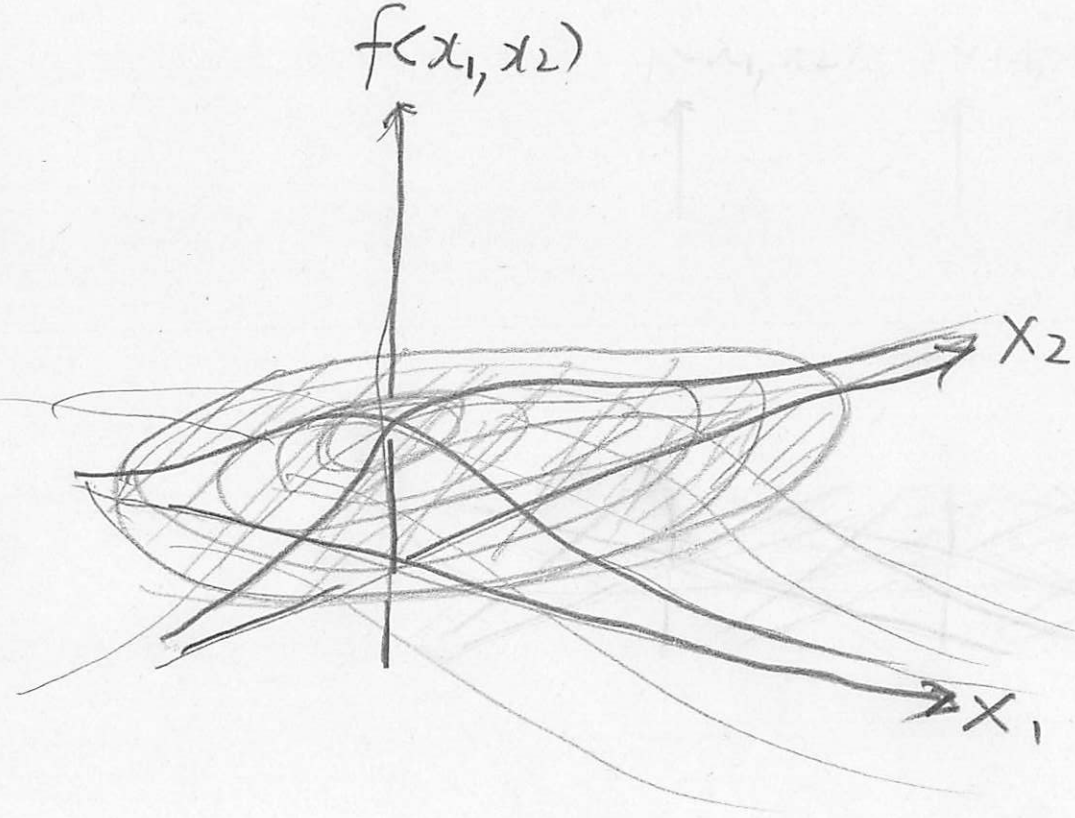
4章 二つ以上の確率変数の分布



○ 2以上は
本質は同じ

○ 連続と離散は
$\int$ と Σ の違い

2つ以上の
確率変数の分布

同時確率密度 f.

$^{\forall} X_1, X_2 \ ^{\exists} f(x_1, x_2)$

① $f(x_1, x_2) \geq 0$

② $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 = 1$

③ $P(a < X_1 < b, c < X_2 < d)$
$= \int_a^b \int_c^d f(x_1, x_2) dx_1 dx_2$

$f(x_1, x_2)$ は X_1, X_2 の
同時確率密度 f —— 確率変数の独立 —— 確率周辺密度 f

X_1 と X_2 に対し、

$\exists f(x_1, x_2) \geq 0$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1\, dx_2 = 1$$

$\forall a, b, c, d:\ a<b,\ c<d$

$$P(a < X_1 < b,\ c < X_2 < d)$$

$$= \int_a^b \int_c^d f(x_1, x_2)\, dx_2\, dx_1$$

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2 \quad = P(a < X_1 < b,\ -\infty < X_2 < \infty) \quad = P(a < X_1 < b)$$

