2.3 多次元確率変数と　　　n変数への拡張（ベクトル）
r.v.
同時確率分布と
周辺 〃

行列の行と列 転置

多次元確率変数

n次元確率変数 $\vec{X} = (X_1, X_2, \cdots, X_n)^t$

一次元 r.v. の組

$$= \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

同時 p.d.

幾つかの r.v. を同時に扱うということ。

この多次元確率変数が従う確率分布を

同時確率分布

離散型と連続型

p.d.f $f(\vec{x}) = P(\vec{X} = \vec{x})$ $(\vec{x} = \vec{x}_1, \vec{x}_2, \cdots)$

$= 0$ $(\vec{x} \neq$ 〃 $)$

$$\begin{cases} f(\vec{x}_i) > 0 \\ \sum f(\vec{x}_i) = 1 \end{cases}$$

連続型分布

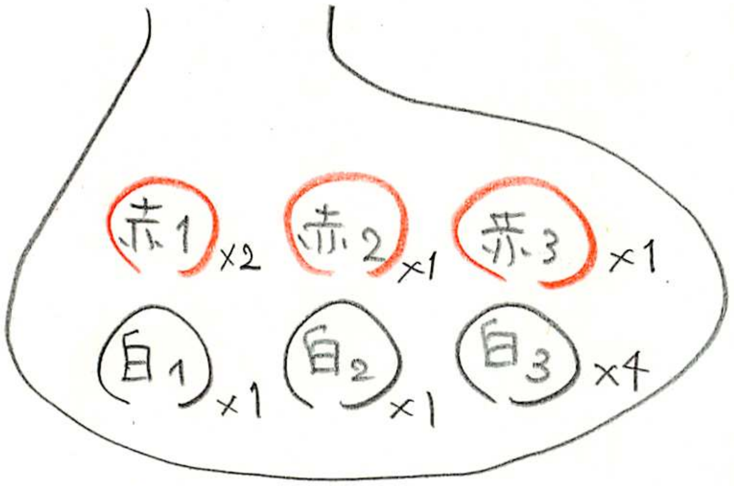
p.d.f の定義) $P(\vec{X} \leq \vec{x}) = P(X_1 \leq x_1, \cdots, X_n < x_n)$

$$= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t_1, \cdots, t_n)\, dt_1 \cdots dt_n$$

$$= \int_{-\vec{\infty}}^{\vec{\infty}} f(\vec{t})\, d\vec{t}$$

$$\begin{cases} f(\vec{x}) = f(x_1, \cdots, x_n) \geq 0 \\ \int_{-\vec{\infty}}^{\vec{\infty}} f(\vec{x})\, d\vec{x} = 1 \end{cases}$$

二次元離散型の例 2.3.1



3種の赤玉 と 3種の白球

赤(1, 2, 3)　　(白1, 2, 3)

一次元確率変数
$$\begin{cases} X_1 = 1, 2 & (\text{赤}=1,\ \text{白}=2) \\ X_2 = 1, 2, 3 & (\text{数字}=1, 2, 3) \end{cases}$$

二次元 r.v. $\vec{X} = (X_1, X_2)^t$

二次元離散型確率変数に対する確率分布の例

	$X_2=1$	$X_2=2$	$X_2=3$	$P(X_1=x_i)$
$x_1 = 1$ (赤)	2/10	1/10	1/10	4/10
$x_1 = 2$ (白)	1/10	1/10	4/10	6/10
$P(X_2=x_2)$	3/10	2/10	5/10	1

周辺確率分布

$\vec{X}$ の同時確率分布が決まっている時,

X_i だけの確率分布は?

例 2.3.1 では,

- 取出した玉が赤である確率
- 〃 1である確率

周辺確率分布と呼ぶ

多変数, 離散型の場合,

$\vec{X} = (X_1, \cdots, X_m)$ の場合で,

$f(\vec{x}) = P(\vec{X} = \vec{x})$ が known

・ $f_{\underset{i}{1}}(x_{\underset{i}{1}}) = P(X_{\underset{i}{1}} = x_{\underset{i}{1}})$

$$= P\left(\bigcup_{x_2, \cdots, x_k} \{(X_1, X_2, \cdots, X_m) = (x_1, x_2, \cdots, x_k)\}\right)$$

$$= \sum_{x_2, \cdots, x_k} P((X_1, X_2, \cdots, X_m) = (x_1, x_2, \cdots, x_m))$$

$$= \sum_{x_2, \cdots, x_k} f(\underbrace{x_1}_{\text{これだけ固定}}, x_2, \cdots, x_k)$$

多変数,連続型の場合

$$\int_{-\infty}^{x} f_{\underset{i}{1}}(x_{\underset{i}{1}})\,dx_{\underset{i}{1}} = P(X_{\underset{i}{1}} \leq x)$$

$$= P(X_{\underset{i}{1}} \leq x,\ -\infty < X_2 < \infty, \cdots)$$

$$= \int_{-\infty}^{x} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_n)\, d\vec{x}$$

周辺密度関数

$$\therefore\ f_1(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_n)\, d\vec{x}^{(1)\,(i)}$$

2.4
Stat

2.4. 多次元確率変数の特性値

$\vec{X} = (X_1, \cdots, X_n)$ の

期待値, 平均, 分散

期待値 (の定義) 一次元からの(段階的)拡張

① $g(\vec{X}) = g_i(X_i)$ の時 (一変数のみ)

$$E[g(\vec{X})] = E[g_i(X_i)]$$

$$\equiv \int^{\rightarrow} g(x_i) \cdot f(\vec{x}) \, d\vec{x}$$

$$= \int g(x_i) f_i(x_i) \, dx_i \cdot \left[\underbrace{\int^{\overrightarrow{(i)}} f(\vec{x}^{(i)}) \, d\vec{x}^{(i)}}_{x_i \text{以外}} \right]$$

② $\vec{g}(\vec{X}) = (g_1(\vec{X}), \cdots, g_n(\vec{X}))$ の時

$$\vec{E}[g(\vec{X})] \equiv (E[g_1(\vec{X})], \cdots, E[g_n(\vec{X})])$$

$\vec{X}$ の平均(ベクトル) $\vec{\mu}$

$$\vec{\mu} \equiv (\mu_1, \cdots, \mu_n)$$

$$\equiv \vec{E}[\vec{X}]$$ E: ベクトルを返す関数

$$= (E[X_1], \cdots, E[X_n])$$ ← 一変数のE

一次元から n次元への拡張の確認.

$$\int x_1 f(\vec{x}) d\vec{x} \overset{?}{=} \int x_1 f_1(x_1) dx_1$$

↑ 一次元

多次元での $E_{\vec{x}}[X_1] \overset{?}{=} E_{X_1}[X_1]$

$$E[X_1] = \vec{\int} x_1 f(\vec{x}) d\vec{x}$$

$g(\vec{X}) = X_1$

$$= \int x_1 \left(\vec{\int}^{(1)} f(\vec{x}) d\vec{x}^{(1)} \right) \cdot dx_1$$

↓

$f_1(x_1)$

$$= \int x_1 \cdot f_1(x_1) dx_1$$

$\vec{X}$の分散は?

