

応用システム工学

第二回 確率統計の復習

平成25年04月26日

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
 - $P(5 \leq X < 10)$ X が5以上10未満となる確率:連続値
 - $P(X \in A)$ 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: $f_X(x)$
 - $x \leq X \leq x + dx$ となる確率 $f_X(x)dx$
 - 確率密度関数の性質

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

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

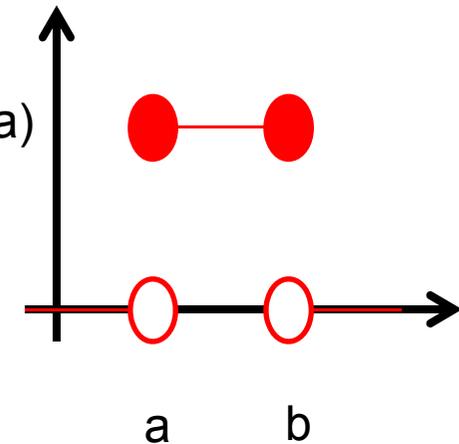
$$f_X(x) \geq 0$$

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

一様分布と確率密度関数

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

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



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

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

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

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(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, \dots, 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 $(X_1 + \dots + X_n) / n$ 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)

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

- 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

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

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

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

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

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

正規分布関数の確率密度関数3

- これを求める
 - これは偶関数 $\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz$
 $\int_{-\infty}^0 e^{-\frac{1}{2}z^2} dz = \int_0^{\infty} e^{-\frac{1}{2}z^2} dz$
 - これを求めればよい $2 \int_0^{\infty} e^{-\frac{1}{2}z^2} dz$
 - 変数変換 $x = \frac{z}{\sqrt{2}} \quad \frac{dz}{dx} = \sqrt{2}$
- $$\int_0^{\infty} e^{-\frac{1}{2}z^2} dz = \sqrt{2} \int_0^{\infty} e^{-x^2} dx$$
- $\int_0^{\infty} e^{-x^2} dx$ 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

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

– 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 $E[X]$

– 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

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

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)

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

- 期待値

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

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

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

正規分布関数の期待値2

- 変数変換

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

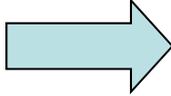
$$x = \sigma z + \mu \quad x: -\infty \leftrightarrow \infty \Leftrightarrow z: -\infty \leftrightarrow \infty$$

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

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

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

$$= \mu$$

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

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