

# 応用システム工学

## 第一回 確率統計の復習

平成21年5月29日

# 確率について1

- 確率変数とは
  - 値が $x$ となる確率 $p$ が与えられている変数 $X$ 
    - $X$ の取り得る値は離散値・連続値どちらでも可
- 事象とは
  - 確率変数 $X$ が取り得る値の部分集合
  - 事象に対する確率の表し方
    - $P(X=1)$  連続値・離散値
    - $P(5 \leq X < 10)$  連続値
    - $P(X \in A)$  離散値

# 確率について2

- 確率分布とは
  - 確率変数 $X$ と, 対応する確率 $P(X=x)$ の対応関係

- 確率分布関数とは
  - 確率変数 $X$ が, ある $x$ までとる確率

$$F_X(x) = P(X \leq x)$$

- 性質

- 単調非減少

- $F_X(-\infty) = 0$

- $F_X(\infty) = 1$

※確率変数の種類(連続・不連続)  
に関係しない

# 確率密度関数

- 連続な確率変数: $X$ に対する確率密度関数: $f_X(x)$ 
  - $x \leq X \leq x + dx$  となる確率を  $f_X(x)dx$  とする時
  - 確率密度関数は以下の性質を持つ

$$F_X(x) = \int_{-\infty}^x f_X(x') dx'$$

$$f_X(x) = \frac{d}{dx} F_X(x)$$

$$f_X(x) \geq 0$$

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

# 一様分布と確率密度関数

- 一様分布の確率密度関数
  - 確率変数 $X$ が区間 $[a,b]$ で一様分布

$$f_X(x) = \begin{cases} \frac{1}{b-a} & (a \leq x \leq b) \\ 0 & x < a, b < x \end{cases}$$

# 正規分布と確率密度関数

- 正規分布
  - 別名ガウス分布
  - 平均  $\mu$  , 分散  $\sigma^2$  の確率変数  $Y$  の正規分布に対する確率密度関数

$$f(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]$$

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

# 標準正規分布と確率密度関数

- 標準正規分布
  - 平均0,分散1の正規分布
  - 標準正規分布の確率密度関数が

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right]$$

$$Y \sim N(0,1)$$

# 正規分布関数の確率密度関数1

- 正規分布した確率変数 $Y$ の確率密度関数

$$f(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]$$

- 性質

$$\begin{aligned} \int_{-\infty}^{\infty} f_Y(y) dy &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy \end{aligned}$$



# 正規分布関数の確率密度関数2

- 変数変換  $z = \frac{y - \mu}{\sigma} \Rightarrow \frac{dz}{dy} = \frac{d}{dy} \frac{y - \mu}{\sigma} = \frac{1}{\sigma}$   
 $y = \sigma z + \mu \quad y: -\infty \leftrightarrow \infty \Leftrightarrow z: -\infty \leftrightarrow \infty$

$$\begin{aligned} \int_{-\infty}^{\infty} f_Y(y) dy &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}z^2\right] \sigma dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}z^2\right] dz \Rightarrow \text{次ページ} \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} = 1 \end{aligned}$$

# 正規分布関数の確率密度関数3

- これを求める
  - 偶関数なので  $\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = 2 \int_0^{\infty} e^{-\frac{1}{2}z^2} dz$

- これを求めればよい  $2 \int_0^{\infty} e^{-\frac{1}{2}z^2} dz$ 
  - 変数変換  $x = \frac{z}{\sqrt{2}} \quad \frac{dz}{dx} = \sqrt{2}$

$$\int_0^{\infty} e^{-\frac{1}{2}z^2} dz = \sqrt{2} \int_0^{\infty} e^{-x^2} dx$$

$$\int_0^{\infty} e^{-x^2} dx = I \quad \text{を求める}$$

# 正規分布関数の確率密度関数4

- 二重積分を考える

$$\int_0^{\infty} e^{-x^2} dx \int_0^{\infty} e^{-y^2} dy = \int_0^{\infty} \int_0^{\infty} e^{-(x^2+y^2)} dx dy = I^2$$

