応用計量経済分析 TA セッション
第 7 回：正規母集団に関する分布

TA：北村友宏 *

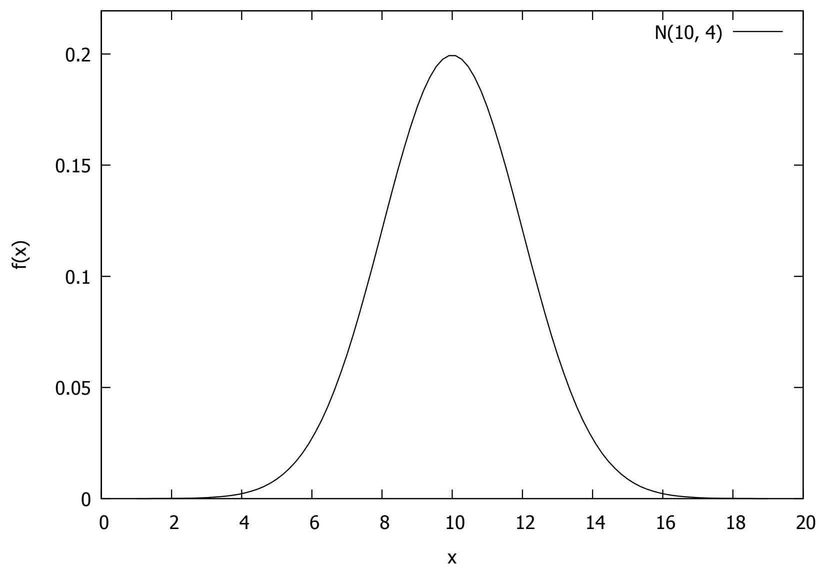
2015 年 11 月 24 日

1　正規分布

- 期待値 μ、分散 σ^2 の**正規分布**（normal distribution）に従う確率変数 X の確率密度関数は、

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

  - ⋆ 指数関数 exp[.] の中に x を標準化（期待値 μ を引き標準偏差 σ で割る変換）したものが入り、それに $\dfrac{1}{\sqrt{2\pi}\sigma}$ を掛けて全積分が 1 になるように調整している。
  - ⋆ $f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right]$ と書いてもよい。
  - ⋆ 連続型の確率分布。
  - ⋆ 確率密度関数 $f(x)$ の形状は左右対称の山型。
  - ⋆ X が期待値 μ、分散 σ^2 の正規分布に従うことを、$X \sim N\left(\mu, \sigma^2\right)$ と表す。
  - ⋆ $X \sim N(10, 4)$ の確率密度関数のグラフは、



* Email: kitamu.tom@gmail.com　URL: http://tomkitamura.html.xdomain.jp

★ 確率変数 $X \sim N\left(\mu, \sigma^2\right)$ と定数 a, b について、

$$aX + b \sim N\left(\underbrace{a\mu + b}_{\text{期待値}}, \underbrace{a^2\sigma^2}_{\text{分散}}\right)$$

が成り立つ（証明略）。

★ **互いに独立な**確率変数 $X \sim N\left(\mu_X, \sigma_X^2\right), Y \sim N\left(\mu_Y, \sigma_Y^2\right)$ について、

$$X + Y \sim N\left(\underbrace{\mu_X + \mu_Y}_{\text{期待値}}, \underbrace{\sigma_X^2 + \sigma_Y^2}_{\text{分散}}\right)$$

が成り立つ（証明略）。これを**正規分布の再生性**とよぶ。

★ 母集団分布を $N\left(\mu, \sigma^2\right)$ とし、そこから抽出した標本 $(X_1, X_2, \ldots, X_n)$ の分布は、

$$\forall i, X_i \sim N\left(\mu, \sigma^2\right)$$

となる。

★ 正規分布に従う母集団を**正規母集団**とよぶ。

例題 1. 任意の i について $X_i \sim N\left(\mu, \sigma^2\right)$ であり、X_i は互いに独立であるとする。このとき、$X_1 + X_2 + X_3$ の分布を求めなさい。

任意の i について $X_i \sim N\left(\mu, \sigma^2\right)$ なので、

$$E(X_1) = E(X_2) = E(X_3) = \mu,$$
$$V(X_1) = V(X_2) = V(X_3) = \sigma^2$$

である。X_1, X_2, X_3 は互いに独立なので、正規分布の再生性より、

$$X_1 + X_2 + X_3 \sim N\left(\mu + \mu + \mu, \sigma^2 + \sigma^2 + \sigma^2\right),$$

