(i, j) \cdot i \qquad \mu_j = \frac{1}{S} \sum_{i=1}^{M} \sum_{j=1}^{N} w(i, j) \cdot j$$

Le premier moment est le centre de gravité de la forme :

Deuxième moment :

$$\sigma_i^2 = \frac{1}{S} \sum_{i=1}^{M} \sum_{j=1}^{N} (w(i, j)) \cdot (i-\mu_i)^2$$

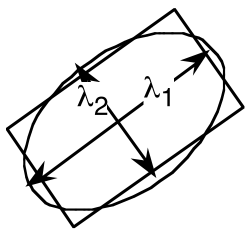
$$\sigma_j^2 = \frac{1}{S} \sum_{i=1}^{M} \sum_{j=1}^{N} w(i, j) \cdot (j-\mu_j)^2$$

$$\sigma_{ji}^2 = \frac{1}{S} \sum_{i=1}^{M} \sum_{j=1}^{N} w(i, j) \cdot (i-\mu_i)(j-\mu_j)$$

Ceci permet de définir les "axes", majeur, λ_1 et mineur, λ_2, de la forme par analyse des composantes principales de la deuxième moment

$$\mathbf{C}_o \triangleq \begin{bmatrix} s_i^2 & s_{ij}^2 \\ s_{ij}^2 & s_j^2 \end{bmatrix}$$

Composantes principales



Les deuxièmes moments sont "invariants" à l'orientation

Les axes sont calculés par une analyse en composantes principales de la matrice C. Il s'agit de trouver une rotation, **Φ,** dans l'espace de caractéristiques $\mathbf{\Phi}\ \mathbf{C}_P\ \mathbf{\Phi}^T = \mathbf{\Lambda}$ telles que $\mathbf{\Lambda}$ soit diagonale.

$$\mathbf{\Lambda} = \begin{pmatrix} l_1 & 0 \\ 0 & l_2 \end{pmatrix} \quad \text{tel que } \lambda_1 > \lambda_2 \qquad \mathbf{\Phi} = \begin{pmatrix} \mathrm{Cos}(q) & \mathrm{Sin}(q) \\ -\mathrm{Sin}(q) & \mathrm{Cos}(q) \end{pmatrix}$$

$$\mathbf{\Phi}\ \mathbf{C}_P\ \mathbf{\Phi}^T = \mathbf{\Lambda} = \begin{pmatrix} l_1 & 0 \\ 0 & l_2 \end{pmatrix} \qquad \mathbf{\Phi}^T\ \mathbf{\Phi} = \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{\Phi}\ \mathbf{C}_P\ \mathbf{\Phi}^T\ \mathbf{\Phi} = \mathbf{\Phi}\ \mathbf{C}_P = \mathbf{\Lambda}\ \mathbf{\Phi} = \begin{pmatrix} l_1 & 0 \\ 0 & l_2 \end{pmatrix} \begin{pmatrix} \mathrm{Cos}(q) & \mathrm{Sin}(q) \\ -\mathrm{Sin}(q) & \mathrm{Cos}(q) \end{pmatrix}$$

Les lignes du **Φ** sont des vecteurs propres du **C.**
La longueur des axes majeurs et mineur est les valeurs propres de la matrice **C**.

θ est l'orientation de l'axe "majeur" et λ_1 / λ_2 est le rapport entre la longueur et la largeur.

λ_1 / λ_2 est une caractéristique invariante de la taille et de l'orientation.

Par exemple, $\vec{X} = \begin{pmatrix} m_i \\ m_j \\ l_1 \\ l_2 \\ q \end{pmatrix}$ est une vecteur de caractéristique pour les "blobs".

Histogrammes de Champs Réceptifs.

Cette méthode peut être généralise en remplaçant la chrominance par une vecteur de champs réceptifs. Mais il faut bien gérer la ration Q/M !

Soit une image p(i,j) et une vecteur de "d" champs réceptifs $\vec{G}$

$\vec{V}(i,j) = < \vec{G}, p(i,j) >$ est un vecteur de caractéristiques de d dimensions.

$h(\vec{V})$ aura $Q = N^d$

N \ d	1	2	3	4	5	6
2	2^1	2^2	2^3	2^4	2^5	2^6
4	2^2	2^4	2^6	2^8	2^{10} =1 Kilo	2^{12} =
8	2^3	2^6	2^9	2^{12}	2^{15}	2^{18}
16	2^4	2^8	2^{12}	2^{16}	2^{20} = 1 Meg	2^{24} =
32	2^5	2^{10} =1 Kilo	2^{15}	2^{20} = 1 Meg	2^{25}	2^{30} =
64	2^6	2^{12}	2^{18}	2^{24}	2^{30} = 1 Gig	236
128	2^7	2^{14}	2^{21} = 2 Meg	2^{28}	2^{35}	2^{42} =
256	2^8	2^{16}	2^{24}	2^{32} = 2 Gig	2^{40} = 1 Tera	248

Soit les champs réceptifs achromatique

$$\vec{G}_{\sigma} = (G_x, G_y, G_{xx}, G_{xy}, G_{yy})$$

d = 5

ou chromatique avec normalisation de l'orientation et echelle :

$$\vec{G}_{\sigma,\theta} = (G_x^L, G_x^{C_1}, G_x^{C_2}, G_{xx}^L, G_{xy}^L, G_{xx}^{C_1}, G_{xx}^{C_2})$$

d = 7.

On peut faire

$$p(\text{objet}(i,j) \mid \vec{V}(i,j)) = \frac{p(\vec{V}(i,j) \mid objet(i,j))\,p(objet(i,j)}{p(\vec{V}(i,j)} = \frac{h_{objet}(\vec{V}(i,j))}{h_{tot}(\vec{V}(i,j))}$$

sur condition de gérer M et Q.

Rappel :

$$P(X=x) \approx \frac{1}{M} \quad h(x)$$

La validité de l'approximation dépend du nombre de valeurs possible et de M.
En règle générale, on dit qu'il faut 10 exemples par cellule de l'histogramme.

On peut démontrer que l'écart type moyenne de l'erreur est en proportion avec la ration du nombre de cellule de h(x), N, sur le nombre d'échantillons, M.

$$\sigma_{MSE} \approx \frac{N}{M}$$

Que faire si la masse d'exemple est insuffisante : M << N?
Que faire si x n'est pas entier ? Il faut une fonction <u>paramétrique</u> pour p(X).

La validité de l'approximation dépend de la nombre de valeurs possible et de M.
En règle générale, on dit qu'il faut 10 exemples par cellule de l'histogramme.

Que faire si la masse d'exemple est insuffisante : $M < 10 (X_{max} - X_{min})$?
Que faire si x n'est pas entier ?
Dans ces cas, on peut faire appel à une fonction <u>paramétrique</u> pour p(X).