画像処理のための統計学用語集

・平均値（期待値）

$$\bar{x}=\frac{1}{n}\sum_{k=1}^{n}x_k$$

・分散値

$$\sigma^2=\frac{1}{n}\sum_{k=1}^{n}(x_k-\bar{x})^2$$

・標準偏差値

$$\sigma=\pm\sqrt{\sigma^2}$$

・ベクトルの内積はスカラー

$$\boldsymbol{x}^T\boldsymbol{x}=(a_1,a_2,\cdots,a_k)\begin{pmatrix}b_1\\b_2\\\vdots\\b_k\end{pmatrix}$$

$\bullet^T$:転置を示す.

ex.

$$(a_1,a_2)\begin{pmatrix}b_1\\b_2\end{pmatrix}=a_1*b_1+a_2*b_2=s$$

・ベクトルの外積は分散共分散行列

$$\boldsymbol{x}\boldsymbol{x}^T=\begin{pmatrix}a_1\\a_2\\\vdots\\a_k\end{pmatrix}(b_1,b_2,\cdots,b_k)$$

$$=\begin{pmatrix}a_1b_1 & a_1b_2 & \cdots & a_1b_k\\ a_2b_1 & a_2b_2 & & \vdots\\ \vdots & & \ddots & \vdots\\ a_k & \cdots & \cdots & a_kb_k\end{pmatrix}$$

ex.

$$\begin{pmatrix}a_1\\a_2\end{pmatrix}(b_1,b_2)=\begin{pmatrix}a_1b_1 & a_1b_2\\ a_2b_1 & a_2b_2\end{pmatrix}$$

・相関があるのかないのか？

$E[\boldsymbol{x}^T\boldsymbol{x}]=0,$あるいは$E[\boldsymbol{x}^T\boldsymbol{y}]=0$ならば

ベクトル $\boldsymbol{x}$ の元（要素）は互いに無相関，
あるいはベクトル $\boldsymbol{x}$ とベクトル $\boldsymbol{y}$ は無相関

・相関関数

定常確率過程に従う 2 つの過程 $\boldsymbol{x}(t), \boldsymbol{y}(t)$ を考えた場合，いま，時刻 t におけるそれぞれの確率密度関数を $p(x,t), p(y,t)$ とすると $\boldsymbol{x}(t), \boldsymbol{y}(t)$ の平均値は時刻 t に無関係な定数となる．

$$E[\boldsymbol{x}(t)] = \int_{-\infty}^{\infty} xp(x,t)dx = m_x$$

$$E[\boldsymbol{y}(t)] = \int_{-\infty}^{\infty} yp(y,t)dy = m_y$$

・自己相関関数（auto-correlation function）

自分自身に自分自身を畳み込むという異色な関数

$$E[\boldsymbol{x}(t)\boldsymbol{x}(t+\tau)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 p(x_1, t : x_2, t+\tau) dx_1 dx_2 = R_{xx}(\tau)$$

・相互相関関数（cross-correlation function）

$$E[\boldsymbol{x}(t)\boldsymbol{y}(t+\tau)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 y_2 p(x_1, t : y_2, t+\tau) dx_1 dy_2 = R_{xy}(\tau)$$

・パワースペクトル

自己相関関数 $R_{xx}(\tau)$ のフーリエ変換 $S_{xx}(\omega)$ はパワースペクトル
確率過程 $\boldsymbol{x}(t)$ のスペクトル密度（spectral density）あるいはパワースペクトル密度(power spectral density)と呼ばれる．

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-i\omega\tau} d\tau$$

$S_{xx}(\omega)$ は，$\int_{-\infty}^{\infty} \left| R_{xx}(\tau) \right| d\tau < \infty$ の条件下で存在する．

また，パワースペクトル $S_{xx}(\omega)$ のフーリエ逆変換 $R_{xx}(\tau) = \dfrac{1}{2\pi}\int_{-\infty}^{\infty} S_{xx}(\omega) e^{i\omega\tau} d\omega$ は自己相関関数 $R_{xx}(\tau)$ のスペクトル表示という．

