第 2 回：確率変数と確率分布
【教科書第 3 章】

北村　友宏

2025 年 10 月 7 日

本日の内容

1. 確率変数と確率分布
2. 確率変数の期待値・分散・標準偏差
3. 同時・周辺確率質量関数
4. 条件付き期待値と条件付き分散
5. 連続型確率分布
6. 計量経済学で使う代表的な確率分布

確率変数

- 試行の結果によって値がランダムに変わり，それぞれの値が実現する確率が割り振られている変数を確率変数（random variable）という．
  - e.g., コイントスの結果，サイコロの目，天気
- 確率変数がランダムに変化した結果，実際に出現した値を実現値（realization）という．
- 統計学では確率変数を大文字で，実現値を小文字で表す．
  - e.g., 「サイコロの目」という確率変数を X とする．出た目を一般化して x として，「サイコロを振ると x が出た」を数式で $X = x$ と表す．
- 各実現値とそれが出る確率の関係を確率分布（probability distribution）という．

- X をコイントスの結果とするならば，表を 1，裏を 0 とすると，

$$X = \begin{cases} 0 & \text{with pr. } \frac{1}{2}, \\ 1 & \text{with pr. } \frac{1}{2}. \end{cases}$$

- X をサイコロの目とするならば，

$$X = \begin{cases} 1 & \text{with pr. } \frac{1}{6}, \\ 2 & \text{with pr. } \frac{1}{6}, \\ \vdots & \\ 6 & \text{with pr. } \frac{1}{6}. \end{cases}$$

離散型確率変数

- 非連続的な値をとる確率変数を離散型確率変数（random variable of discrete type）といい，その確率分布を離散型確率分布（probability distribution of discrete type）という．
  - e.g., サイコロを振って出る目
  - サイコロを振ると，1, 2, 3, 4, 5, 6 のみが出る．0.5 や 1.5 などは出ない．

連続型確率変数

- 連続的な値をとる確率変数を連続型確率変数（random variable of continuous type）といい，その確率分布を連続型確率分布（probability distribution of continuous type）という．
  - e.g., ルーレットを回して針が止まる位置
  - ルーレットの円周上に，0 から 1 までの数値を定義すると，円周上での位置を表す点が無限個現れる．
  - 針が止まる可能性のある位置は、無限個考えられる．
  - ある特定の点に針がピッタリ止まる確率は 0 とみなせる．
  - 「何%の確率で，針の位置がどこからどこまでの間に止まる」のような考え方をするのが適切である．例えば，「ルーレットを回して針が 0.2 の点から 0.8 の点までの間に止まる確率は？」などを考える．

確率質量関数

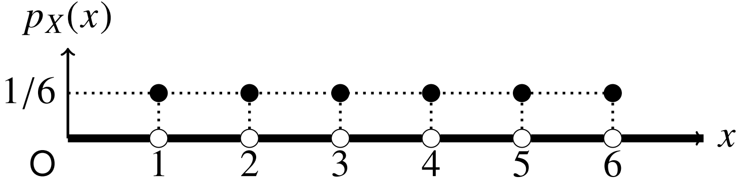
- 離散型確率変数における各実現値と，それが出る確率の関係を表す関数を確率質量関数（probability mass function, pmf）という．
  - 確率関数ともいう．
- $X = x$ となる確率を $\Pr(X = x)$ と書けば，確率質量関数（の書き方の一例）は

$$p_X(x) = \Pr(X = x).$$

  - とりうる値を $x_i (i = 1, 2, \cdots, n)$ として，確率関数を $p_X(x_i) = \Pr(X = x_i)$ などと書くこともある．
- サイコロの目の確率質量関数は，

$$p_X(x) = \begin{cases} \frac{1}{6} & \text{for } x = 1, 2, \ldots, 6, \\ 0 & \text{elsewhere.} \end{cases}$$

サイコロの目の確率質量関数をグラフで表すと,



- 離散型確率変数では，確率質量関数のグラフの高さが確率を表す.