「X_iとX_jの間の関係」が生まれる

X_iの分散は(一次元と同様)

$$\sigma_i^2 \equiv V[X_i] = E[(X_i - \mu_i)^2]$$ i方向の

X_iとX_jの共分散(covariance)

i方向とj方向のかさなり

$$\sigma_{ij} \equiv Cov[X_i, X_j]$$
$$= E[(X_i - \mu_i)(X_j - \mu_j)]$$ 的内積

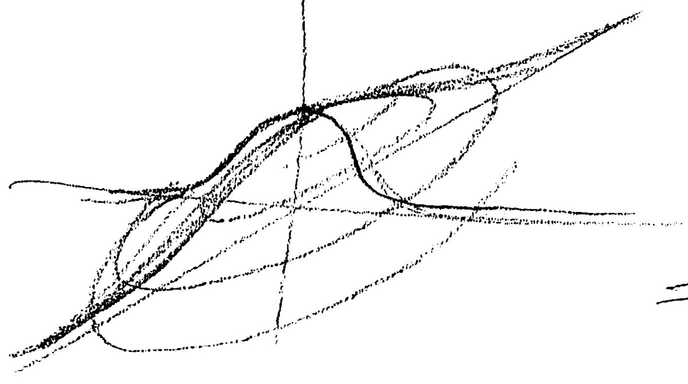
$\vec{X}$の共分散行列(covariance matrix)

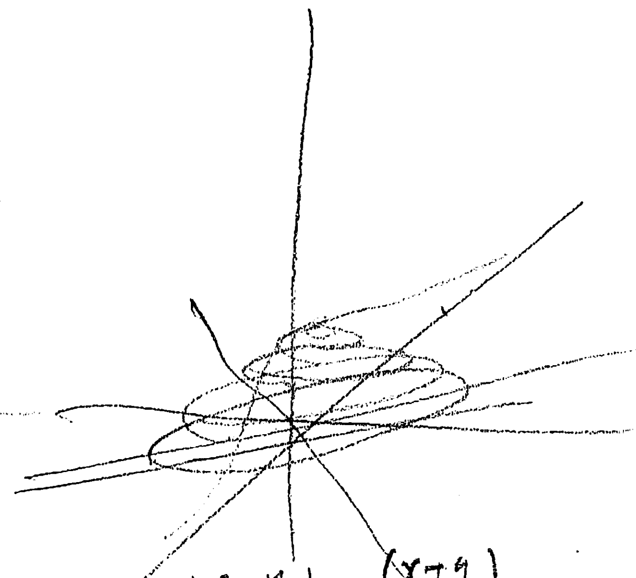
$$\Sigma = V[\vec{X}] = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \sigma_{ii} & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix}$$

$$\sigma_{ii} = \sigma_i^2$$

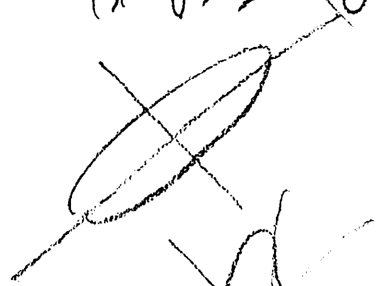
分散

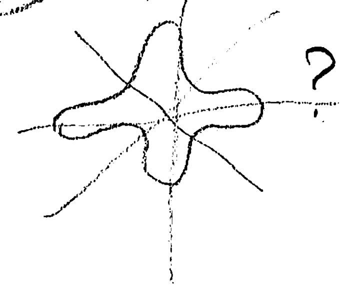
共分散の意味





$(x-y)$, $(x+y)$





隋円

$x^2+y^2=a^2$

$a^2x^2+b^2y^2=c^2$

$a^2(x-y)^2+b^2(x^2+y)^2=c^2$

任意ベクトル

行列　ベクトル

$$\vec{a}'\Sigma\vec{a} = (a_1 \cdots a_n)\begin{pmatrix} \text{---} \\ i \quad \Sigma \\ \quad \end{pmatrix}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

二次形式

$$= (a_1 \cdots a_n)\begin{pmatrix} \sum_{k=1}^{n} a_k \sigma_{1k} \\ \vdots \\ \sum_{k=1}^{n} a_k \sigma_{ik} \\ \vdots \\ \sum_{k=1}^{n} a_k \sigma_{nk} \end{pmatrix}$$

j

$$= \sum_{j=1}^{n} a_j \left(\sum_{k=1}^{n} a_k \sigma_{jk}\right)$$

$$= \sum_{j=1}^{n}\left(\sum_{k=1}^{n} a_j \cdot a_k \cdot \sigma_{jk}\right)$$

$$= \sum_{j=1}^{n}\sum_{k=1}^{n} a_j a_k E\left[(X_j - \mu_j)(X_k - \mu_k)\right]$$

$$= E\left[\sum\sum a_j a_k (X_j - \mu_j)(X_k - \mu_k)\right]$$

$$= E\left[\sum\sum \{a_j (X_j - \mu_j)\}\{a_k (X_k - \mu_k)\}\right]$$

$$= E\left[\left\{\sum a_i (X_i - \mu_i)\right\}^2\right] \geq 0$$

∴ 行列 Σ は半正定値

二次形式 ≥ 0

$\overset{=0}{(半)}$正定値行列　M　　$\left\{ \begin{array}{l} \text{実対称行列} \\ \text{エルミート 〃 (複素)} \end{array} \right.$

$$\vec{z}^t M \vec{z} \underset{(=)}{\geq} 0$$

$$\vec{z}^* M z \geq 0$$

全ての固有値 $\lambda \underset{(=)}{\geq} 0$

正定値なら
Mは正則で，M^{-1}と正定値

正定値なら $|M| > 0$

Mが線形独立なベクトルに対するグラム行列

$M = P^{-1} D P$　　Pはユニタリ行列
　　　　　　　　$D = (d_{ij})$　$d_{ij} = 0$　$i \neq j$

相関係数 (correlation coefficient)

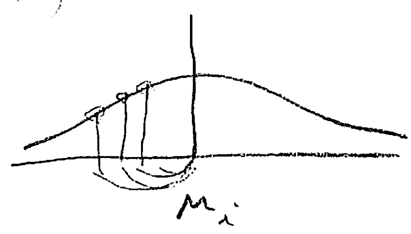
$$\rho_{ij} = \mathrm{Corr}[X_i, X_j] = \frac{\mathrm{Cov}[X_i, X_j]}{\sqrt{V[X_i]\ V[X_j]}}$$

