統計学ノートNo. 5
連続確率分布

担当教員：蛭川雅之

研究室：紫英館3階345研究室

メール： hirukawa@econ.ryukoku.ac.jp

オフィスアワー：毎週金曜日3講時

1　連続確率変数

連続確率変数の性質

- **連続確率変数はある区間（例：$[0,1], (-\infty, \infty)$）内の任意の実数をとる。**
- **連続確率変数の確率は区間に対して定義される。**
  - **連続確率変数が特定の一点をとる確率はゼロである。**

分布関数

- **連続確率変数Xに対して関数$F_X(x)$を**

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

と定義する。

- X**が区間**$[a, b]$ $(a \leq b)$**に入る確率は次のように表現される。**

$$Pr(a \leq X \leq b)$$
$$= Pr(a < X \leq b)$$
$$= Pr(X \leq b) - Pr(X \leq a)$$
$$= F_X(b) - F_X(a)$$

分布関数（つづき）

- $F_X(x)$は連続確率変数Xの（累積）分布関数と呼ばれる。

- 分布関数は以下の性質を持つ。

1. $0 \leq F_X(x) \leq 1$
2. $F_X(-\infty) = 0, F_X(\infty) = 1$
3. $x < x^*$を満たすに対し、$F_X(x) \leq F_X(x^*)$

確率密度関数

- **分布関数$F_X(x)$の微小な増加分**

$$f_X(x) = \frac{d}{dx}F_X(x) = \lim_{dx \to 0} \frac{F_X(x + dx) - F_X(x)}{dx}$$

を確率密度関数という。

- **確率密度関数は以下の性質を持つ。**

1. $f_X(x) \geq 0$
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$

分布関数と確率密度関数との対応

- **確率密度関数の定義から**

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du$$

が成り立つ。

- **ところで、**

$$Pr(a \leq X \leq b) = F_X(b) - F_X(a)$$

であった。

分布関数と確率密度関数との対応（つづき）

- **従って、**

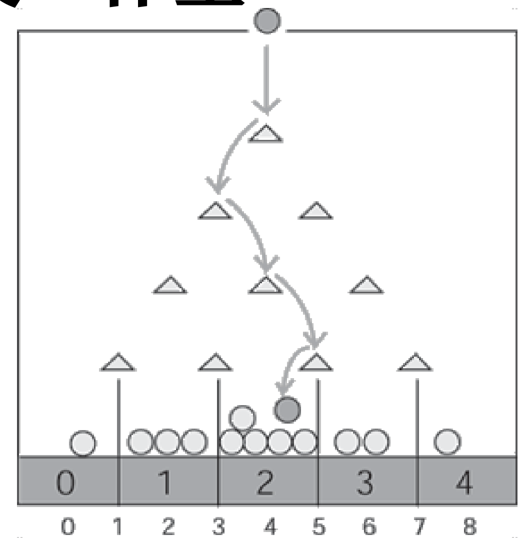
$$Pr(a \leq X \leq b)$$
$$= \int_{-\infty}^{b} f_X(x)dx - \int_{-\infty}^{a} f_X(x)dx$$

が成り立つ。

2　正規分布

正規分布とは…

- **正規分布（ガウス分布）は理論的にも実用的にも最も重要（あるいは代表的）な分布である。**
- **どの程度「代表的」か？**
  1. **人間・動物の身長・体重**
  2. **試験の得点**
  3. **株価収益率**
  4. Galton Board → →