確率質量関数の性質：

1. $0 \le p_X(x) \le 1.$
2. $\displaystyle\sum_x p_X(x) = 1.$

累積分布関数

- 実現値が小さい順に確率を積み重ねた関数を累積分布関数（cumulative distribution function, cdf）という．
  - 分布関数ともいう．
  - e.g., サイコロを振り、3 以下が出る確率 $F_X(3) = \Pr(X \leq 3)$ は，1, 2, 3 が出る確率を足し合わせたもの．つまり，

$$
\begin{aligned}
F_X(3) &= \Pr(X \leq 3) \\
&= \Pr(X = 1) + \Pr(X = 2) + \Pr(X = 3) \\
&= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\
&= \frac{3}{6} \\
&= \frac{1}{2}.
\end{aligned}
$$

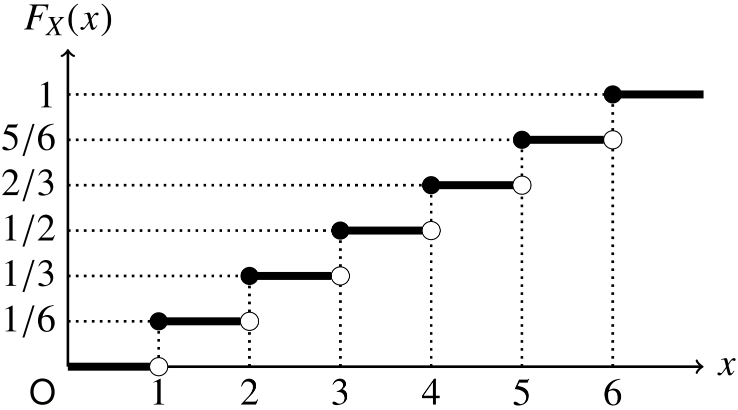
サイコロの目を X とすると，

$$
\begin{aligned}
F_X(1) &= \Pr(X \le 1) = \frac{1}{6}, \\
F_X(2) &= \Pr(X \le 2) = \frac{2}{6} = \frac{1}{3}, \\
F_X(3) &= \Pr(X \le 3) = \frac{3}{6} = \frac{1}{2}, \\
F_X(4) &= \Pr(X \le 4) = \frac{4}{6} = \frac{2}{3}, \\
F_X(5) &= \Pr(X \le 5) = \frac{5}{6}, \\
F_X(6) &= \Pr(X \le 6) = \frac{6}{6} = 1.
\end{aligned}
$$

よって，サイコロの目の累積分布関数は，

$$F_X(x) = \begin{cases} 0 & \text{for } x < 1, \\ \frac{1}{6} & \text{for } 1 \le x < 2, \\ \frac{1}{3} & \text{for } 2 \le x < 3, \\ \frac{1}{2} & \text{for } 3 \le x < 4, \\ \frac{2}{3} & \text{for } 4 \le x < 5, \\ \frac{5}{6} & \text{for } 5 \le x < 6, \\ 1 & \text{for } 6 \le x. \end{cases}$$

サイコロの目の累積分布関数をグラフで表すと，



- 離散型確率変数では，累積分布関数のグラフの高さが確率を表す．

離散型確率変数の累積分布関数の性質：

1. $0 \leq F_X(x) \leq 1$.
2. 単調非減少関数である．すなわち，

$$x_1 < x_2 \Rightarrow F_X(x_1) \leq F_X(x_2).$$

3. 階段状である．

離散型確率変数の期待値

- $(x_1, x_2, \cdots, x_M)$ の実現値をとりうる離散型確率変数 X の確率質量関数を $p_X(x)$ とすると，離散型確率変数 X の期待値（expectation）は，

$$
\begin{aligned}
E(X) &= x_1 p_X(x_1) + x_2 p_X(x_2) + \cdots + x_M p_X(x_M) \\
&= \sum_{i=1}^{M} x_i p_X(x_i).
\end{aligned}
$$