– 変数変換(極形式)

$$x = r \cos \theta, y = r \sin \theta$$

- コーシーリーマン条件

$$dx dy = r dr d\theta$$

$$x, y : 0 \leftrightarrow \infty \Leftrightarrow r : 0 \leftrightarrow \infty, \theta : 0 \leftrightarrow \frac{\pi}{2} \quad \text{第一象限}$$

# 正規分布関数の確率密度関数5

$$\begin{aligned}\int_0^\infty \int_0^\infty e^{-(x^2+y^2)} dx dy &= \int_0^{\frac{\pi}{2}} \int_0^\infty e^{-(r^2 \cos^2 \theta + r^2 \sin^2 \theta)} r dr d\theta \\&= \int_0^{\frac{\pi}{2}} \int_0^\infty e^{-r^2} r dr d\theta = \int_0^{\frac{\pi}{2}} \int_0^\infty r e^{-r^2} dr d\theta \\&= \int_0^{\frac{\pi}{2}} \left[ \frac{-1}{2} e^{-r^2} \right]_0^\infty d\theta = \frac{-1}{2} \int_0^{\frac{\pi}{2}} [0 - (1)] d\theta \\&= \frac{1}{2} \int_0^{\frac{\pi}{2}} d\theta = \frac{1}{2} [\theta]_0^{\frac{\pi}{2}} = \frac{\pi}{4} = I^2\end{aligned}$$

$$I = \frac{\sqrt{\pi}}{2} \quad \Rightarrow \quad \int_{-\infty}^\infty e^{-\frac{1}{2}z^2} dz = 2 \int_0^\infty e^{-\frac{1}{2}z^2} dz = 2\sqrt{2}I = 2\sqrt{2} \frac{\sqrt{\pi}}{2} = \sqrt{2\pi}$$

# 期待値

- 確率変数の平均みたいなもの
- 期待値 $E[X]$

– 確率変数 $X$ が離散分布の場合

- $X$ の取り得る値の集合 $M$

$$E[X] = \sum_{x \in M} xP(X = x)$$

– 確率変数 $X$ が連続分布の場合

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)dx$$

# 正規分布関数の期待値1

- 正規分布した確率変数 $Y$ の確率密度関数

$$f(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]$$

- 期待値

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy$$

$$= \int_{-\infty}^{\infty} y \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy$$

# 正規分布関数の期待値2

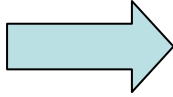
- 変数変換  $z = \frac{y - \mu}{\sigma} \Rightarrow \frac{dz}{dy} = \frac{d}{dy} \frac{y - \mu}{\sigma} = \frac{1}{\sigma}$   
 $y = \sigma z + \mu \quad y: -\infty \leftrightarrow \infty \Leftrightarrow z: -\infty \leftrightarrow \infty$

$$\begin{aligned} E[Y] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (z\sigma + \mu) e^{-\frac{1}{2}z^2} \sigma dz \\ &= \frac{1}{\sqrt{2\pi}} \left[ \sigma \int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz + \mu \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz \right] \end{aligned}$$

# 正規分布関数の期待値3

- 右辺第一項  $\frac{d}{dz} e^{-\frac{1}{2}z^2} = -\frac{1}{2} 2ze^{-\frac{1}{2}z^2} = -ze^{-\frac{1}{2}z^2}$

$$\int_{-\infty}^{\infty} ze^{-\frac{1}{2}z^2} dz = \left[ e^{-\frac{1}{2}z^2} \right]_{-\infty}^{\infty} = e^{-\infty} - e^{-\infty} = 0$$

- 右辺第二項  $\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = \sqrt{2\pi}$   正規分布関数の  
確率密度関数3  
参照

- 全体 
$$E[Y] = \frac{1}{\sqrt{2\pi}} \left[ \sigma \int_{-\infty}^{\infty} ze^{-\frac{1}{2}z^2} dz + \mu \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz \right]$$
$$= \frac{1}{\sqrt{2\pi}} [\sigma 0 + \mu \sqrt{2\pi}]$$
$$= \mu$$



# 正規分布関数の分散1

- 正規分布した確率変数 $Y$ の確率密度関数

$$f(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]$$

- 分散

$$V[Y] = \int_{-\infty}^{\infty} (y - \mu)^2 f_Y(y) dy$$

$$= \int_{-\infty}^{\infty} (y^2 - 2y\mu + \mu^2) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (y^2 - 2y\mu + \mu^2) \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy$$