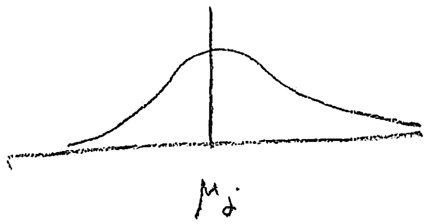
標準化変数 z_i と z_j の共分散と等しい。

$$z_i = \frac{X_i - \mu_i}{\sigma_i}$$

$X_i \in \{x_{i1}, \cdots, x_{in}\}$

$X_j \in \{x_{j1}, \cdots, x_{jn}\}$





$$\mu_i = E[X_i] = \frac{1}{n}\sum_{k=1}^{n} x_{ik}$$

標本相関係数

データ列 $\{(x_i, y_i)\}$ $(i=1 \sim n)$ が与えられた時,

標本共分散 $S_{xy} = \sum_{1}^{n} (x_i-\bar{x})(y_i-\bar{y})$

〃 標準偏差 S_x, S_y

2.5 確率変数の独立性

例　コイン投げを k 回行なう。

すべてが表である確率は $\left(\frac{1}{2}\right)\times\left(\frac{1}{2}\right)\times\cdots\left(\frac{1}{2}\right)=\left(\frac{1}{2}\right)^k$

一回毎に表である確率を k 回かけ合せることで、確率が求められる。

一回の（ベルヌーイ試行）　X,　$X=0$, 表
$X=1$ 裏

k 回の　〃　$X_1, X_2, \cdots, X_k$ で表わす。

この時 $(X_1,\cdots,X_k)=(x_1,\cdots,x_k)$ となる確率は、

$$
\begin{aligned}
&P((X_1,\cdots,X_k)=(x_1,\cdots,x_k))\\
&\quad = P(X_1=x_1)\cdot P(X_2=x_2)\cdots P(X_k=x_k)
\end{aligned}
$$

連続型、密度関数の表現では

$$f(x_1,\cdots,x_k)=f_1(x_1)\cdots f_k(x_k)$$

$$
\begin{aligned}
F(x_1,\dots,x_k) &= P(X_1\le x_1,\cdots,X_n\le x_n)\\
&= F_1(x_1)\cdots F_n(x_n)
\end{aligned}
$$

独立

$$\begin{cases} f(\vec{x}) = \prod_i f_i(x_i) \\ F(\vec{x}) = \prod_i F_i(x_i) \end{cases}$$

$X_1, \cdots, X_n$が独立で、同一分布に従う　← それぞれ独立な確率的試行

$$X_1, \cdots, X_n \underset{i.i.d}{\sim} F(x)$$

変数変換と独立性

XとYが独立
⇓
$g(X)$と$h(Y)$も独立。

独立性と期待値

$$E[X_1 \cdots X_n] = E[X_1] \cdots E[X_n]$$
$$\left(E\left[\prod_i X_i\right] = \prod_i E[X_i] \right)$$

二変数の場合の確認

$$\begin{aligned} E[X_1 X_2] &= \iint x_1 x_2 f_1(x_1) f_2(x_2)\, dx_1 dx_2 \\ &= \int x_1 f_1(x_1)\, dx_1 \cdot \int x_2 f_2(x_2)\, dx_2 \\ &= E[X_1] \cdot E[X_2] \end{aligned}$$

独立性と無相関性

XとYが独立,

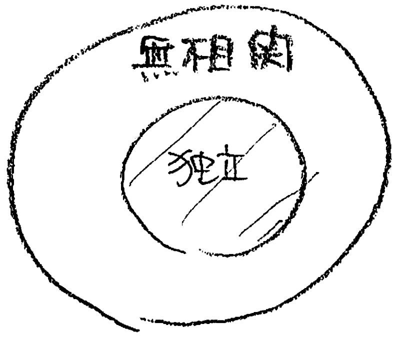
⇓

XとYは無相関

$$\because \quad Cov[X,Y] = E[(X-\mu_x)(Y-\mu_y)]$$

$$= E[X-\mu_x]\cdot E[Y-\mu_y]$$

$$= 0$$

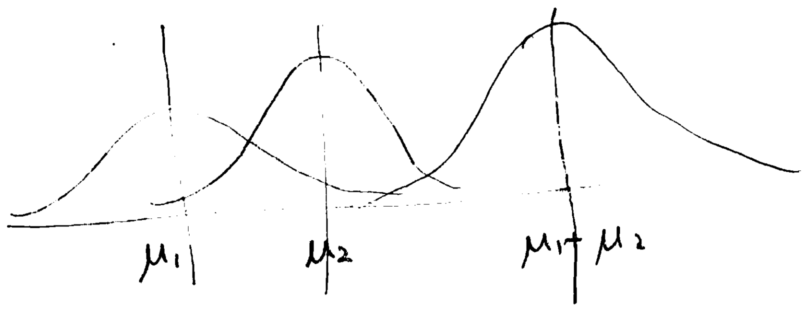


2.6 確率変数の和の平均と分散

$\left(\sum_{i=1}^{n} X_i\right)$ の平均と分散について考える

和の平均

$$E[X_1+X_2] = \int (x_1+x_2) f(\vec{x})\, d\vec{x}$$
$$= \int x_1 f(\vec{x})\, d\vec{x} + \int x_2 f(\vec{x})\, d\vec{x}$$
$$= E[X_1] + E[X_2]$$



和の分散

$$V[X_1+X_2] = E[\{(X_1+X_2)-(\mu_1+\mu_2)\}^2]$$
$$= E[\{(X_1-\mu_1)+(X_2-\mu_2)\}^2]$$
$$= E[(X_1-\mu_1)^2] + E[(X_2-\mu_2)^2] + 2E[(X_1-\mu_1)(X_2-\mu_2)]$$
$$= V[X_1] + V[X_2] + 2\,Cov[X_1, X_2]$$

(n)
一般化

$$E[\sum X_i] = \sum E[X_i]$$
$$V[\sum X_i] = \sum V[X_i] + \underbrace{2\sum_{i<j} cov[X_i, X_j]}_{\text{無相関の時. } 0}$$

算術平均

$$\bar{X} = \frac{1}{n}\sum_{1}^{n} X_i$$

〃の
平均

$$E[\bar{X}] = E\left[\frac{1}{n}\sum X_i\right] = \frac{1}{n}\sum E[X_i] = \mu$$