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1 \qquad \text{〃}$$

X_1 と X_2 が互いに
独立 $\equiv$ $f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2) \quad (x_1, x_2) \in R^2$

$$P(a < X_1 < b,\ c < X_2 < d)$$

$$= \int_a^b \int_c^d f(x_1, x_2)\, dx_2\, dx_1$$

$$= \int_a^b f_1(x_1)\, dx_1 \int_c^d f_2(x_2)\, dx_2$$

$$= P(a < X_1 < b) \cdot P(c < X_2 < d)$$

$f(x_1, x_2)$ が given

$$P(a < X_1 < b)$$

$$= P(a < X_1 < b, -\infty < X_2 < +\infty)$$

$$= \int_a^b \left\{ \underbrace{\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2}_{} \right\} dx_1$$

X_1の密度f. → $f_1(x_1)$

X_1の周辺確率密度f.

X_2について

$$\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_1$$

③

X_1, X_2 の同時密度 f

$f(x_1, x_2)$

X_1 と X_2 が独立の時

周辺密度 $f_1(x_1), f_2(x_2)$

$P(a<X_1<b, c<X_2<d)$ if $f(x_1, x_2) = f_1(x_1) f(x_2)$

の時は、X_1, X_2 は独立。

$= \int_a^b \int_c^d f(x_1, x_2)\, dx_1 dx_2$

$= \int_a^b f_1(x_1)\, dx_1 \cdot \int_c^d f_2(x_2)\, dx_2$

$= P(a<X_1<b) \cdot P(c<X_2<d)$

確率変数

n個

$X_1, \cdots, X_n$

同時密度関数

$$f(\overset{\vec{x}}{x_1, x_2, \cdots, x_n})$$

$$f(\overset{\vec{x}}{x_1, \cdots, x_n}) \geq 0$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\vec{x}) d\vec{x} = 1$$

$$\vec{a} < \vec{b}$$

$$P(\vec{a} < \vec{X} < \vec{b})$$

$$= \int_{\vec{a}}^{\vec{b}} f(\vec{x}) d\vec{x}$$

$$f_1(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_n) dx_2 \cdots dx_n$$

X_1の周辺密度 f.

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

$X_1, X_2, \cdots, X_n$は互いに独立

$$P(\vec{a} < \vec{X} < \vec{b}) = P(a_1 < X_1 < b_1) \cdots \cdots P(a_n < X_n < b_n)$$

$$M_{X_1 + \cdots + X_n}(t) = M_{X_1}(t) \cdot \cdots \cdots M_{X_n}(t)$$

$\varphi(X_1, X_2)$の
平均

$$E\{\varphi(X_1, X_2)\} \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \varphi(x_1, x_2) f(x_1, x_2)\, dx_1 dx_2$$

定理1　X_1とX_2は独立
$\varphi(X_1)$, $\varphi(X_2)$も独立

$$\boxed{E(\varphi_1(X_1)\cdot\varphi_2(X_2)) = E(\varphi_1(X_1))\cdot E(\varphi_2(X_2))}$$

系1.　$E(X_1 X_2) = E(X_1)(X_2)$

独立性より

$$f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$$

同辺

$$\varphi(x_1, x_2) = \varphi_1(x_1) \cdot \varphi_2(x_2)$$

定理1の証明

$$E(\varphi_1(X_1) \cdot \varphi_2(X_2))$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi_1(x_1)\varphi_2(x_2) f(x_1\ x_2) dx_1 dx_2$$

$$= \int_{-\infty}^{\infty} \varphi_1(x_1) f_1(x_1) dx_1 \int_{-\infty}^{\infty} \varphi_2(x_2) f_2(x_2) dx_2$$

$$= E(\varphi_1(X_1)) \cdot E(\varphi_2(X_2))$$

系2 X_1とX_2は互独立, 積率母関数

$$M_{X_1+X_2}(t) = M_{X_1}(t) \cdot M_{X_2}(t)$$

$$M_X(t) = E(e^{tX})$$

$$= \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$M_{X_1+X_2}(t) = E(e^{t(X_1+X_2)})$$

$$= E(e^{tX_1} \cdot e^{tX_2})$$

$$= E(e^{tX_1}) \cdot E(e^{tX_2})$$

$$= M_{X_1}(t) \cdot M_{X_2}(t)$$

例12

X_1, X_2: 独立

$X_1 \sim N(\mu_1, \sigma_1^2)$ $\quad f_1(x_1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_1-\mu_1)^2}{\cancel{\sigma} \sigma_1^2}\right)$