正規分布の確率密度関数

- **確率変数Xが平均μと分散σ^2の正規分布に従うとき、Xの確率密度関数は**

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

で与えられる。

正規分布の確率密度関数（つづき）

- **確率変数Xが平均μと分散σ^2の正規分布に従うことを**

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

と表現する。

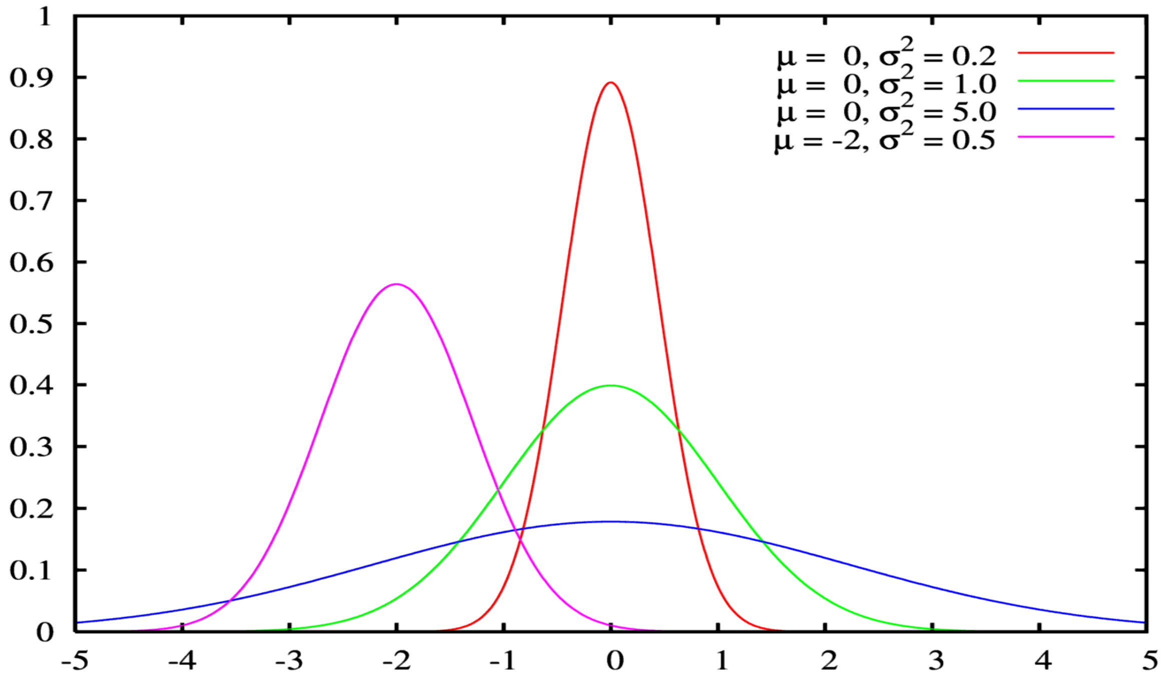
- **当然ながら、**

$$E(X) = \mu$$
$$Var(X) = \sigma^2$$

である。

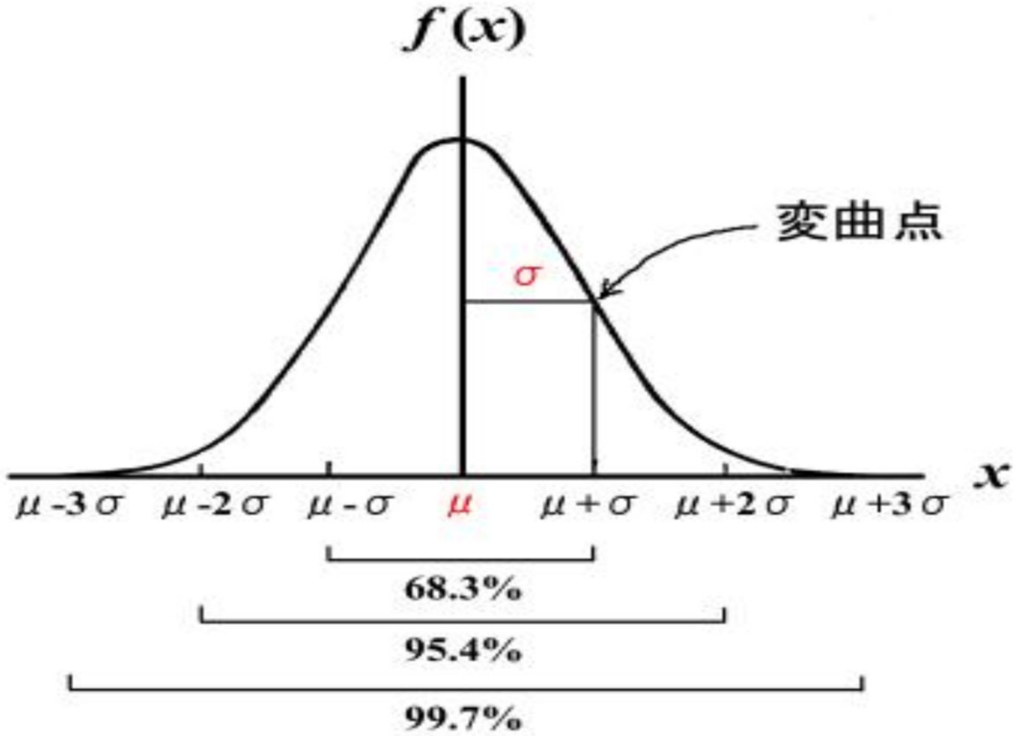
正規分布の形状

- **正規分布は、平均μ と分散σ^2の２つの数値を与えると一種類に決まる。**



正規分布の形状（つづき）

- **平均μに関して****左右対称の釣鐘型****をしている。**



標準正規分布

- **平均$\mu = 0$と分散$\sigma^2 = 1$の正規分布$N(0,1)$を標準正規分布という。**
  - **しばしば、標準正規分布に従う確率変数をZ、さらに、Zの確率密度関数および分布関数を以下のように表記する。**

