応用システム工学
第二回　確率統計の復習

平成26年4月18日

E6-111

分布を表す量
モーメント

- モーメント(積率) → 分散の発展形
- 変数Xの度数分布を考える
  - 階級値xの階級の相対度数f(x)
    - 平均$\mu = E(X) = \sum_x xf(x)$
    - 分散$\sigma^2 = V(X) = E\{(x-\mu)^2\} = \sum_x (x-\mu)^2 f(x)$
    - 一般形 変数Xの関数g(x)の平均

$$E\left(g(x)\right) = \sum_x g(x)f(x)$$

  - 関数g(x)がk次 → 変数Xのk次のモーメント

分布を表す量
モーメント

- 原点の周りのk次のモーメント　$\mu'_k = E\left(X^k\right)$
  - 平均μは原点の周りの1次のモーメントに相当
- 平均の周りのk次のモーメント　$\mu_k = E\left((X-\mu)^k\right)$
  - 分散σ^2は平均の周りの2次のモーメントに相当
- 高次のモーメントを用いた分布の特徴量
  - 3次のモーメント　歪度　$\alpha_3 = \frac{\mu_3}{\sigma_3}$
    - ヒストグラムの正または負への偏り具合
  - 4次のモーメント　尖度　$\alpha_4 - 3 = \frac{\mu_4}{\sigma_4} - 3$
    - 正規分布の尖度を0とする（正規分布の値が3）
    - 平均値付近での分布の集中の度合い

標本の取り出し方

- 復元抽出
  - 取り出した標本を母集団に戻してから，次の標本を取り出す
    - 母集団より取り出される確率は，データによらず同じ
    - データが取り出される確率は，他に選ばれたデータの影響を受けない
- 非復元抽出
  - 取り出した標本を母集団に戻さずに，次の標本を取り出す
    - 母集団の数が小さい場合は，補正が必要

確率分布モデル

- 確率分布モデル
  - 母集団の分布を表す関数
  - 連続数値⇒ヒストグラム⇒区間(階級)
- 離散型確率分布
  - 確率変数が飛び飛びの値をとる
    - 人数, 回数等
- 連続型確率分布
  - 確率変数が連続値
    - 身長, 体重等

- ある範囲の値をとる確率
  - ヒストグラムの階級の面積
  - ヒストグラムの極限の上端部
    ⇒確率密度関数

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

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

    - 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

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

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

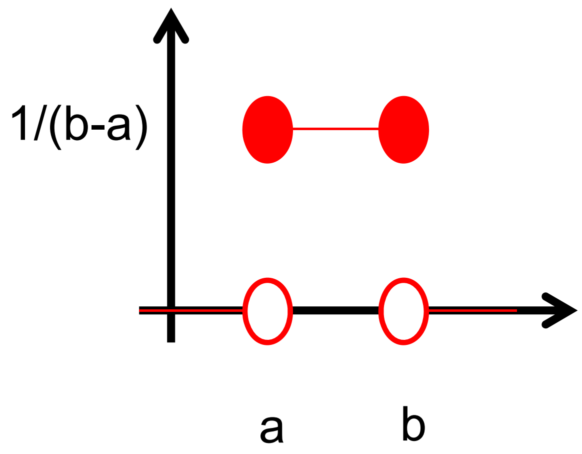
$$f_X(x) \ge 0$$

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

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

- 中心極限定理
  - 無数の独立した原因によるデータの分布の合計は正規分布となる
  - 母集団分布が正規分布ならば，標本平均も正規分布となる
- 正規分布の性質
  - 確率変数Xが期待値μ，分散σ^2の正規分布$N(\mu, \sigma^2)$に従うとき，確率変数$(X - \mu\ )/\sigma$は正規分布$N(0,1)$に従う
  - $X_1, \cdots, X_n$が独立で，各々正規分布$N(\mu, \sigma^2)$にしたがうとき，平均$(X_1 + \cdots + X_n)/n$は$N(\mu, \sigma^2/n)$に従う

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

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

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

正規分布関数の確率密度関数2

- 変数変換

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

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

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

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

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

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

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

$$= \int_0^{\frac{\pi}{2}} \int_0^\infty e^{-r^2} r dr d\theta = \int_0^{\frac{\pi}{2}} \int_0^\infty r e^{-r^2} dr d\theta$$

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

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

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