〃の
分散

$$V[X] = V\left[\frac{1}{n}\sum X_i\right]$$

$$= \frac{1}{n^2} V\left[\sum X_i\right]$$

$$= \frac{1}{n^2}(n\sigma^2)$$

$$= \frac{\sigma^2}{n}$$

[A.2.2] 演習問題

$(X_1, \cdots, X_n)$ は連続型として、X_i の

独立性を表わす二つの表現の同値性を示せ。

(1) $f(x_1, \cdots, x_n) = f_1(x_1) \cdots f_n(x_n)$

(2) $F(x_1, \cdots x_n) = F_1(x_1) \cdots F_n(x_n)$

一変数の時 $F(x) = \int_{-\infty}^{x} f(t)\,dt$

$F'(x) = f(x)$

$$\left(\qquad F(\vec{x}) = \prod \int_{-\infty}^{x_i} f_i(t_i)\,dt_i \qquad \right)$$

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(x_1, \ldots, x_n)\, d\vec{x}$$

$$dF(\vec{x}) = \sum \frac{\partial}{\partial x_i} F(\vec{x}) \cdot dx_i$$

$$\frac{dF_i(x_i)}{dx_i} \cdot dx_i$$

(1) → (2)

$$\int f(\vec{x})\,d\vec{x} = \underbrace{\int f_1(x_1)dx_1}\cdot\underbrace{\cdots}\underbrace{\int f_n(x_n)dx_n}$$

$$\int_{-\infty}^{x_i} f_i(t_i)dt_i = F_i(x_i) \quad \therefore f_i(x_i) = \underset{\frac{\partial}{\partial x_i}}{F_i'(x_i)}$$

$$\therefore \int^{(i)} f(\vec{t})\,d\vec{t}^{(i)} = f_i(t_i)\underbrace{\int^{(i)} \prod_j f(t_j)\,dt_j}_{1}$$

$$F(x_1,\cdots,x_n) = \int_{-\infty}^{x_1} f_1(t_1)dt_1\cdot\cdots\cdots = \prod_j \int_{-\infty}^{x_j} f_j(t_j)dt_j$$

$$\frac{\partial}{\partial x_i}F(\vec{x}) = \prod_j^{(i)} \int_{-\infty}^{x_j} f_j(t_j)dt_j \cdot f_i(x_i)$$

stat 2.5
exA-3

変数変換後の独立性

XとYは独立,

$g(X)$, $h(Y)$は実数値連続関数

$\Downarrow$

$g(X)$と$h(Y)$も独立.

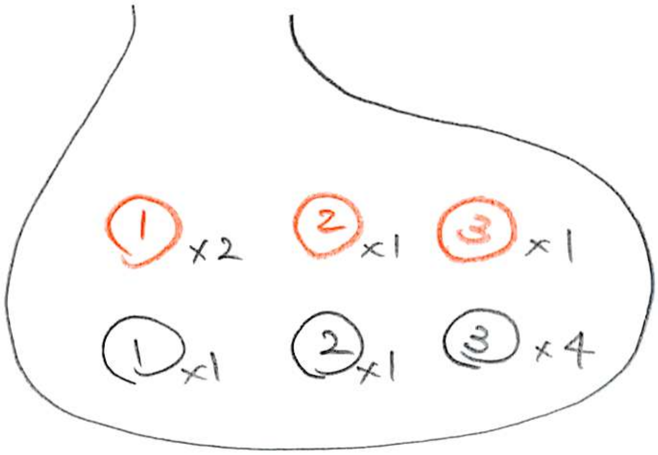
証明(i) $g(x) = x+a$, $h(y) = y+b$ の時

$f(x, y) = f_1(x) \cdot f_2(y)$ が成り立っている

$f($

2.7 確率変数の条件付確率分布

図2.5
(同じ)



問題

袋から取り出した玉が赤色であった時

⇓

数字が3である確率

確率空間(モデル)記述

玉の色を $X \in \{1=$赤$, 2=$白$\}$
玉の数字を $Y \in \{1, 2, 3\}$

} 確率モデル

問題の記述

$$P(\underbrace{Y=3}_{\text{事後}} \mid \underbrace{X=1}_{\text{事前}}) = \frac{P(X=1, Y=3)}{P(X=1)}$$

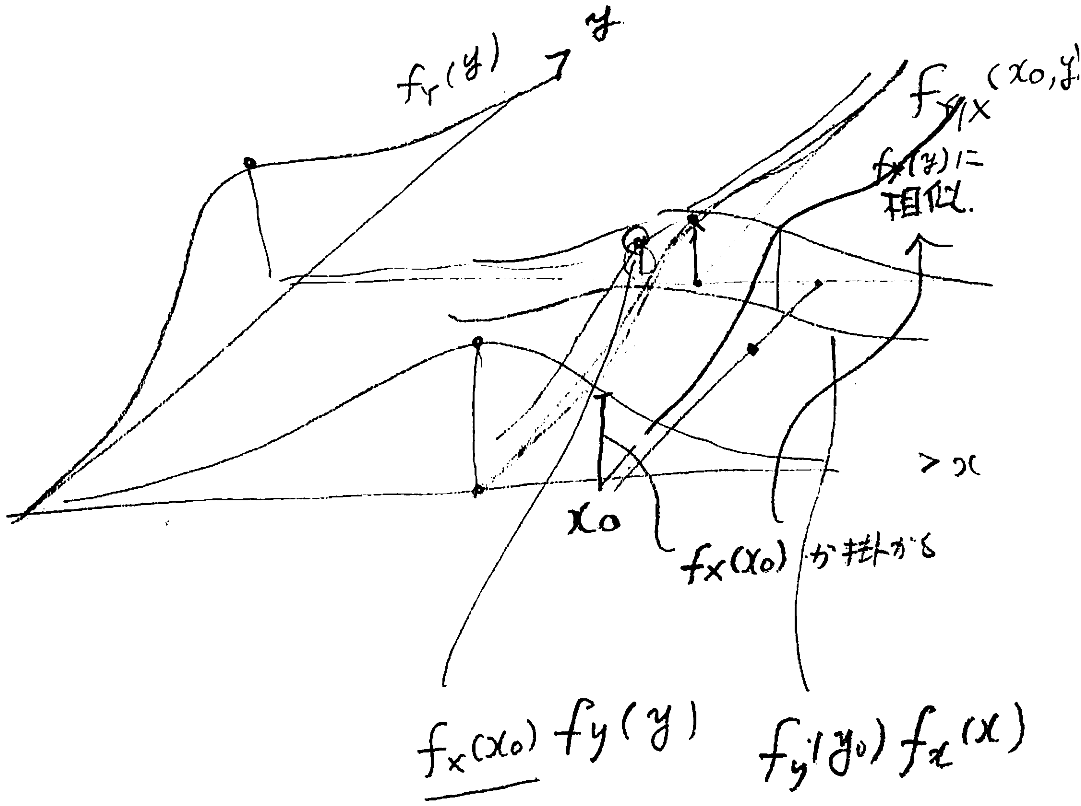
$$= \frac{f(1,3)}{f_X(1)}$$

← 同時 (分子)　← 周辺 (分母)