# 正規分布関数の分散2

$$\begin{aligned} V[Y] = & \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y^2 \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy \\ & - \frac{2\mu}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy \\ & + \frac{\mu^2}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy \end{aligned}$$

# 正規分布関数の分散3

- 正規分布関数の確率密度関数1参照

$$\begin{aligned}\frac{\mu^2}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy &= \mu^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy \\ &= \mu^2\end{aligned}$$

- 正規分布関数の期待値2参照

$$\begin{aligned}\frac{-2\mu}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy &= \frac{-2\mu}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy \\ &= -2\mu^2\end{aligned}$$

# 正規分布関数の分散4

• 変数変換  $z = \frac{y - \mu}{\sigma} \Rightarrow \frac{dz}{dy} = \frac{d}{dy} \frac{y - \mu}{\sigma} = \frac{1}{\sigma}$

$y = \sigma z + \mu \quad y: -\infty \leftrightarrow \infty \Leftrightarrow z: -\infty \leftrightarrow \infty$

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y^2 \exp\left[-\frac{1}{2}\left(\frac{y - \mu}{\sigma}\right)^2\right] dy = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (z\sigma + \mu)^2 e^{-\frac{1}{2}z^2} \sigma dz$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (z^2\sigma^2 + 2z\sigma\mu + \mu^2) e^{-\frac{1}{2}z^2} \sigma dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (z^2\sigma^2 + 2z\sigma\mu + \mu^2) e^{-\frac{1}{2}z^2} dz$$

# 正規分布関数の分散5

$$\int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz = 0 \qquad \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}z^2\right] dz = \sqrt{2\pi}$$

$$\begin{aligned} \int_{-\infty}^{\infty} z^2 e^{-\frac{1}{2}z^2} dz &= \int_{-\infty}^{\infty} z z e^{-\frac{1}{2}z^2} dz = \left[ z \left( -e^{-\frac{1}{2}z^2} \right) \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -e^{-\frac{1}{2}z^2} dz \\ &= [0 - 0] + \sqrt{2\pi} = \sqrt{2\pi} \end{aligned}$$

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left( z^2 \sigma^2 + 2z\sigma\mu + \mu^2 \right) e^{-\frac{1}{2}z^2} dz &= \frac{1}{\sqrt{2\pi}} \left( \sigma^2 \sqrt{2\pi} + 2\sigma\mu \times 0 + \mu^2 \sqrt{2\pi} \right) \\ &= \sigma^2 + \mu^2 \end{aligned}$$

# 正規分布関数の分散6

$$\begin{aligned} V[Y] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y^2 \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy \\ &\quad - \frac{2\mu}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy \\ &\quad + \frac{\mu^2}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] dy \\ &= \{\sigma^2 + \mu^2\} + \mu^2 - 2\mu^2 = \sigma^2 \end{aligned}$$

# 多変数の確率分布3

## 2変数

- 周辺密度関数
  - 二つの確率変数 $X, Y$
  - $Y$ を無視した $X$ のみに関する確率密度関数を指す

$f_X(x)$

- 同時確率密度関数  $f_{XY}(x, y)$   
と周辺密度関数  $f_X(x)$   
の関係

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$$

# 分散

- 確率変数 $X$ のばらつきの程度を表す
- 分散 $V[X]$ の定義

$$V[X] = E[(X - E[X])^2]$$



# 分散

- サンプルに対する算出

- 相加平均

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

- 分散

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

分散は母集団の  
分散より常に小さい

- 不偏分散

$$\sigma'^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

バラツキができるのは  
二個以上

- 標準偏差

$\sigma, \sigma'$

- (不偏)分散の平方根

# 分散

- 期待値の線形性

$$E[A + B] = E[A] + E[B]$$

- 期待値の線形性を用いた分散  $\sigma^2$  の性質

$$\begin{aligned}\sigma^2 &= V[X] = E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2] \\ &= E[X^2] - 2E[X]E[X] + E[X]^2 \\ &= E[X^2] - E[X]^2\end{aligned}$$