  - 確率変数がランダムに変化する結果，実現値が平均的にどの程度の値になるかを表す．

- 確率変数の関数の期待値は，

$$
E(g(X)) = \sum_{i=1}^{M} g(x_i) p_X(x_i).
$$

離散型確率変数の分散

- $(x_1, x_2, \cdots, x_M)$ の実現値をとりうる離散型確率変数 X の確率質量関数を $p_X(x)$ とすると，離散型確率変数 X の分散（variance）は，

$$V(X) = E\left((X - E(X))^2\right) = \sum_{i=1}^{M}(x_i - E(X))^2 p_X(x_i).$$

※ 確率は二乗しない．

※ 前回出てきた，観測値の分散とは別物である．

- 確率変数がランダムに変化する結果，実現値がどの程度ばらつくかを表す．
- 0 以上の値をとる．
- 分散が大きいほど実現値のばらつきが大きく，分散が小さいほど実現値のばらつきが小さい．
- 単位も二乗される．

確率変数の標準偏差

- 確率変数の分散の正の平方根を標準偏差（standard deviation）という．

$$D(X) = \sqrt{V(X)}.$$

※ 前回出てきた，観測値の標準偏差とは別物である．
- 確率変数がランダムに変化する結果，実現値がどの程度ばらつくかを表す．
- 0 以上の値をとる．
- 標準偏差が大きいほど実現値のばらつきが大きく，標準偏差が小さいほど実現値のばらつきが小さい．
- 単位は，元の単位に戻る．

同時確率質量関数

- 離散型確率変数 X と Y の同時確率質量関数（joint probability mass function）は，

$$p_{X,Y}(x, y) = \Pr(X = x, Y = y).$$

  - 同時確率関数，結合確率関数，結合確率質量関数ともいう．

周辺確率質量関数

- X の周辺確率質量関数（marginal probability mass function）は，

$$p_X(x) = \Pr(X = x) = \sum_y p_{X,Y}(x, y).$$

また，Y の周辺確率質量関数は，

$$p_Y(y) = \Pr(Y = y) = \sum_x p_{X,Y}(x, y).$$

  - 周辺確率関数ともいう．
  - $X = x$ となる確率は，$p_{X,Y}(x, y)$ を Y のとりうるすべての実現値について足し合わせたもの．
  - $Y = y$ となる確率は，$p_{X,Y}(x, y)$ を X のとりうるすべての実現値について足し合わせたもの．

確率変数の共分散

- 離散型確率変数 X と Y の共分散（covariance）は，

$$
\begin{aligned}
&\mathrm{Cov}(X, Y)\\
=&E((X - E(X))(Y - E(Y)))\\
=&\sum_x \sum_y (x - E(X))(y - E(Y)) p_{X,Y}(x, y).
\end{aligned}
$$

  - 符号の向きが，2 つの確率変数の相関の方向を表す．
  - $\mathrm{Cov}(X, Y) > 0 \Leftrightarrow$ 正の相関関係
  - $\mathrm{Cov}(X, Y) < 0 \Leftrightarrow$ 負の相関関係
  - $\mathrm{Cov}(X, Y) = 0 \Leftrightarrow$ 無相関
  - 相関関係の強さは表さない．

期待値の性質

a と b を定数とし，X と Y を確率変数とする．

1. $E(a) = a$.
2. $E(aX + b) = aE(X) + b$.
3. $E(aX + bY) = aE(X) + bE(Y)$.
4. $\mathrm{Cov}(X, Y) = 0$ ならば，$E(XY) = E(X)E(Y)$.

