応用システム工学
第一回　確率統計の復習

平成22年05月07日

概要

- 確率変数
- 事象
- 確率分布
- 確率分布関数
- 確率密度関数
  - 一様分布
  - 正規分布
  - 標準正規分布
- 期待値
- 分散
  - 不偏分散
- 多変数の確率分布
  - 同時確率密度関数
  - 周辺密度関数
  - 確率変数の独立性
  - 条件付確率密度

確率について

- 確率変数
  - 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
  - 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)　　　Xが1となる確率:連続値・離散値
    - P(5≦X<10)　Xが5以上10未満となる確率:連続値
    - P(X∈A)　XがAに含まれる確率:離散値

確率分布関数

- 確率分布
  - 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
  - 確率変数Xが，x以下の値をとる確率

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

  - 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:$f_X(x)$
  - $x \le X \le x + dx$ となる確率 $f_X(x)dx$
  - 確率密度関数の性質

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

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

$$f_X(x) \ge 0$$

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

一様分布と確率密度関数

- 一様分布
  - 確率変数Xが区間[a,b]に一様分布
  - 確率密度関数

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

正規分布と確率密度関数

- 正規分布
  - 別名ガウス分布
  - 確率変数Xの平均μ，分散σ^2
  - 確率密度関数

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

$$X \sim N\left(\mu,\sigma^2\right)$$

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(確率変数X)
  - 確率密度関数

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

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

正規分布関数の確率密度関数1

- 正規分布の確率密度関数(確率変数X)

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

- 性質
  - 確率密度関数を確率変数の取りうる範囲で積分すると1となる(必ず事象が生じる)
    - 確認 $$\int_{-\infty}^{\infty} f_X(x)dy=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]dx$$

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

正規分布関数の確率密度関数2

- 変数変換

$$z=\frac{x-\mu}{\sigma} \Rightarrow \frac{dz}{dx}=\frac{d}{dx}\frac{x-\mu}{\sigma}=\frac{1}{\sigma}$$

$$x=\sigma z+\mu \qquad x:-\infty \leftrightarrow \infty \Leftrightarrow z=:-\infty \leftrightarrow \infty$$

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

$$=\frac{1}{\sqrt{2\pi}}\boxed{\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}z^2\right]dz} \Rightarrow \text{次ページ}$$

$$=\frac{1}{\sqrt{2\pi}}\sqrt{2\pi}=1$$

正規分布関数の確率密度関数3

- これを求める $\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz$
  - これは偶関数 $\int_{-\infty}^{0} e^{-\frac{1}{2}z^2} dz = \int_{0}^{\infty} e^{-\frac{1}{2}z^2} dz$
    - これを求めればよい $2\int_{0}^{\infty} e^{-\frac{1}{2}z^2} dz$
      - 変数変換 $x = \frac{z}{\sqrt{2}}$ $\frac{dz}{dx} = \sqrt{2}$

$$\int_{0}^{\infty} e^{-\frac{1}{2}z^2} dz = \sqrt{2}\int_{0}^{\infty} e^{-x^2} dx$$

$\int_{0}^{\infty} e^{-x^2} dx$ を求める $\int_{0}^{\infty} e^{-x^2} dx = I$ とする

正規分布関数の確率密度関数4

- 二重積分を考える

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

  - 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

$$dxdy = rdrd\theta$$

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

正規分布関数の確率密度関数5

$$\int_0^\infty\int_0^\infty e^{-\left(x^2+y^2\right)}dxdy = \int_0^{\frac{\pi}{2}}\int_0^\infty e^{-\left(r^2\cos^2\theta+r^2\sin^2\theta\right)}rdrd\theta$$

$$= \int_0^{\frac{\pi}{2}}\int_0^\infty e^{-r^2}rdrd\theta = \int_0^{\frac{\pi}{2}}\int_0^\infty re^{-r^2}drd\theta$$

$$= \int_0^{\frac{\pi}{2}}\left[\frac{-1}{2}e^{-r^2}\right]_0^\infty d\theta = \frac{-1}{2}\int_0^{\frac{\pi}{2}}\left[0-(1)\right]d\theta$$

$$= \frac{1}{2}\int_0^{\frac{\pi}{2}}d\theta = \frac{1}{2}\left[\theta\right]_0^{\frac{\pi}{2}} = \frac{\pi}{4} = I^2$$

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

期待値

- 確率変数の平均みたいなもの
- 期待値E[X]
  - 離散分布の確率変数X
    - Xの取り得る値の集合M

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

  - 連続分布の確率変数X

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

正規分布関数の期待値1

- 正規分布の確率密度関数(確率変数X)

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

- 期待値

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

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

$$=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}xe^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx$$

正規分布関数の期待値2

- 変数変換 $z = \dfrac{x-\mu}{\sigma}$ ⇨ $\dfrac{dz}{dx} = \dfrac{d}{dx}\dfrac{x-\mu}{\sigma} = \dfrac{1}{\sigma}$

$x = \sigma z + \mu \qquad x : -\infty \leftrightarrow \infty \Leftrightarrow z =: -\infty \leftrightarrow \infty$

$$E[X] = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

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

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

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

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

$$= \frac{1}{\sqrt{2\pi}}\left[\sigma 0 + \mu\sqrt{2\pi}\right]$$

$$= \mu$$

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

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

分散

- サンプルに対する算出
  - 相加平均 $$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$
  - 分散 $$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2$$
    分散は母集団の
    分散より常に小さい
  - 不偏分散 $$\sigma'^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2$$
    バラツキがでるのは
    二個以上
  - 標準偏差 σ, σ'
    - (不偏)分散の平方根

期待値と分散

- 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σ^2の性質

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