$X_2 \sim N(\mu_2, \sigma_2^2)$ $\quad f_2(x_2) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)$

$Y = X_1 - X_2$ はどんな分布

$M_{X_1}(t) = E(e^{tX_1}) = \exp(\mu_1 t + \frac{1}{2}\sigma_1^2 t^2)$

$M_{X_2}(t) = E(e^{tX_2}) = \exp(\mu_2 t + \frac{1}{2}\sigma_2^2 t^2)$

$M_{X_1 - X_2}(t) = E(e^{t(X_1 - X_2)})$

$= \cancel{\exp} E(e^{tX_1}) \cdot E(e^{-tX_2})$

$= \exp(\mu_1 t + \frac{1}{2}\sigma_1^2 t^2) \cdot \exp(-\mu_2 t + \frac{1}{2}\sigma_2^2 t^2)$

$= \exp((\mu_1 - \mu_2)t + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2)$

$Y = X_1 - X_2$

$\sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$

例13 独立 $X_1, X_2 \sim N(0,1)$

$$Y = X_1^2 + X_2^2$$

の密度f $g(y)$ を求める

$X_i^2 \sim \chi^2(1)$ なので $M_{X_i^2}(t) = (1-2t)^{-\frac{1}{2}}$

$t < \frac{1}{2}, \quad i=1,2$

$$M_Y(t) = M_{X_1^2}(t) \cdot M_{X_2^2}(t)$$

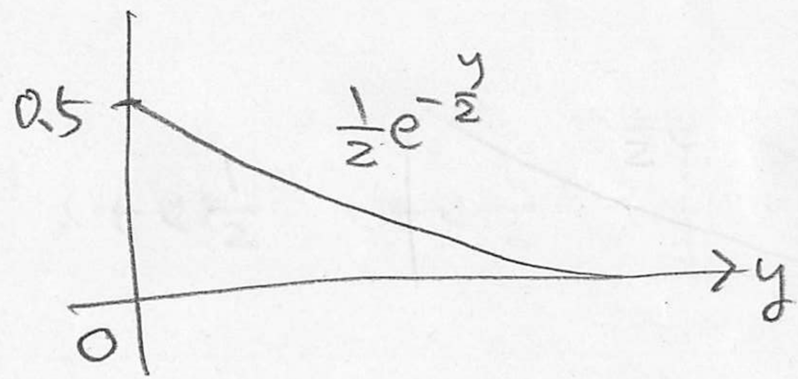
$$= (1-2t)^{-1}, \quad t < \frac{1}{2}$$

$\therefore Y \sim \chi^2(2)$

$$f_r(x) = \frac{1}{\Gamma(\frac{r}{2}) 2^{\frac{r}{2}}} x^{\frac{r}{2}-1} \cdot e^{-\frac{x}{2}} \quad 0<x<\infty$$

$$g(y) = \frac{1}{2} \exp\left(-\frac{y}{2}\right)$$

$0<y<\infty$



※独立な確率変数の和の分布

※独立 $X_1, X_2 \sim N(\mu_i, \sigma_i^2)$

定理2

$$Y = k_1X_1 + k_2X_2$$

$$\sim N(k_1\mu_1 + k_2\mu_2,\ k_1^2\sigma_1^2 + k_2^2\sigma_2^2).$$

定理3

X_1とX_2は独立でなくてよい

① $E(k_1X_1 + k_2X_2) = k_1E(X_1) + k_2E(X_2)$

② $V(k_1X_1 + k_2X_2)$

$$= k_1^2V(X_1) + k_2^2V(X_2)$$

$$E(k_1X_1+k_2X_2)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(k_1x_1+k_2x_2)f(x_1,x_2)dx_1dx_2$$

$$= k_1\int_{-\infty}^{\infty}x_1\left\{\int_{-\infty}^{\infty}f(x_1,x_2)dx_2\right\}dx_1$$

$$+k_2\int_{-\infty}^{\infty}x_2\left\{\int_{-\infty}^{\infty}f(x_1,x_2)dx_1\right\}dx_2$$

$$= k_1\int_{-\infty}^{\infty}x_1f_1(x_1)dx_1 + k_2\int_{-\infty}^{\infty}x_2f_2(x_2)dx_2$$