つまり、

$$X_1 + X_2 + X_3 \sim N\left(3\mu, 3\sigma^2\right)$$

である。

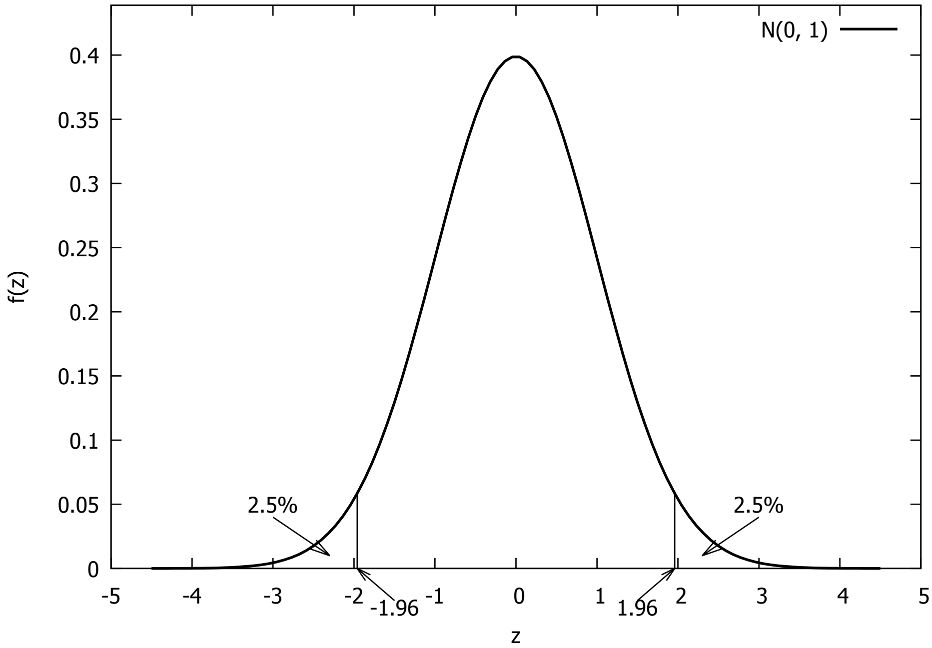
- 期待値 0、分散 1（標準偏差も 1）の正規分布を**標準正規分布**（standard normal distribution）とよぶ。

★ $Z \sim N(0, 1)$ の確率密度関数は、

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right).$$

※ $(*)$ に $\mu = 0, \sigma = 1$ を代入すれば得られる。

★ $Z \sim N(0, 1)$ の絶対値 $|Z|$ が 1.96 を超える確率 $P(|Z| > 1.96)$ はわずか 5%（仮説検定や区間推定の際に重要）。



※連続型の確率分布なので、確率密度関数の下部の面積が確率を表す。

⋆ 母分散が**既知**のとき、$N\left(\mu, \sigma^2\right)$ の母集団から無作為抽出した大きさ n の標本 $(X_1, X_2, \ldots, X_n)$ について、標本平均を標準化したものの分布は

$$\frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0, 1)$$

となる（証明は練習問題参照）。

例題 2. $X \sim N\left(\mu, \sigma^2\right)$ とし、X を標準化したものを Z とする。つまり、$Z = \dfrac{X - \mu}{\sigma}$ である。このとき、Z が標準正規分布に従うことを証明しなさい。

（証明）

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma} \cdot X - \frac{\mu}{\sigma}.$$

$X \sim N\left(\mu, \sigma^2\right)$ なので、$E(X) = \mu, V(X) = \sigma^2$ である。よって、

$$Z = \frac{1}{\sigma} \cdot X - \frac{\mu}{\sigma} \sim N\left(\frac{1}{\sigma} \cdot \mu - \frac{\mu}{\sigma}, \left(\frac{1}{\sigma}\right)^2 \cdot \sigma^2\right),$$

つまり、

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

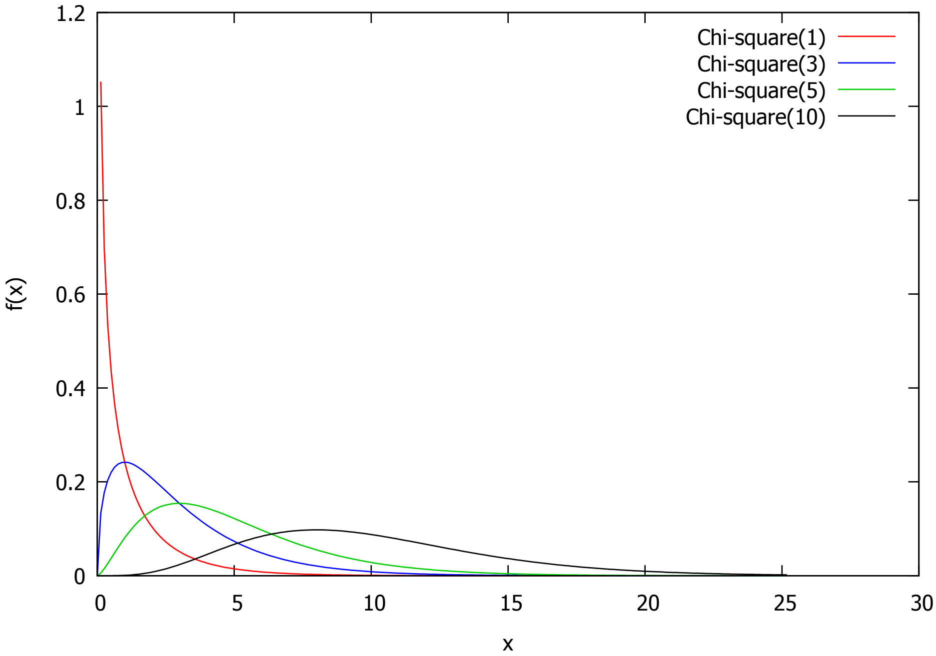
である。したがって、Z は標準正規分布に従う。（証明終）

2 カイ 2 乗分布

- k 個の確率変数 $Z_1, Z_2, \ldots, Z_k$ が全て $N(0, 1)$ に従い、**互いに独立**であるとき、$Z_1^2 + Z_2^2 + \cdots + Z_k^2$ は自由度 k の**カイ 2 乗分布**（chi-square distribution）に従う。