条件付密度関数

$X=x$のときに、$Y=y$である条件付密度関数

$$f_{Y|X}(\underbrace{y|x}_{?}) = \frac{f(x,y)}{f_X(x)} \quad \left(\overset{\text{独立なら}}{=} f_Y(y) \right)$$



条件付期待値

$$E_{Y|X}[g(X,Y) \mid X=x]$$

$$= \int g(x,y)\, f_{Y|X}(y|x)\, dy$$

$$E_X[E_{Y|X}[g(X,Y) \mid X=x]]$$

$$= \int \left(\int g(x,y)\, f_{Y|X}(y|x)\, dy \right) f_X(x)\, dx$$

$$= \iint g(x,y)\, \underline{f_{Y|X}(y|x)\, f_X(x)}\, dy\, dx$$

$$= \iint g(x,y)\, f(x,y)\, dy\, dx$$

$$= E[g(X,Y)]$$

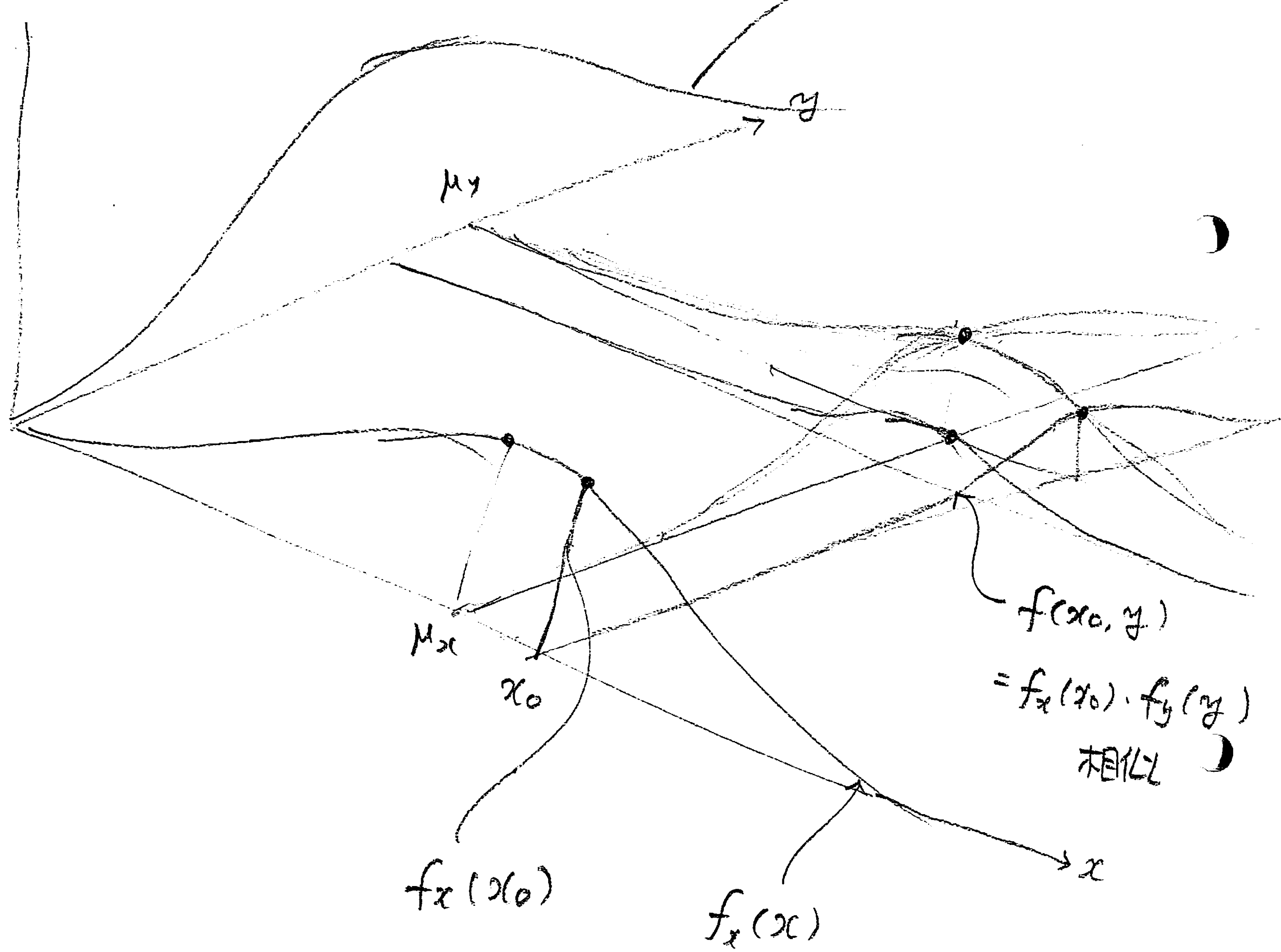
$N = e^{-(x^2+y^2)/2} \frac{1}{\sqrt{2\pi}}$

$N_{(x,y)} = e^{-((x-5)^2+(y-3)^2)/2} \frac{1}{\sqrt{2\pi}}$

$N(0,y)$, $(x=0)$

$N(x,0)$ & $(y=0)$

curve 3D



sageで描画しよう.

父親の身長から息子の伸長を予測する (A 2.5 演習問題)

model

父親の伸長 X

息子の身長 Y

$g(X)$: Yを予測する関数

誤差

$E[\{Y-g(X)\}^2]$ 平均的な誤差

最小誤差

誤差を最小にする $g(X)$ は?

条件付平均 $E_{Y|X}[Y|X]$ となる

$(Y-g(x))^2 = [Y^2 - 2Y \cdot g(x) + g^2(X)$

$Y = g(X)$ ならば最小　Yの分布と $g(X)$ の分布が等しければ

$f_x(y)$

A.2.5の解

ポイントは $E_{Y|X}[g(X)|\overset{=x}{X}] = g(X)$

$$E[h(X,Y)] = E_X[E_{Y|X}[h(X,Y)|\overset{=x}{X}]]$$

$$\int \underset{f_X(x)}{)} \left[\int h(x,y) \frac{f(x,y)}{f_X(x)} dy\right] dx$$