分散の性質

a と b を定数とし，X と Y を確率変数とする．

1. $V(a) = 0.$
2. $V(aX + b) = a^2 V(X).$
3. $V(aX + bY) = a^2 V(X) + b^2 V(Y) + 2ab\mathrm{Cov}(X, Y).$
4. $\mathrm{Cov}(X, Y) = 0$ ならば，

$$V(aX + bY) = a^2 V(X) + b^2 V(Y).$$

条件付き確率質量関数

- $Y = y$ が与えられたときの離散型確率変数 X の条件付き確率質量関数（conditional probability mass function）は，

$$p_{X|Y}(x \mid Y = y) = \frac{p_{X,Y}(x, y)}{p_Y(y)}.$$

確率変数の独立性

- 2 つの離散型確率変数 X, Y がとりうる任意の実現値 x, y について,

$$p_{X,Y}(x, y) = p_X(x)p_Y(y)$$

ならば,「X と Y は互いに独立（independent）である」という.
  - 任意の実現値 x, y について,

$$p_{X|Y}(x \mid Y = y) = p_X(x)$$

ならば,「X と Y は互いに独立である」ということもできる.
  - X と Y が互いに独立ならば, $\mathrm{Cov}(X, Y) = 0$（X と Y は無相関）である.
  - $\mathrm{Cov}(X, Y) = 0$ であっても, X と Y が互いに独立であるとは限らない.

条件付き期待値

- $Y = y$ が与えられたときの離散型確率変数 X の条件付き期待値（conditional expectation）は，

$$E(X \mid Y = y) = \sum_{x} x p_{X|Y}(x \mid Y = y).$$

  - $E(X \mid Y)$ と書いてもよい．

条件付き分散

- $Y = y$ が与えられたときの離散型確率変数 X の条件付き分散（conditional variance）は，

$$V(X \mid Y = y) = \sum_x (x - E(X \mid Y = y))^2 p_{X|Y}(x \mid Y = y).$$

  - $V(X \mid Y)$ と書いてもよい．

連続型確率変数の累積分布関数

- 連続型確率変数を X とすると，連続型確率変数の累積分布関数（cumulative distribution function, cdf）は，

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

- 連続型確率変数 X が a を超えて b 以下の値をとる確率は，

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

- 連続型確率変数においては，ある 1 点の値をとる確率は 0 とみなせる．つまり，

$$\Pr(X = a) = \Pr(X = b) = 0.$$

- よって，

$$\begin{aligned}
\Pr(a < X \leq b) &= \Pr(a < X < b) + \underbrace{\Pr(X = b)}_{=0} \\
&= \Pr(a < X < b) \\
&= \underbrace{\Pr(X = a)}_{=0} + \Pr(a < X < b) \\
&= \Pr(a \leq X < b) \\
&= \Pr(a \leq X < b) + \Pr(X = b) \\
&= \Pr(a \leq X \leq b).
\end{aligned}$$

- つまり,

$$\begin{aligned}\Pr(a < X \le b) &= \Pr(a < X < b) \\ &= \Pr(a \le X < b) \\ &= \Pr(a \le X \le b).\end{aligned}$$

$$\Downarrow$$

連続型確率変数においては,「以上」と「超える」,「以下」と「未満」を区別しなくてよい.

$$\Downarrow$$

連続型確率変数 X の累積分布関数を,

$$F_X(x) = \Pr(X < x),$$

と定義してもよい.