$$= k_1E(X_1) + k_2E(X_2)$$

$$E(k_1X_1+\cdots+k_nX_n)$$

$$= k_1E(X_1)+k_2E(X_2)+\cdots+k_nE(X_n)$$

$= \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$

$M_X(t) = E\{e^{tX}\}$

$Y = k_1 X_1 + k_2 X_2$

$\sim N(k_1\mu_1 + k_2\mu_2, k_1^2\sigma_1^2 + k_2^2\sigma_2^2)$

の証明

$$M_Y(t) \overset{\text{定義}}{=} E\{\exp(t(k_1X_1 + k_2X_2))\}$$

$$= E\{\exp(tk_1X_1)\cdot\exp(tk_2X_2)\}$$

$$= \exp\left\{(k_1\mu_1 + k_2\mu_2)t + \frac{1}{2}(k_1^2\sigma_1^2 + k_2^2\sigma_2^2)t^2\right\}$$

(2) 分散

$$V(k_1X_1+k_2X_2)=E\{[(k_1X_1+k_2X_2)-E(k_1X_1+k_2X_2)]^2\}$$

$$=E\{[k_1(X_1-\mu_1)+k_2(X_2-\mu_2)]^2\}$$

$$=k_1^2E\{(X_1-\mu_1)^2\}+k_2^2E\{(X_2-\mu_2)^2\}$$

$$+2k_1k_2\underbrace{E\{(X_1-\mu_1)(X_2-\mu_2)\}}_{\text{共有散 (co variance)}} \rightarrow X_1X_2-\mu_1\underbrace{X_2}_{\mu_2}-\mu_1\underbrace{X_1}_{\mu_1}+\underbrace{\mu_1\mu_2}_{E(X_1)E(X_2)}$$

$$Cov(X_1,X_2)=E(X_1X_2)-\underbrace{E(X_1)E(X_2)}_{0}$$

X_1 と X_2 が互独立なら

$$Cov(X_1,X_2)=0$$

$$\therefore V(k_1X_1+k_2X_2)=k_1^2V(X_1)+k_2^2V(X_2)$$

定理4 $X_1, \cdots, X_n \sim \chi^2(r_i)$ $i=1:n$

この時 $Y = X_1 + X_2 + \cdots + X_n$ は

自由度 $r_1 + \cdots + r_n$ の χ^2 分布

再生的

$$M_{X_i}(t) = E\{e^{tX_i}\} = (1-2t)^{-\frac{r_i}{2}}$$

$$t < \frac{1}{2}$$

$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdots M_{X_n}(t)$ なので

$$M_Y(t) = (1-2t)^{-\frac{1}{2}\sum r_i}, \quad t < \frac{1}{2}$$

例4 X_1とX_2は独立

$X_1 \sim$ パラメータ2の指数分布

$X_2 \sim \chi^2(4)$

$Y = 4X_1 - X_2$の時

$E(X_1) = \frac{1}{2}$ $E(X_2) = 4$

$V(X_1) = \frac{1}{4}$ $V(X_2) = 8$

$E(4X_1 - X_2) = 4E(X_1) - E(X_2) = -2$

$V(4X_1 - X_2) = 4^2 V(X_1) + V(X_2) = 12$

確率変数の和の確率密度関数

X_1 と X_2 は互いに独立

$f(x_1), f(x_2)$ が密度 f.

$Y = X_1 + X_2$, その密度 $f \cdot g(y)$

$$g(y) = \int_{-\infty}^{\infty} f_2(y - x_1) \cdot f_1(x_1)\, dx_1$$

$$= \int_{-\infty}^{\infty} f_1(y - x_2) \cdot f_2(x_2)\, dx_2$$

証明

X_1, X_2, Y の分布関数

$F_1(x), F_2(x), G(y)$ とする

$$G(y) = P\{X_1 + X_2 \leq y\}$$

$$= \iint_{x_1+x_2 \leq y} f(x_1, x_2)\, dx_1 dx_2$$

$$= \int_{-\infty}^{\infty} f_1(x_1) \left\{ \int_{-\infty}^{y-x_1} f_2(x_2)\, dx_2 \right\} dx_1$$