・ベクトルを用いた $R_{xx}(\tau)$ と $S_{xx}(\omega)$ の関係

ある関数同士の畳み込みは，フーリエ変換後の領域において互いの関数のフーリエ変換の積で表現できる．

$$F[\int_{-\infty}^{\infty} g(x) h(x-\tau) d\tau] = F[g(x)]F[h(x)] = G(\omega)H(\omega)$$

F :フーリエ変換を示す線形作用素

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-i\omega\tau} d\tau$$

$$= \int_{-\infty}^{\infty} E[\boldsymbol{x}^T \boldsymbol{x}] e^{-i\omega\tau} d\tau$$

$$= \int_{-\infty}^{\infty} E[\boldsymbol{x}^T \boldsymbol{x}] e^{-i\omega\tau} d\tau$$

・直交変換

ある変換

$$\boldsymbol{y} = \boldsymbol{A}^T \boldsymbol{x}$$

$\boldsymbol{A}$:変換行列（例えばフーリエ変換基底）

を考えた場合，

$$E[\boldsymbol{y}\,\boldsymbol{y}^T] = \boldsymbol{A}^T E[\boldsymbol{x}\,\boldsymbol{x}^T]\boldsymbol{A} = \boldsymbol{\Lambda}$$

$\boldsymbol{\Lambda}$:固有値λ_kを成分に持つ対角行列

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_k \end{pmatrix}$$

上記のような変換を直交変換という

$\boldsymbol{A}^T \boldsymbol{A} = \boldsymbol{I}, \boldsymbol{I}$:単位行列

$\boldsymbol{A}^T = \boldsymbol{A}^{-1}$:従って直行行列の逆行列は転置となる．

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

直交変換は種々知られており Karhunen-Loeve 変換，Fourier 変換，JPEG 圧縮等で利用される離散コサイン変換，ウェーブレット変換なども直交変換の一種である．

確率論および確率過程を理解するための覚書

試行（trial）：事象（event）を生起させるための入力あるいは行動など etc.サイコロを振ること

事象（event）：試行（trial）の結果得られる事柄

根元事象（elementary event）：これ以上分解できない事象 etc.サイコロを振る試行の結果“1 の目が出る”，“6 の目が出る”こと

複合事象（compound event）：根元事象を組み合わせた事象 etc.“奇数の目が出る”，“1 と 2 の目のいずれかが出る”こと

標本空間（sample space）：全事象を全ての標本点（sample points）で示したもの

標本点（sample points）：試行（trial）の結果として与えられるもの

全事象（whole event）：全ての標本を示す

空事象（null event）ϕ：生起しえない事象 etc.1 回のサイコロの試行において 2 つの目が同時に出る

互いに排反（exclusive）な事象：$A\cap B=\phi$ を満たす事象 A および事象 B

和事象（sum event）；$A\cup B$

結合事象あるいは積事象（product event）：$A\cap B$

結合確率密度関数の性質においてベクトル x およびベクトル y が互いに独立であれば

$F(x,y)=F(x)F(y)$

$p(x,y)=p(x)p(y)$

である

確率変数 x と y が独立ということは，2 つの事象$\{x<=x1\}$と$\{y<=y1\}$が独立であることをも意味する．よって

$P(x<=x1\wedge y<=y1)=P\{x<=x1\}P\{y<=y1\}$

であり，これは結合確率分布関数 $F(x,y)=F(x)F(y)$と同義である．さらにこの式の両辺を x と y で偏微分すれば結合確率密度関数 $p(x,y)=p(x)p(y)$となる．

確率の公理

1.ある試行の結果として起こりうるすべての事象を考えたとき，いかなる事象 A に対しても，確率（probability）と呼ぶ負でない実数値 P(A)が対応している．

$0\leqq P(A)\leqq 1$

2.必ず起こる事象 S（標本空間）の確率を 1 とする．

$P(S)=1$

3.事象 A,B,C が互いに排反な事象であれば $P(A\cup B\cup C)=P(A)+P(B)+P(C)$ である．この公理を有限加法性の公理と呼ぶ.