連続型確率変数の累積分布関数の性質：

1. $0 \leq F_X(x) \leq 1$.
2. 単調非減少関数である．すなわち，

$$x_1 < x_2 \Rightarrow F_X(x_1) \leq F_X(x_2).$$

確率密度関数

- 連続型確率変数 X において，

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt,$$

となる $f_X(.)$ を X の確率密度関数（probability density function, pdf）という．

- 連続型確率変数 X が a を超えて b 以下の値をとる確率を，確率密度関数を用いて表すと，

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

$\Rightarrow$ 確率が積分で表されているので，$f_X(x)$ のグラフの線よりも下の部分の面積が確率を表す．

- 累積分布関数 $F_X(x)$ が微分可能なら，$F_X(x)$ を微分すると確率密度関数 $f_X(x)$ となる．つまり，

$$F'_X(x) = f_X(x).$$

確率密度関数の性質：

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

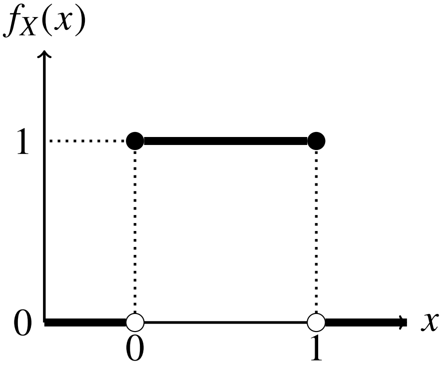
一様分布

- ルーレットの円周上に 0 から 1 までの数値を定義して，ルーレットを回して針が止まる位置を X とすると，X は一様分布に従う連続型確率変数であると考えることができる．

- ルーレットを回して針が止まる位置の確率密度関数は，

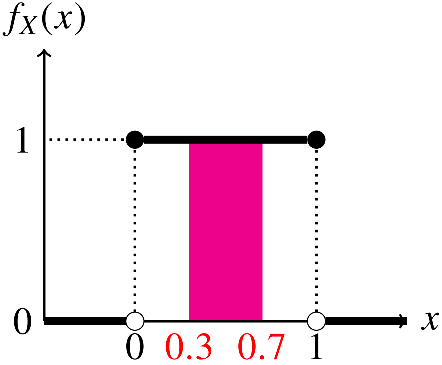
$$f_X(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq 1, \\ 0 & \text{elsewhere.} \end{cases}$$

ルーレットを回して針が止まる位置の確率密度関数をグラフで表すと,



- 連続型確率変数では，確率密度関数のグラフの線よりも下の部分の面積が確率を表す.

ルーレットを回して針が 0.3 の点から 0.7 の点までの間に止まる確率 $\Pr(0.3 < x \leq 0.7)$ は，下のグラフの塗りつぶされている部分.

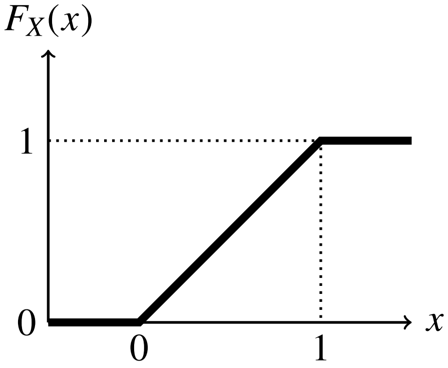


- 連続型確率変数では，確率密度関数のグラフの線よりも下の部分の面積が確率を表す.

- ルーレットを回して針が止まる位置の累積分布関数は,

$$F_X(x) = \begin{cases} 0 & \text{for } x < 0, \\ x & \text{for } 0 \leq x \leq 1, \\ 1 & \text{for } x > 1. \end{cases}$$

ルーレットを回して針が止まる位置の累積分布関数をグラフで表すと,



離散型確率変数と連続型確率変数の違いのまとめ

離散型確率変数：

- 非連続的な値をとる.
- 確率質量関数のグラフの高さが確率を表す.

連続型確率変数：

- 連続的な値をとる.
- 確率密度関数のグラフの線よりも下の部分の面積が確率を表す.

正規分布

- 確率密度関数が,
$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right),$$
の連続型確率分布を，期待値 μ, 分散 σ^2 の正規分布（normal distribution）という.
  - π は円周率（$3.14159\cdots$）.
  - exp(.) はネイピア数 $e = 2.71828\cdots$ を底とする指数関数.
$$\exp(a) = e^a.$$
- 確率変数 X が期待値 μ, 分散 σ^2 の正規分布に従うことを,
$$X \sim N(\mu, \sigma^2),$$
と表す.