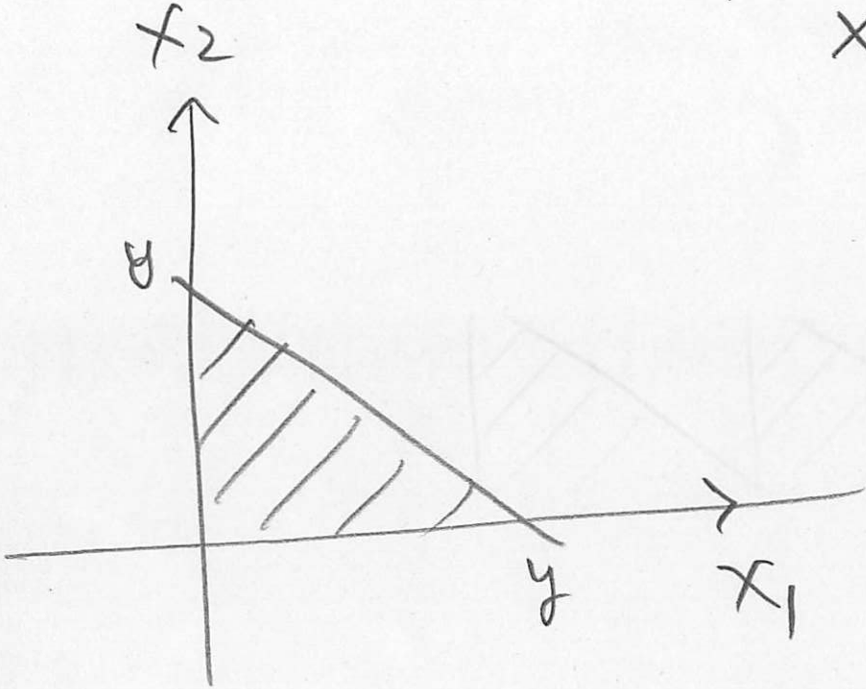
$$= \int_{-\infty}^{\infty} f_1(x_1)\, F(y - x_1)\, dx_1$$