$E[\{Y-g(X)\}^2]$

$= E[\{(Y-E_{Y|X}[Y|X]) + (E_{Y|X}[Y|X] - \underline{g(X)})\}^2]$

? $= E_X[E_{Y|X}[\{ \quad〃 \quad \}^2 | X]]$

$$E_{Y|X}\left[\{\underbrace{(Y - E_{Y|X}[Y|X])}_{} + \underbrace{(E_{Y|X}[Y|X] - g(X))}_{}\}^2\right]$$

$$\{Y - g(X)\}^2 = \{(Y - E_{XY}) + (E_{YX} - g(X))\}^2$$

$$E[\{(Y - E_{YX}) + (\underline{E_{YX} - g(X)})\}^2$$ → Xだけの式

$$= E_X[E_{YX}[\{\cancel{(Y - E_{YX}) + (\quad)}\}^2 | X]]$$

$$= E_X$$

$$E_{YX}[\underline{(Y - E_{YX})(E_{YX} - g(X))}|X]$$

$$Y \cdot g(X) - Y E_Y + g(X E_{XY} + (E_{YX})^2$$

2.8 確率とモーメントに関連した不等式

Chebyshev's

$$\mu = E[X],\ \sigma^2 = V[X] = E[X^2] - E[X]^2$$

$$P(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

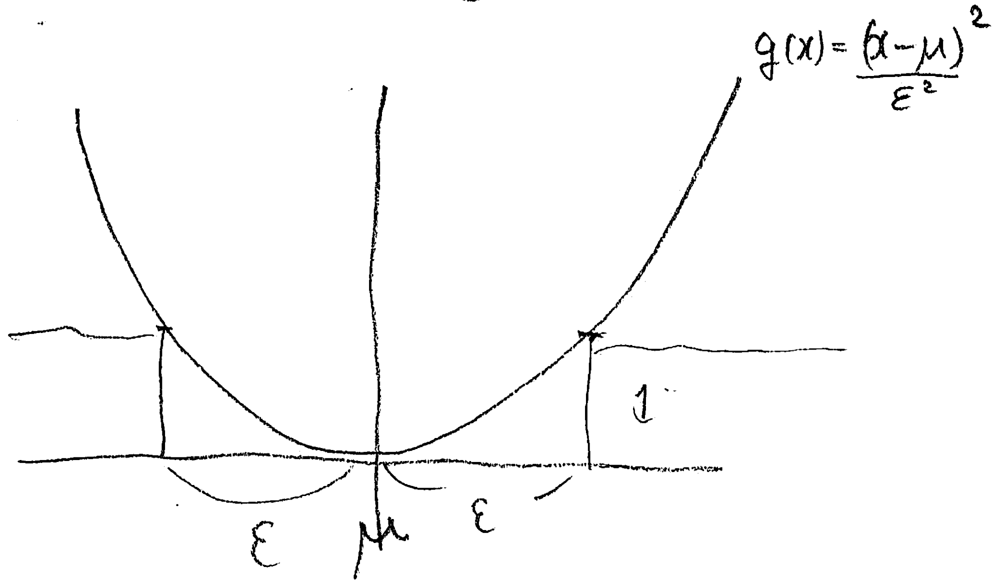
証明

$$P(|X-\mu| \geq \varepsilon) = \int_{|x-\mu| \geq \varepsilon} f(x)\,dx$$

$$\leq \int_{|x-\mu| \geq \varepsilon} \left(\frac{|x-\mu|}{\varepsilon}\right)^2 f(x)\,dx$$

$$\leq \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\varepsilon^2} f(x)\,dx$$

$$= \frac{\sigma^2}{\varepsilon^2}$$



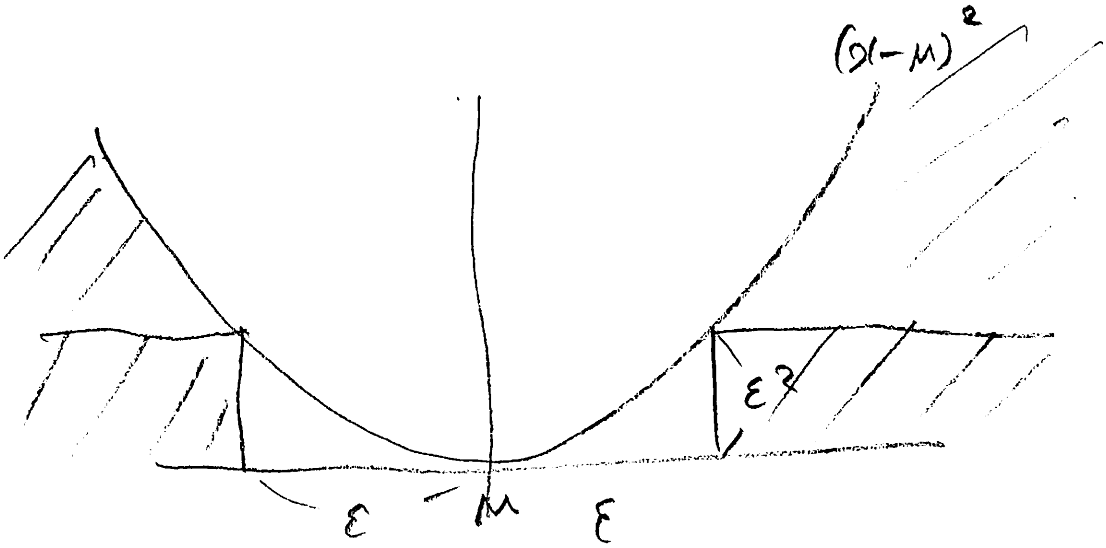
Chebyshev's
こっちの方が
わかり易い

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

$$\geq \int_{|x-\mu| \geq \varepsilon} (x-\mu)^2 f(x)\,dx$$

$$\geq \int_{|x-\mu| \geq \varepsilon} \varepsilon^2 f(x)\,dx$$

$$= \varepsilon^2 \int_{|x-\mu| > \varepsilon} f(x)\,dx = \varepsilon^2 P(|x-\mu| \geq \varepsilon)$$



平均 μ から ε 以上離れる確率は $\dfrac{\sigma^2}{\varepsilon^2}$

Cauchy-Schwartz's $(\sigma_{xy})^2 \leq \sigma_{xx} \cdot \sigma_{yy}$

X と Y：確率変数

$$\overset{\sigma_{xy}^2}{Cov[X,Y]^2} \leq \overset{\sigma_{xx}\ ((\sigma_x)^2)}{V[X]} \cdot \overset{\sigma_{yy}\ ((\sigma_y)^2)}{V[Y]}$$

共分散　　分散

$$E[(X-\mu_X)(Y-\mu_Y)]^2 \leq E[(X-\mu_X)^2] \cdot E[(Y-\mu_Y)^2]$$

証明

$$a = \sigma_{yy},\quad b = -\sigma_{xy},\quad Z = a(X-\mu_x) + b(Y-\mu_y)$$

この時、次式が成り立つ

$$0 \leq E[Z^2] = a^2(\sigma_x)^2 + b^2(\sigma_y)^2 + 2ab\,\sigma_{xy}$$

$$= (\sigma_{yy})^2(\sigma_{xx}) + (\sigma_{xy})^2 \cdot (\sigma_{yy}) - 2(\sigma_{yy})(\cancel{\sigma_{xy}})(\sigma_{xy})$$

$$= (\sigma_{yy})(\sigma_{yy}\sigma_{xx} + \sigma_{xy}^2 - 2\sigma_{xy}\sigma_{xy})$$

$$= \sigma_{yy}(\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2)$$

$\sigma_{yy} > 0$

$$\sigma_{yy} > 0 \rightarrow \sigma_{xx}\sigma_{yy} - \sigma_{xy}^2 \geq 0$$

$$\therefore V[X] \cdot V[Y] \geq Cov[X,Y]$$

$\sigma_{yy} = 0$

$\sigma_{yy} = 0 \rightarrow Y = E[Y]$ なので、

$$\sigma_{xy} = 0$$

Cauchy-schwaltz' の続き.