標準正規分布

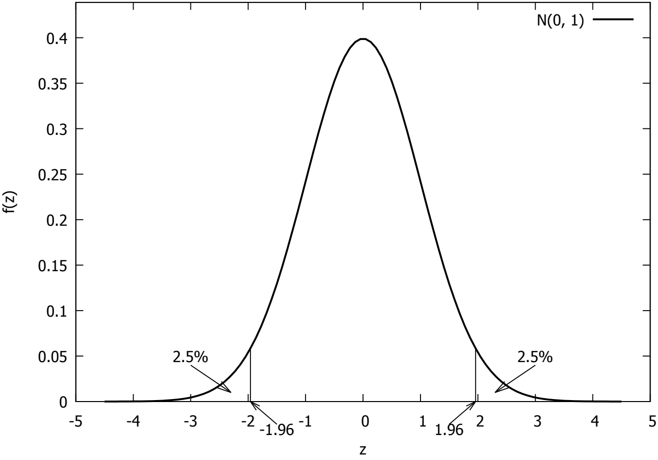
- 期待値 0, 分散 1 の正規分布を標準正規分布（standard normal distribution）という.
  - $N(0,1)$ と表す.
  - 確率密度関数は,

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

  - 確率密度関数のグラフが 0 を中心に左右対称のベル型.
  - 標準正規分布に従う確率変数を Z とすると,

$$\begin{aligned}\Pr(|Z| \geq 1.96) &= \underbrace{\Pr(Z \leq -1.96)}_{=0.025} + \underbrace{\Pr(Z \geq 1.96)}_{=0.025} \\ &= 0.05.\end{aligned}$$

標準正規分布の確率密度関数



自由度

- 自由に値を選ぶことのできる変数の数を自由度（degree of freedom）という．
  - e.g., 変数 X_1 と X_2 の間に，

$$0.5X_1 + 0.5X_2 = 3,$$

という関係があるとき，X_1 の値を自由に決めると X_2 の値は自動的に決まる．

$$\Downarrow$$

X_2 の値は自由に選ぶことができない．

$$\Downarrow$$

この場合，自由度は 1.

カイ二乗分布

- 連続型確率変数 $Z_1, Z_2, \cdots, Z_n$ が互いに独立に $N(0,1)$ に従っているとき，

$$X = Z_1^2 + Z_2^2 + \cdots + Z_n^2,$$

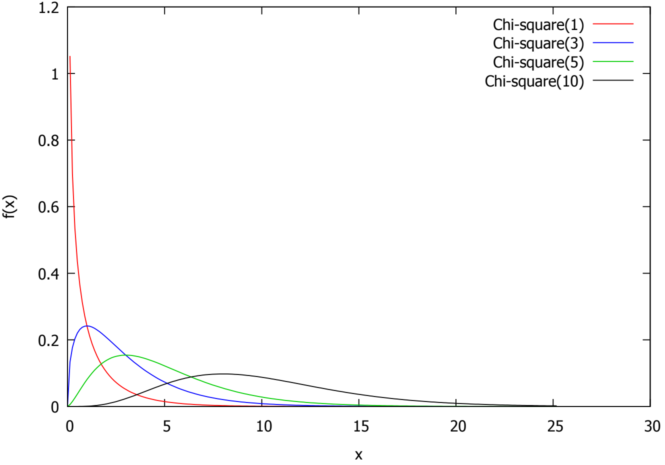
  の確率分布を自由度 n のカイ二乗分布（chi-square distribution）という．