$$\because F(\infty) = 0$$

$$g(y) = G'(y)$$

$$= \int_{-\infty}^{\infty} f_1(x_1) f_2(y - x_1)\, dx_1$$

$$g(y) = \int_{-\infty}^{\infty} f_2(x_2) f_1(y - x_2)\, dx_2$$



$X_1, X_2 > 0$

$$g(y) = \int_0^y f_1(x_1) f_2(y - x_1)\, dx_1$$

$$= \int_0^y f_2(x_2) f_1(y - x_2)\, dx_2$$

例5 X_1, X_2 互独立

$\sim \lambda$ の指数分布

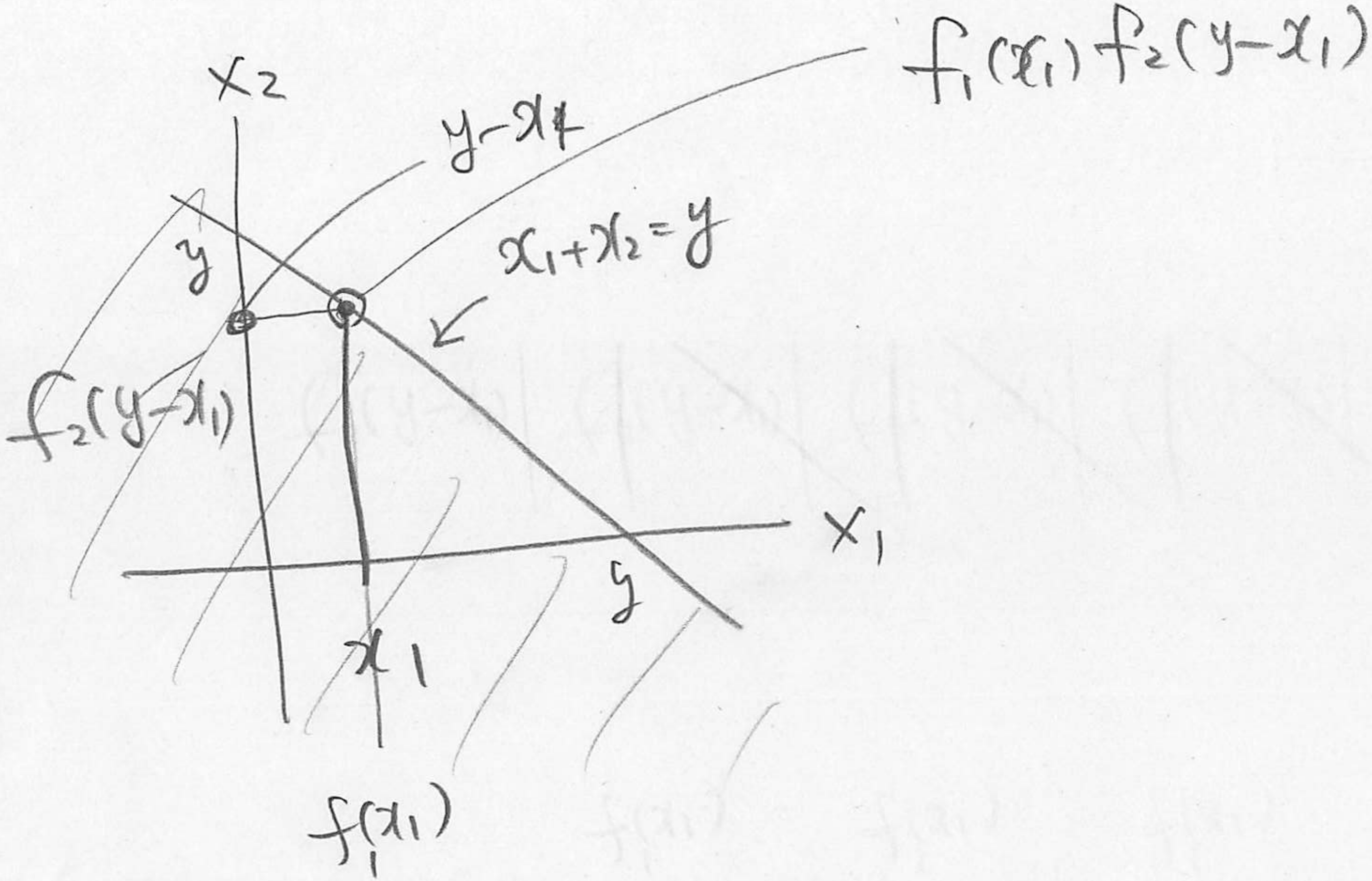
$Y = X_1 + X_2$

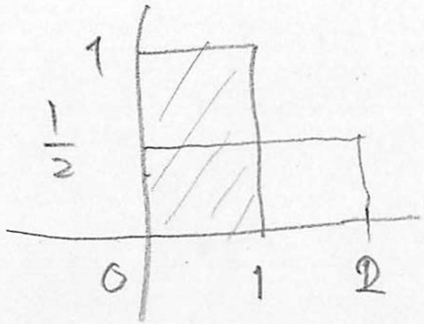
X_i の密度 f $\lambda e^{-\lambda x_i}$

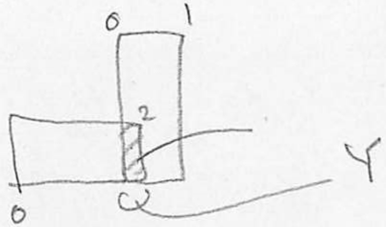
$$g(y) = \int_0^y \lambda e^{-\lambda(y-x_1)} \lambda e^{-\lambda x_1} dx_1$$

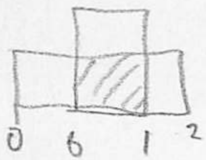
$$= \int_0^y \lambda^2 e^{-\lambda y} dx_1$$

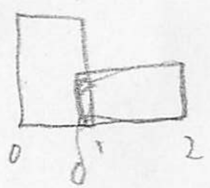
$$= \lambda^2 e^{-\lambda y} \quad (0 < y < \infty)$$


$f_1(x_1) f_2(y-x_1)$
X_2
$y-x_1$
$x_1+x_2=y$
y
$f_2(y-x_1)$
X_1
y
x_1
$f_1(x_1)$