等号が成り立つ場合を考える

$$Z = a(X-\mu_x) + b(Y-\mu_y)$$

$$E[Z^2] = a^2\sigma_{xx} + b^2\sigma_{yy} + 2ab\sigma_{xy}$$

$a=\sigma_{yy}$, $b=-\sigma_{xy}$,

$$= \sigma_{yy}^2\sigma_{xx} + \sigma_{xy}^2\sigma_{yy} - 2\sigma_{yy}\sigma_{xy}^2$$

$$= \sigma_{yy}(\sigma_{xx}\sigma_{yy} - \sigma_{xy}^2)$$

∴ 等号が成立すれば

$E[Z^2]=0$ なので 確率1で $Z=0$

$c = -a\mu_x - b\mu_y$ とおけば

$aX + bY + c = 0$ が確率1で成り立つ.

$$\therefore\ a(X-\mu_x) + b(Y-\mu_y) = 0$$

$$\sigma_{yy}(X-\mu_x) - \sigma_{xy}(Y-\mu_y) = 0$$

直線の式.

Jensen's

X : 確率変数

h(x) : 連続, 下に凸　　x^2とか $-\log(x)$

$\Downarrow$

$$E[h(X)] \geq h(\underbrace{E[X]}_{\mu})$$

証明

二回微分可能なら

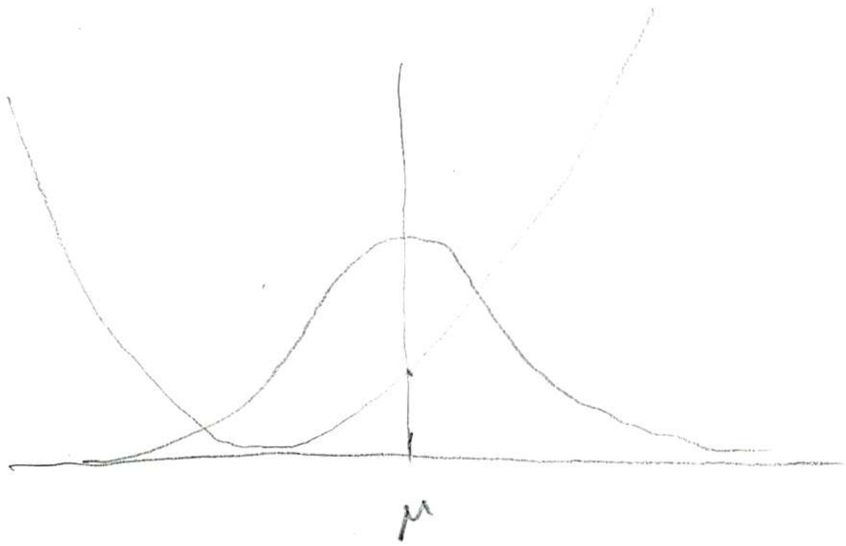
$h(x)$が下凸 $\Leftrightarrow$ $h''(x) \geq 0$