結合事象
積事象の確率 P（A∩B）を結合確率（joint probability）とよぶ

独立性
事象 A と B の結合確率が P（A∩B）=P（A）P（B）であるとき，事象 A と B は統計的に独立（statistically independent）であるという.

条件付確率
零でない，ある正の確率をもつ事象 A が起こったという付帯条件のもとでの，ある事象 B の確率を条件付確率（conditional probability）といい，P（B|A）で表す．ここで注意しなければならないのはこの確率は B に関する確率であって，A はあくまでも，条件としての事象であることである.

乗法定理
P（A∩B）=P（A）P（B|A）=P（B）P（A|B）

条件付確率の性質
a.0≦P（B|A）≦1
b.P（B|A）≧P（A∩B）
c.事象 A と B が独立であれば P（B|A）=P（B）
A と B が独立ならば，A の生起を条件にしようがしまいが B の確率に変化がない.

ここで疑問
P（A∩B）=P（B|A）P（A）の場合 A と B は独立でない（従属である）が B|A と A は独立であるといえるか？

またここで疑問
独立の定義から
『A が B に対して統計的に独立であるということは，A の分布がどうであれ P（B）が一定であるということ.』
『逆に B が A に対して統計的に独立であるということは，B の分布に対して P（A）が一定であるということ.』
Q.「B が A に対して独立ならば、A が B に対して独立である」ことが言えるであろうか????????? → 独立である!!!!!

自然共役事前分布

共役とは、共役複素数という言葉からも分かるように、基本的に同じ構造を持ち合わせていることを意味する。ベイズの定理における共役とは、事前確率と事後確率とが同じような分布にしたがうことをいう。

因数（いんすう）

因数とは，掛け算（積）の結果の原因となる数のことである。整数は，いくつかの整数の掛け算（積）で表すことができるが，掛け合わせているそれぞれの整数が因数となる。
ある整数の因数とある整数の約数は同じことを意味する。
ある整式は，いくつかの整式の掛け算（積）で表すことができるが，掛け合わせている整式も因数という。

ベイズの定理

事象 A の起こったことが既知であるときの事象 Bi の確率 P(Bi|A)
乗法定理より P(B|A)P(A)=P(A|B)P(B)

$$P(Bi \mid A) = \frac{P(Bi)P(A \mid Bi)}{\sum_{j=1}^{n} P(Bj)P(A \mid Bj)}$$

$$P(A) = \sum_{j=1}^{n} P(Bj)P(A \mid Bj)$$

ベイズの定理は P(Bi|A)が直接求まらないときすでに求まっている P(A|Bi)と P(Bi)を用いてこれを知ることができるので，特に，与えられたデータや結果から原因を知ろうとする推定法に用いられることが多い.

確率変数

標本点 s に対する $\boldsymbol{x}$(s)という実関数を仮定した場合，$\boldsymbol{x}$(s)が取りうる値を確率変数と呼ぶ. 単に $\boldsymbol{x}$ で表すことも多い. {$\boldsymbol{x}$(s)= $\boldsymbol{x}_1$}は $\boldsymbol{x}$(s)が $\boldsymbol{x}_1$ という標本値をとる標本点による事象である. 標本点 s とは試行の結果与えられる点（数値等）である. これに対して事象とは，試行の結果得られる事柄である.
ex.)サイコロの目の確率変数を $\boldsymbol{x}$ とすると
事象{ $\boldsymbol{x}$ =1}の確率 P($\boldsymbol{x}$ =1)=1/6
事象{ $\boldsymbol{x}$ =1 ∪ $\boldsymbol{x}$ =2}=の確率 P($\boldsymbol{x}$ =1 ∪ $\boldsymbol{x}$ =2)=P($\boldsymbol{x}$ =1) + P($\boldsymbol{x}$ =2)=1/6 + 1/6 = 1/3

