

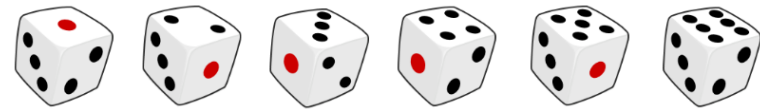
IV. 確率・統計

確率変数

✓ 確率変数とは、ある変数の値をとる確率が存在する変数をいう。

➤ サイコロを投げて出る目は確率変数

→ $\{1, 2, 3, 4, 5, 6\}$ のいずれかであり、それぞれの目が出る確率 $\left(\frac{1}{6}\right)$ が存在



➤ コインを投げて出る面は表 = 1、裏 = 0 とすれば確率変数

→ $\{0, 1\}$ のいずれかであり、それぞれの面が出る確率 $\left(\frac{1}{2}\right)$ が存在



* このように、数値でないデータを数値に変換した変数をダミー変数という。

➤ 金融商品のレート・価格の変化率は確率変数

→ あるレート・価格の変化率が一定の範囲に収まる確率が存在



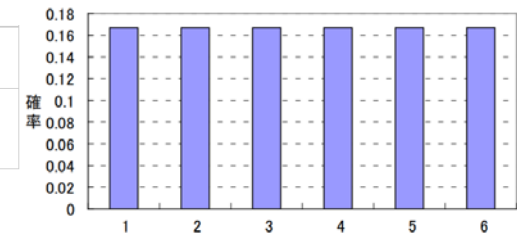
確率分布

✓ 確率分布とは、確率変数がそれぞれどのような確率で発生するかを示すものをいう。

➤ サイコロ1個を投げて出る目の確率分布

$$P(X) = \frac{1}{6} \quad (X = 1, 2, 3, 4, 5, 6)$$

サイコロ1個の目	1	2	3	4	5	6
確率	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

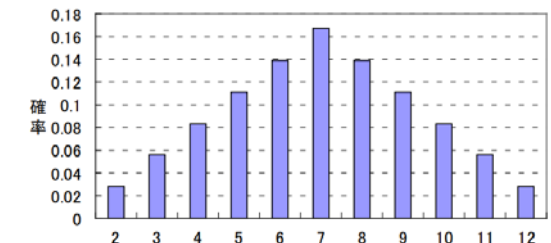


➤ サイコロ2個を投げて出る目の和の確率分布

$$P(X) = \frac{1}{36} \quad (X = 2, 12), \quad P(X) = \frac{2}{36} \quad (X = 3, 11), \quad P(X) = \frac{3}{36} \quad (X = 4, 10),$$

$$P(X) = \frac{4}{36} \quad (X = 5, 9), \quad P(X) = \frac{5}{36} \quad (X = 6, 8), \quad P(X) = \frac{6}{36} \quad (X = 7)$$

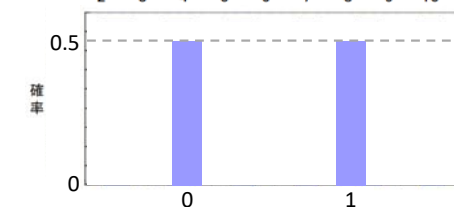
サイコロ2個の目の和	2	3	4	5	6	7	8	9	10	11	12
確率	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$



➤ コインを投げて出る面(表 = 1、裏 = 0)の確率分布

$$P(X) = \frac{1}{2} \quad (X = 0, 1)$$

コインの面	表(=1)	裏(=0)
確率	$\frac{1}{2}$	$\frac{1}{2}$



➤ 金融商品のレート・価格の変化率の確率分布 → 正規分布

期待値

- ✓ 期待値とは、確率変数の平均的なレベルを示す数値。
- ✓ 確率変数の、変数値と発生確率の積の和。
- ✓ 数式で表すと下記の通り。

$$E(X) = \sum_{i=1}^n x_i p_i$$

X : 確率変数 n 個のデータの集合

x_i : X に属するデータ ($i = 1, 2, 3, \dots, n$)

p_i : x_i の発生確率

- ✓ $E(X)$ の代わりに μ_X と表されることもある。
- ✓ 慣用的に「平均(値)」ともいうが、厳密には期待値と平均値は違う。 → スライド41で後述