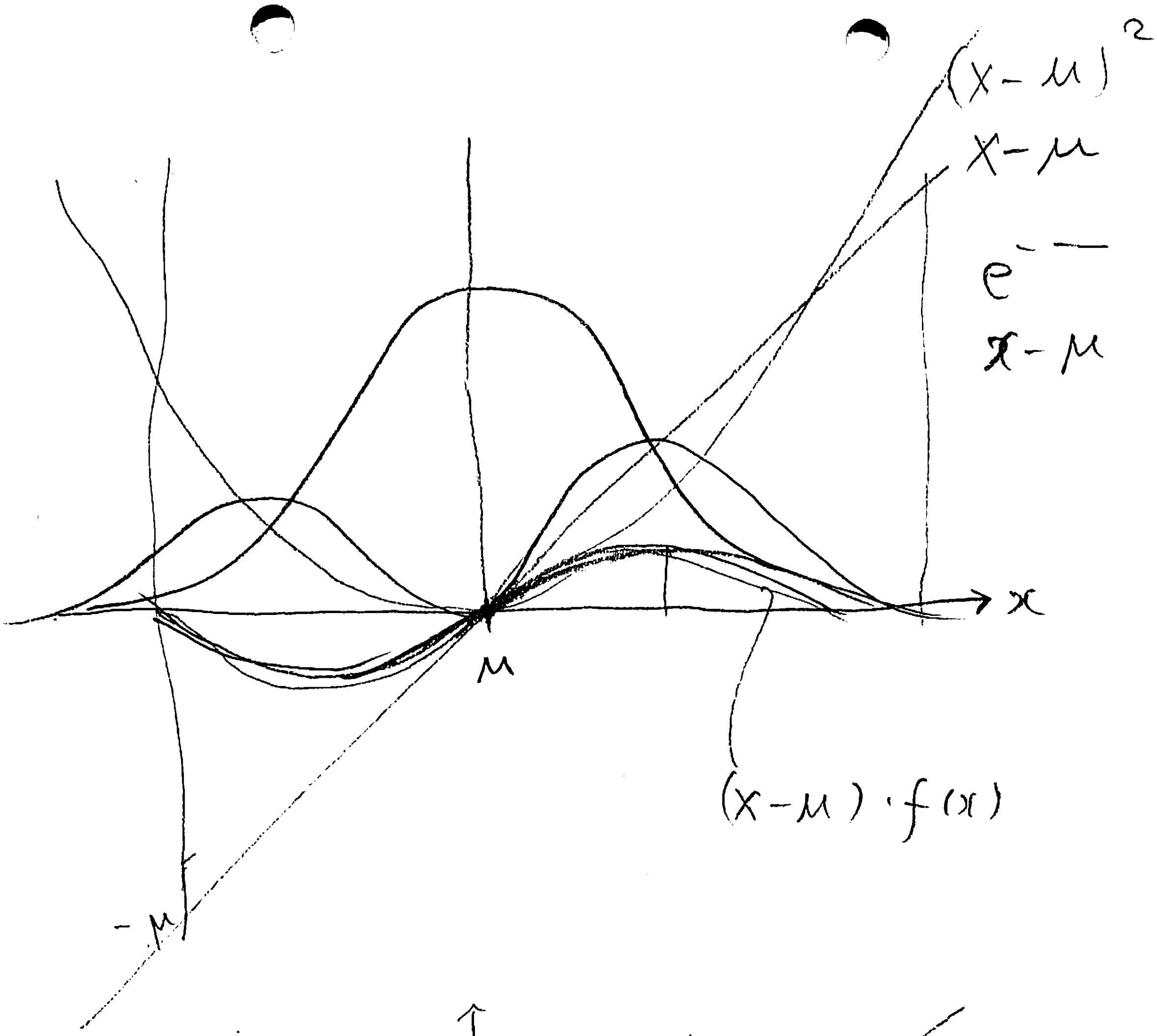
$E[X]=\mu$とおいて,　$h''(x) \geq 0$

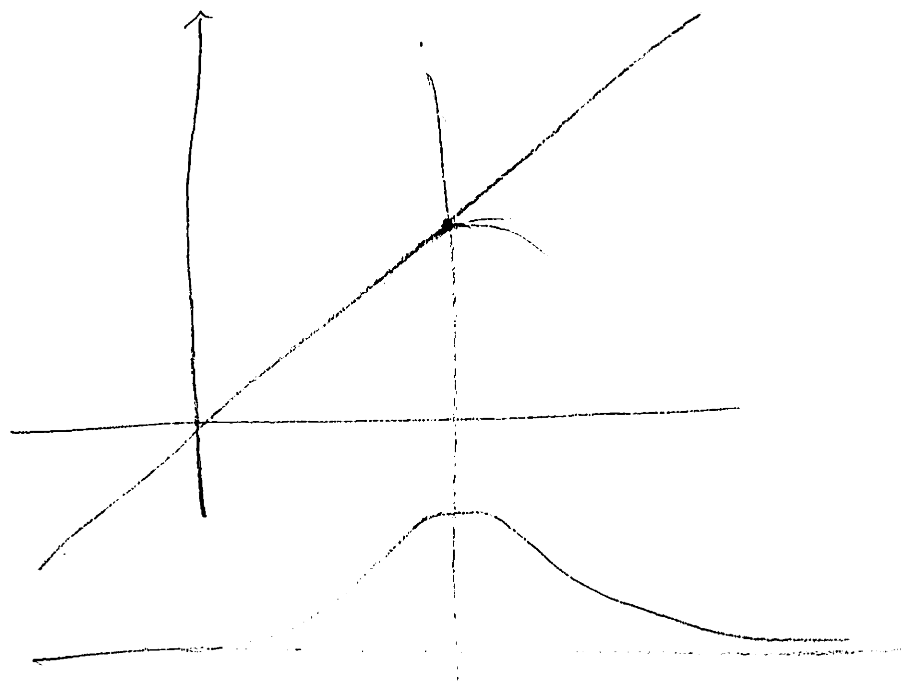
$$h(X) \geq h(\mu) + h'(\mu)(\underline{X-\mu}) \quad + \underbrace{h''(\mu)}_{\geq 0}(X-\mu)^2$$

$$E[h(X)] \geq h(E[X]) + h'(\mu)\underbrace{E[X-\mu]}_{=0}$$

$$\therefore E[h(X)] \geq h(E[X])$$




$(x-\mu)^2$
$x-\mu$
e^{-}
$x-\mu$
x
μ
$(x-\mu)\cdot f(x)$
$-\mu$



2.9 確率model

サイコロ投げに対するmodel

確率変数による確率の定ギ model へ

$$\left(\begin{array}{ll} \text{確率変数} & X \in \{1, \cdots, 6\} \\ \text{確率分布} & P(X=k) = \frac{1}{6} \end{array} \right.$$

標本空間 $\Omega = \{\omega_1, \omega_2, \cdots, \omega_6\}$

サイコロの目が x である標本点 ε ω_x

シグマ集合体（事象の集合）

$$\mathcal{A} = 2^{\Omega}$$

確率model上での確率をどう自然に誘導するか？

xの目が出る事象

$$X(\omega_x) = x, \quad (x = 1, 2, \cdots, 6)$$

$$A_x = \{\omega_x\} = \{\omega \in \Omega : X(\omega) = x\}$$

サイコロの目 x がでる事象.

$$P(A_x) = P(X = x) = 1/6 \quad (x = 1 \sim 6)$$

偶数の目がでる事象

$$A_{2,4,6} = A_2 \cup A_4 \cup A_6$$

$$\begin{aligned} P(A_{2,4,6}) &= P(A_2 \cup A_4 \cup A_6) \\ &= P(A_2) + P(A_4) + P(A_6) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\ &= 1/2 \end{aligned}$$

全ての事象に対する確率を定める

全ての事象は 2^Ω の要素 e で，$e \subset \Omega$，
$e_\alpha = \{w_i \mid i = \mathrm{index}(e_\alpha)\}$
↳ この各 w_i に対する $P(w_i)$ の和．

離散型確率 model

確率変数 X が可算個の離散値をとり，
X の値に対応する pdf が定っている

連続型

$(\Omega, \mathcal{A}, P)$ が定っていたとする

R　区間

確率変数 X は，関数で

$$X : \Omega \to R$$

$$X^{-1}((-\infty, x]) = \{\omega : X(\omega) \leq x\} \in \mathcal{A}$$
$$(x \in R)$$

連続確率モデル

$\Omega = \mathbb{R}$　実数直線

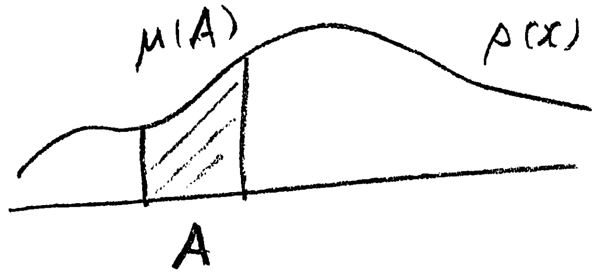
$\mathbb{B}$ = ボレル集合族.

$\rho(x)$ は積分可能な正値関数で

$$\int_{-\infty}^{\infty} \rho(x)\,dx = 1$$

$A \in \mathbb{B}$ に対して

$$\mu(A) = \int_A \rho(x)\,dx$$



n次元拡張

$$\Omega = \mathbb{R}^n$$

$\mathbb{B}$ = n次元のボレル集合族.

$\rho(\vec{x})$:　$\int \rho(\vec{x})\,d\vec{x} = 1$

$$\mu(A) = \int_A \rho(\vec{x})\,d\vec{x}$$