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

$$\Phi(z) = Pr(Z \leq z) = \int_{-\infty}^{z} \phi(u)du$$

標準正規分布（つづき）

- $Z \sim N(0,1)$**に対し、以下が成り立つ。**

$$Pr(-1.65 \leq Z \leq 1.65) \approx 0.90$$

$$Pr(-1.96 \leq Z \leq 1.96) \approx 0.95$$

$$Pr(-2.58 \leq Z \leq 2.58) \approx 0.99$$

- **これらの事実を記憶しておくとよい。**

正規分布の確率計算

- $Z \sim N(0,1)$ であるとき、Z が区間 $[a,b]$ $(a \leq b)$に入る確率

$$Pr(a \leq Z \leq b) = \Phi(b) - \Phi(a)$$

を求めるには、正規分布表を用いる。

  - Excelでは、関数`normsdist`が分布関数 $\Phi(z)$の数値を与える。

正規分布の確率計算（つづき）

- **一般の正規分布の場合はどうするか？**
- $X \sim N(\mu, \sigma^2)$**を標準化すると**

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

となることを利用する。

正規分布の確率計算（つづき）

- **具体的には、**$X \sim N(\mu, \sigma^2)$**であるとき、**X**が区間**$[a, b]$ $(a \leq b)$**に入る確率は**

$$Pr(a \leq X \leq b) = Pr\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right)$$

$$= Pr\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

となることから、この右辺に正規分布表（またはnormsdist**）を適用する。**

正規分布の確率計算：例題

$X \sim N(10,16)$であるとき、次の確率を求めよ。

(1) X**が13以下である確率**

(2) X**が16を上回る確率**

(3) X**が7以上11以下である確率**

正規分布の確率計算：例題（つづき）

(1)

$$Pr(X \leq 13)$$

$$= Pr\left(\frac{X-10}{4} \leq \frac{13-10}{4}\right)$$

$$= Pr(Z \leq 0.75)$$

$$= \Phi(0.75)$$

$$= 0.7734$$

正規分布の確率計算：例題（つづき）

(2)

$$
\begin{aligned}
&Pr(X > 16) \\
&= 1 - Pr(X \leq 16) \\
&= 1 - Pr\left(\frac{X - 10}{4} \leq \frac{16 - 10}{4}\right) \\
&= 1 - Pr(Z \leq 1.50) \\
&= 1 - \Phi(1.50) \\
&= 1 - 0.9332 \\
&= 0.0668
\end{aligned}
$$

正規分布の確率計算：例題（つづき）

(3)

$$
\begin{aligned}
&Pr(7 \leq X \leq 11) \\
&= Pr\left(\frac{7-10}{4} \leq \frac{X-10}{4} \leq \frac{11-10}{4}\right) \\
&= Pr(-0.75 \leq Z \leq 0.25) \\
&= Pr(Z \leq 0.25) - Pr(Z \leq -0.75) \\
&= \Phi(0.25) - \Phi(-0.75) \\
&= \Phi(0.25) - \{1 - \Phi(0.75)\} \\
&= 0.5987 - (1 - 0.7734) \\
&= 0.3721
\end{aligned}
$$

正規分布の分位点

- しばしば、$X \sim N(\mu, \sigma^2)$および0以上1以下の定数qに対し、

$$Pr(X \leq c) = q$$

を満たす点cを求める必要が生じる。

  - cは$N(\mu, \sigma^2)$の**100q%分位点**と呼ばれる。

正規分布の分位点（つづき）

- **cの具体的な求め方は以下の通りである。**

1. **まず、**

$$q = Pr(X \leq c)$$
$$= Pr\left(\frac{X-\mu}{\sigma} \leq \frac{c-\mu}{\sigma}\right)$$
$$= Pr(Z \leq c')$$

を満たす$N(0,1)$の$100q\%$分位点c'を求める。

   - **c'を求める際、正規分布表（またはExcelの関数normsinv）を利用する。**

正規分布の分位点（つづき）

2. **次に、c'とcとの関係式**

$$\frac{c-\mu}{\sigma}=c'$$

をcについて解き、$N(\mu,\sigma^2)$の$100q\%$分位点

$$c=\mu+\sigma c'$$

を得る。

正規分布の分位点：例題

ある試験の平均点は55点、分散は196であった。試験の得点が正規分布に従うと仮定して、得点の上位10%に入るには何点以上が必要であったかを求めよ。

正規分布の分位点：例題（つづき）

- **試験の得点をXと表記すると、$X \sim N(55, 14^2)$である。**
- **上位10%に入るために必要な点数がc点である場合、**

$$Pr(X > c) = 0.1$$
$$\Rightarrow Pr(X \leq c) = 0.9$$

が成り立つ。

正規分布の分位点：例題（つづき）

- ここで、

$$0.9 = Pr\left(\frac{X-55}{14} \leq \frac{c-55}{14}\right) = Pr(Z \leq c')$$

とおき、正規分布表を利用すると

$$c' = 1.28$$

が得られる。

正規分布の分位点：例題（つづき）

- さらに、

$$c' = \frac{c - 55}{14} = 1.28$$

であるから、求める得点は

$$c = 55 + 14 \times 1.28 = 72.92$$

である。