確率分布関数

$\mathrm{F}(x) = \mathrm{P}(\boldsymbol{x} \leq x)$

モーメントとは
平均値?　それとも‥‥
モーメントは積率ともいわれ、分布の平均値やバラツキ、ひずみや尖り度を数値化するものである。分布のモーメントは、密度関数が不明であるときに分布の特徴を説明するのに有用である。
確率論では k 次モーメントあるいは結合中心モーメントとして知られている.
結合モーメントにおいて次式$\mathrm{k}_1 = \mathrm{k}_2 = 1$の場合を相関とよぶ.
$m_{1,1} = \boldsymbol{E}[\boldsymbol{x_1 x_1}] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2) dx_1 dx_2$

確率変数の関数
1 変数関数の確率密度関数
確率変数$\boldsymbol{x}$と$\boldsymbol{y}$に関して 1 対 1 の関係があるとき，$\boldsymbol{x}$の確率密度関数$P_x(x)$と$\boldsymbol{y}$の$P_y(y)$の間には，対応する微小区間の確率が等しいことから$P_x(x)|\mathrm{d}x| = P_y(y)|\mathrm{d}y|$の関係が成り立つ.
$\boldsymbol{y} = f(\boldsymbol{x})$ならば$P_y(y) = P_x(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| = P_x(f^{-1}(y))\left|\frac{\mathrm{d}f^{-1}(y)}{\mathrm{d}y}\right|$の関係より $P_y(y)$を求めることができる.

EM アルゴリズムの理論

・収束性（Convergence property）

完全データ Y の分布は次式のように因数分解される．

$f(Y|\theta) = f(Y_{\mathrm{obs}}, Y_{\mathrm{mis}} \mid \theta) = f(Y_{\mathrm{obs}}|\theta) f(Y_{\mathrm{mis}} \mid Y_{\mathrm{obs}}, \theta)$

但し，

$f(Y_{\mathrm{obs}}|\theta)$は観測データ$Y_{\mathrm{obs}}$の密度

$f(Y_{\mathrm{mis}} \mid Y_{\mathrm{obs}}, \theta)$は観測データが与えられた下での欠損データの密度

である．

対数尤度の分解を用いれば次式で表される．

$l(\theta|Y) = l(\theta \mid Y_{\mathrm{obs}}, Y_{\mathrm{mis}}) = l(\theta \mid Y_{\mathrm{obs}}) + \ln f(Y_{\mathrm{mis}} \mid Y_{\mathrm{obs}}, \theta)$

EM アルゴリズムは固定されたY_{obs}のためのパラメータθに関して不完全データの対数尤度$l(\theta \mid Y_{\mathrm{obs}})$を最大化させることにより$\theta$を推定することを目的としている．

手順として

先ず

$l(\theta \mid Y_{\mathrm{obs}}) = l(\theta|Y) - \ln f(Y_{\mathrm{mis}} \mid Y_{\mathrm{obs}}, \theta)$

$l(\theta \mid Y_{\mathrm{obs}})$が最大化されるべき（観測される）対数尤度である．$l(\theta|Y)$は完全データの対数尤度であり容易に最大化できる．$\ln f(Y_{\mathrm{mis}} \mid Y_{\mathrm{obs}}, \theta)$は完全データの対数尤度における欠損部分である．

また，上式における両辺の期待値は次式で表される．

$l(\theta \mid Y_{\mathrm{obs}}) = Q(\theta \mid \theta^{(t)}) - H(\theta \mid \theta^{(t)})$

但し，

$$Q(\theta|\theta^{(t)}) = \int [l(\theta \mid Y_{\mathrm{obs}}, Y_{\mathrm{mis}})]\, f(Y_{\mathrm{mis}} \mid Y_{\mathrm{obs}}, \theta^{(t)}) dY_{\mathrm{mis}}$$

$$H(\theta|\theta^{(t)}) = \int [\ln f(Y_{\mathrm{mis}} \mid Y_{\mathrm{obs}}, \theta)]\, f(Y_{\mathrm{mis}} \mid Y_{\mathrm{obs}}, \theta^{(t)}) dY_{\mathrm{mis}}$$

$H(\theta|\theta^{(t)}) \le H(\theta^{(t)}|\theta^{(t)})$

Jensen’s inequality による．