例6

独立 X_1, X_2

$$X_1 \sim f_1(x_1) = 1,\quad 0 < x_1 < 1$$

$$X_2 \sim f_2(x_2) = \frac{1}{2},\quad 0 < x_2 < 2$$

の時

$Y = X_1 + X_2$ の確率密度 f

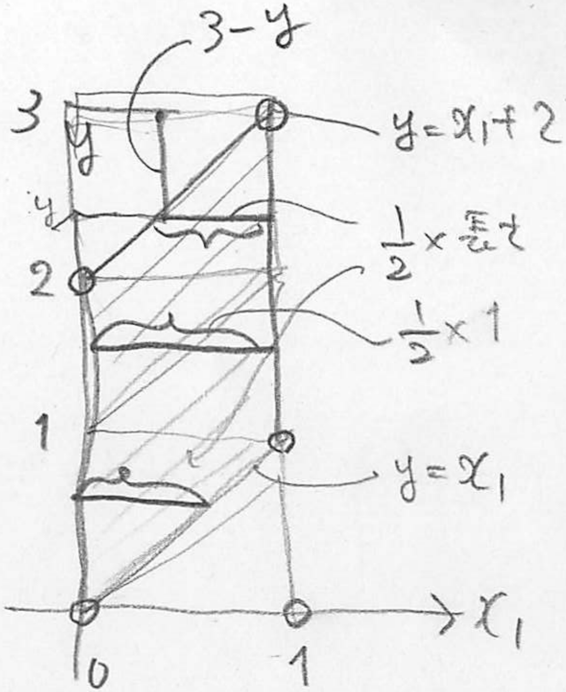
$$g(y) = \int_{-\infty}^{\infty} \underbrace{f(y - x_1)}_{\frac{1}{2}[0 < y - x_1 < 2]} \underbrace{f(x_1)}_{1\,[0 < x_1 < 1]} dx_1$$

$$\frac{1}{2}[0 < y - x_1 < 2] \iff x_1 < y < x_1 + 2$$

$$= \int_0^y \frac{1}{2} dx_1 = \frac{y}{2} \quad 0 < y \le 1$$

$$= \int_0^1 \frac{1}{2} dx_1 = \frac{1}{2} \quad 1 \le y < 2$$

$$= \int_{3-y}^1 \frac{1}{2} dx_1 = \frac{-1}{2}(y - 3) \quad 2 \le y < 3$$



例16 独立 X_1, X_2

$$X_1 \sim f_1(x_1) = 1 \quad 0 < x_1 < 1$$

$$X_2 \sim f_2(x_2) = \frac{1}{2} \quad 0 < x_2 < 2$$

の時

$Y = X_1 + X_2$ の確率分布fは

$0 \leq y \leq 3$ で $\neq 0$

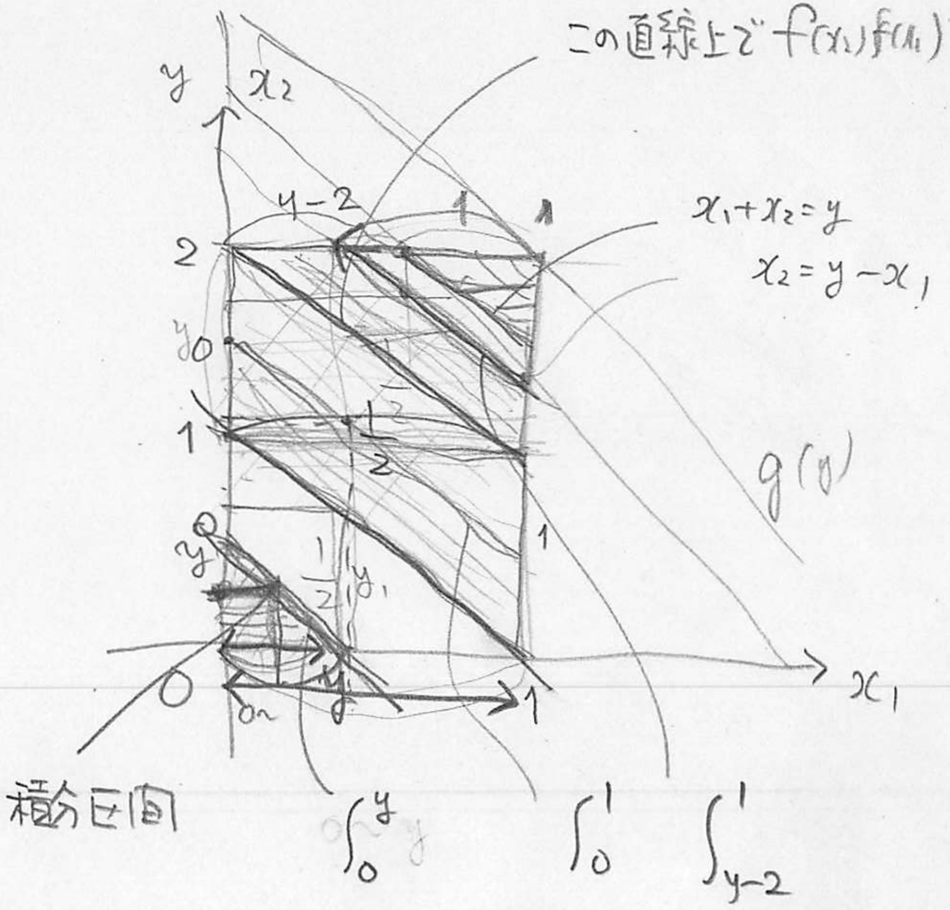
$$g(y) = \int_{-\infty}^{\infty} \underbrace{\overbrace{\boxed{f(y-x_1)}}^{\frac{1}{2}}}_{0<y-x_1<2\text{で定ぎ}} \overbrace{\boxed{f(x_1)}}^{1,\ 0<\ <1} dx_1$$