# 多変量の確率分布1

## 2変数

- 同時確率密度関数
  - 二つの確率変数 $X, Y$

$$\begin{aligned} x \leq X \leq x + dx \\ y \leq Y \leq y + dy \end{aligned} \quad \begin{array}{l} \text{かつ} \\ \text{となる確率が} \end{array}$$

$$f_{XY}(x, y) dx dy \quad \text{と書ける}$$

- $X$ と $Y$ の同時確率密度関数

$$f_{XY}(x, y)$$

# 多変数の確率分布2

## n変数

- 同時確率密度関数

- n個の確率変数

$$X_1, X_2, \dots, X_n$$

- 各々  $x_i \leq X_i \leq x_i + dx_i (1 \leq i \leq n)$  となる確率が

$$f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n$$

- $X_1, X_2, \dots, X_n$  の同時確率密度関数

$$f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots, x_n)$$

# 多変数の確率分布4

## n変数

- 周辺密度関数
  - n個の確率変数

$$X_1, X_2, \dots, X_n$$

- m個のみに注目した同時確率密度

$$f_{X_1 X_2 \dots X_m}(x_1, x_2, \dots, x_m)$$

- $X_1, X_2, \dots, X_m$  についての周辺密度関数

$$f_{X_1 \dots X_m}(x_1, \dots, x_m) = \int_{-\infty}^{\infty} dx_{m+1} \cdots \int_{-\infty}^{\infty} dx_n f_{X_1 \dots X_n}(x_1, \dots, x_n)$$

# 確率変数の独立性1

- 二つの確率変数 $X, Y$
- 独立性の定義

$$\Leftrightarrow P(X \in A \text{かつ} Y \in B) = P(X \in A)P(Y \in B)$$

$$\Leftrightarrow f_{XY}(x, y) = f_X(x)f_Y(y)$$

# 確率変数の独立性2

- n個の確率変数

$$X_1, X_2, \dots, X_n$$

- 独立性の定義

$$\Leftrightarrow P(X_1 \in A_1 \text{かつ} X_2 \in A_2 \text{かつ} \dots \text{かつ} X_n \in A_n)$$

$$= P(X_1 \in A_1)P(X_2 \in A_2) \cdots P(X_n \in A_n)$$

$$\Leftrightarrow f_{X_1 X_2 \cdots X_n}(x_1, x_2, \dots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n)$$

- 自由度

– 独立に選べる確率変数の数

# 条件付き確率1

- 二つの確率変数 $X, Y$ が必ずしも独立でない

$$f_{XY}(x, y) = f(y|x)f_X(x)$$

–  $Y$ の $X$ に関する条件付き確率密度  $f(y|x)$

- $X$ が値 $x$ を取る条件下で $Y$ が値 $y$ を取る確率密度

–  $X$ と $Y$ が独立な場合  $f(y|x) = f_Y(y)$

- 同時確率密度 $f_{XY}(x, y)$ と条件付確率密度 $f(y|x)$ の関係

$$f(y|x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{f_{XY}(x, y)}{\int_{-\infty}^{\infty} f_{XY}(x, y) dy}$$



# 条件付き確率2

- 条件  $X_1 = x_1, X_2 = x_2, \dots, X_m = x_m$   
に対する,  $X_{m+1} = x_{m+1}, X_{m+2} = x_{m+2}, \dots, X_n = x_n$   
となる確率密度関数

$$f(x_{m+1}, x_{m+2}, \dots, x_n \mid x_1, x_2, \dots, x_m)$$

$$\begin{aligned} & f(x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_n) \\ &= f(x_{m+1}, x_{m+2}, \dots, x_n \mid x_1, x_2, \dots, x_m) \\ &\times f(x_1, x_2, \dots, x_m) \end{aligned}$$

# 確率変数の和の分布

- 二つの独立な確率変数 $X, Y$
- $Z=X+Y$ の分布

$X, Y$ が独立の時,  $Y=z-x$ かつ $X=x$ となる確率

– 事象  $Z = z$

$$\Leftrightarrow X + Y = z$$

$$P_X(x)P_Y(z-x)$$

$$\Leftrightarrow Y = z - x \text{ かつ } X = x$$

– 離散分布  $X \in M$

–  $Z=z$ となる確率

$$P(Z = z) = \sum_{x \in M} P_X(x)P_Y(z-x)$$