★ $Z_1^2 + Z_2^2 + \cdots + Z_k^2 \sim \chi^2(k)$ と書く。

★ 自由度によって確率密度関数の形状が異なる。



★ 母平均が未知のとき、分散 σ^2 の正規母集団から無作為抽出した大きさ n の標本 $(X_1, X_2, \ldots, X_n)$ について、

$$\sum_{i=1}^{n} \left(\frac{X_i - \overline{X}}{\sigma} \right)^2 \sim \chi^2(n-1)$$

が成り立つ（証明略）。自由度は $n-1$ であることに注意！

※ $s^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2$ とすれば $\sum_{i=1}^{n} \left(\dfrac{X_i - \overline{X}}{\sigma} \right)^2 = (n-1)\dfrac{s^2}{\sigma^2}$ となる（証明略）ので、

$(n-1)\dfrac{s^2}{\sigma^2} \sim \chi^2(n-1)$ と書いてもよい。

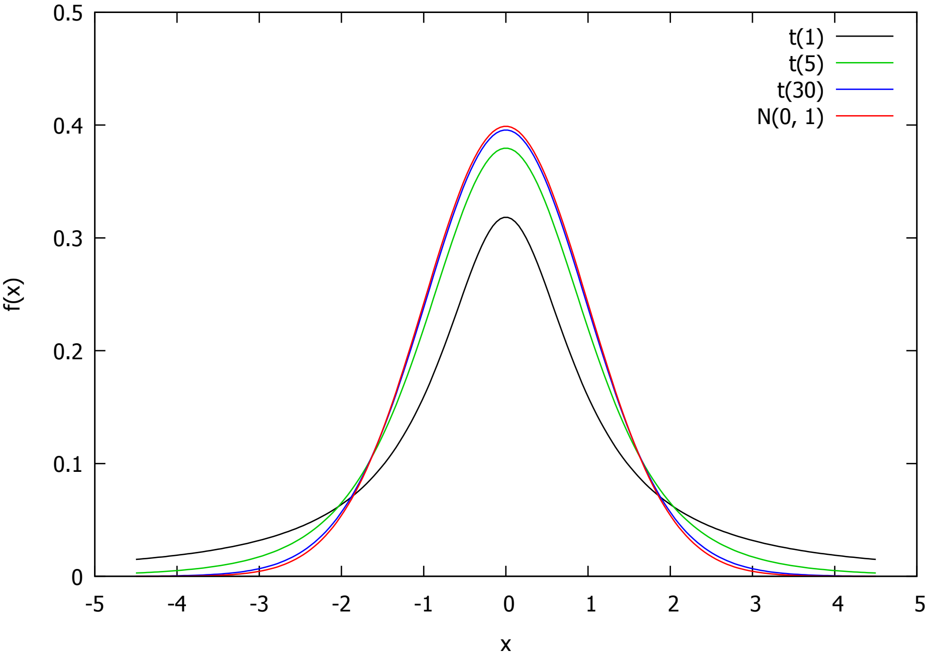
3 *t* 分布

- 確率変数 $Z \sim N(0,1)$ と $W \sim \chi^2(k)$ が独立のとき、$\dfrac{Z}{\sqrt{W/k}}$ は自由度 k の t 分布（t-distribution）に従う。

  ★ $\dfrac{Z}{\sqrt{W/k}} \sim t(k)$ と書く。

  ★ 自由度 1 の t 分布 $t(1)$ をコーシー分布（Cauchy distribution）とよぶ。

  ★ 確率密度関数の形状は左右対称の山型で、自由度によって形状が異なる。



★ 自由度が大きくなるにつれ、分布の裾が薄くなり、形状が標準正規分布 $N(0,1)$ に近づく。$t(30)$ は形状がほぼ $N(0,1)$ と同じで、$t(\infty)$ は標準正規分布 $N(0,1)$ となる。

★ 母分散が**未知**のとき、$N\left(\mu,\sigma^2\right)$ の母集団から無作為抽出した大きさ n の標本 $(X_1,X_2,\ldots,X_n)$ について、$s^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ とすると、

$$\frac{\overline{X}-\mu}{\sqrt{s^2/n}} \sim t(n-1)$$

が成り立つ（証明略）。分母の平方根の中は $\dfrac{s^2}{n}$ であり、自由度は $n-1$ であることに注意！