$f(y-x_1)$ $0 < y-x_1 < 2$ で定義

$$x_1 < y < 2 + x_1$$

$$0 < x_1 < 1$$

$$0 < y < 3$$



4.3 無作為標本

定義3.

互独立 $X_1, \cdots, X_n$ 同じ分布に従う $\Rightarrow$ 大きさ n の無作為標本 (random sample)

とする $\begin{cases} X_1, \cdots X_n \text{ の同時密度 } f(x_1, \cdots, x_n) \\ X_i \text{ の周辺密度f } f_i(x_i) \\ \text{共通密度f } f(x) \end{cases}$

$X_1 \cdots X_n$ が互いに独立

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

同じ分布に従う $= f(x_1) \cdots f(x_n)$

無作為標本ならば

$f(x_1, \cdots, x_n) = f(x_1) \cdots f(x_n)$

定義4. 無作為標本 $X_1 \ldots X_n$ だけの関数 $Y = \varphi(X_1, \cdots, X_n)$ を統計量 (statistic)

例えば

$$X = \frac{1}{n}(X_1 + \cdots + X_n)$$

標本平均 sample mean

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

標本分散

無作為標本

定理6 $X_i \sim N(\mu, \sigma^2)$ にしたがう無作為標本とする

標本平均

$$\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n)$$

$$\sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

平均は同じで
分散は $1/n$ になる

定理2

$$E(\textstyle\sum k_i X_i) = \sum_i k_i E(X_i) \longrightarrow k_i = \frac{1}{n} \qquad E(\sum \frac{1}{n} X_i) = \frac{1}{n}\sum \mu_i$$

$$V(\textstyle\sum k_i X_i) = \sum k_i^2 V(X_i) \qquad V(\sum \frac{1}{n} X_i) = \frac{1}{n^2}\sum \sigma_i^2 = \frac{1}{n}\sigma^2$$

定理7

母集団

「$X_1, \cdots X_n$ は $N(\mu, \sigma^2)$ からの

大きさnの無作為標本」とする

$$\sum\left(\frac{X_i-\mu}{\sigma}\right)$$

$$\chi^2 = \frac{1}{\sigma^2}\left\{\sum(X_i-\mu)^2\right\}$$ は

自由度 n の χ^2 分布に従う

$X_i \sim N(\mu, \sigma^2)$

$\rightarrow \left(\frac{X_i-\mu}{\sigma}\right) \sim N(0,1)$

$\rightarrow \left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi_1^2(1)$

定理4

$X_i \sim \chi^2(r_i)$ の時

$Y = \sum X_i$ は 自由度 $\sum r_i$ の χ^2 分布に従う

$\chi^2_{\sum r_i}$

$$\sum_1^n \left(\frac{X_i-\mu}{\sigma}\right) \sim \chi^2_{\underline{n}}$$

再生性