- 確率変数 X が自由度 n のカイ二乗分布に従うことを，

$$X \sim \chi^2(n),$$

  と表す．

カイ二乗分布の確率密度関数



t 分布

- $Z \sim N(0,1)$, $X \sim \chi^2(n)$ として，Z と X が互いに独立のとき，

$$T = \frac{Z}{\sqrt{X/n}},$$

の確率分布を自由度 n の t 分布（t distribution）という．

- 確率変数 T が自由度 n の t 分布に従うことを，

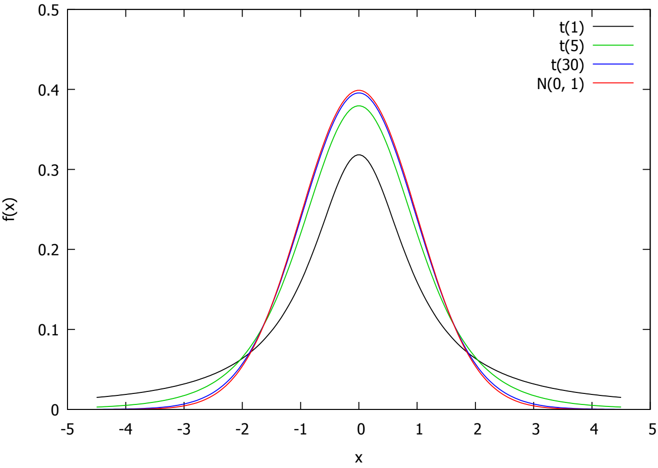
$$T \sim t(n),$$

と表す．

t 分布の特徴：

- 確率密度関数のグラフが 0 を中心に左右対称のベル型で，標準正規分布に似ている．
- $t(1)$ はコーシー分布と呼ばれる．
- 自由度 n が小さいときは分布の裾が分厚いが，n が大きくなるにつれて分布の裾がしだいに薄くなっていき，グラフの形が標準正規分布に近づいていく．
- $t(\infty)$ は標準正規分布 $N(0, 1)$ と同じになる．

t 分布と標準正規分布の確率密度関数



F 分布

- $U \sim \chi^2(m), V \sim \chi^2(n)$ として，U と V が互いに独立のとき，

  $$X = \frac{U/m}{V/n},$$

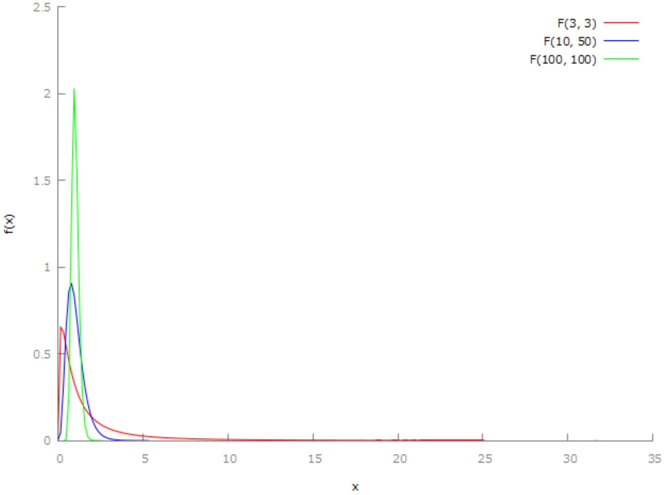
  の確率分布を自由度 (m, n) の F 分布（F distribution）という．

- 確率変数 X が自由度 (m, n) の F 分布に従うことを，

  $$X \sim F(m, n),$$

  と表す．

F 分布の確率密度関数



今日のキーワード

確率変数，実現値，確率分布，離散型確率変数，離散型確率分布，連続型確率変数，連続型確率分布，確率質量関数，（離散型確率変数の）累積分布関数，期待値，分散，標準偏差，同時確率質量関数，周辺確率質量関数，共分散，条件付き確率質量関数，独立，条件付き期待値，条件付き分散，（連続型確率変数の）累積分布関数，確率密度関数，正規分布，標準正規分布，自由度，カイ二乗分布，t 分布，F 分布

次回までの準備

- 今回の講義スライドを読み直す.
- 教科書第 4 章第 1 節〜第 3 節を読む.