分散

✓ 分散とは、確率変数の値のちらばり状況を示す数値。

✓ 確率変数の、偏差の自乗と発生確率の積の和。

* 偏差とは、期待値との差のこと。

➤ 偏差の自乗の期待値という言い方も可能

✓ 数式で表すと下記の通り。

$$Var(X) = \sum_{i=1}^n (x_i - \mu_X)^2 p_i$$

X : 確率変数 n 個のデータの集合

x_i : X に属するデータ ($i = 1, 2, 3, \dots, n$)

μ_X : X の期待値

p_i : x_i の発生確率

* 偏差の自乗の期待値なので
 $Var(X) = E[(x_i - \mu_X)^2]$ とも表せる。

✓ 単位を持たない。

予測値と実測値

- ✓ これまでの説明は、予測値が前提。 * サイコロであれば振る前、コインであれば投げる前。
- ✓ データ x_i の発生確率が i にかかわらず同じ場合、

期待値

$$E(X) = \sum_{i=1}^n x_i p_i$$

分散

$$Var(X) = \sum_{i=1}^n (x_i - \mu_X)^2 p_i$$

は、 $p_i = \frac{1}{n}$ として、それぞれ下記のように表せる。

期待値

$$E(X) = \sum_{i=1}^n \left(x_i \cdot \frac{1}{n} \right) = \frac{1}{n} \sum_{i=1}^n x_i = \frac{\sum_{i=1}^n x_i}{n}$$

分散

$$Var(X) = \sum_{i=1}^n \left\{ (x_i - \bar{x})^2 \cdot \frac{1}{n} \right\} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

n : データ数

x_i : i 番目のデータ ($i = 1, 2, 3, \dots, n$)

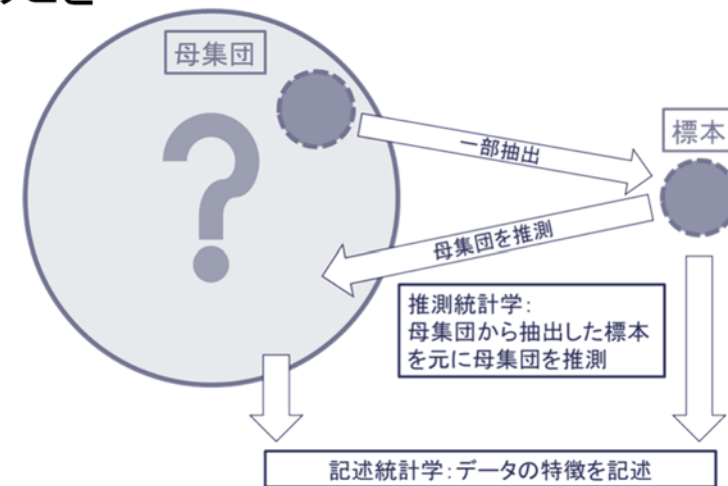
- ✓ これらの式はデータが n 個の実測値の場合に適用でき、その際の期待値はデータの平均値。
 - * サイコロであれば n 回振った後、コインであれば n 回投げた後。
 - * 平均値を $\bar{x}$ と表すこともある。

推測統計学と記述統計学

- ✓ そもそも統計学には母集団の捉え方によって2種類の立場がある。

→ 母集団: 本来知りたいと思っている集団全体のこと

- 推測統計学: 観測データを母集団から抽出した標本と考えて、統計量を算定、母集団の特性を推測、検証
- 記述統計学: 観測データを母集団全体と考えて、統計量を算定、観測データを分析



- ✓ 金融リスク管理の実務での統計的分析においては、いずれの立場も想定されうる。

	推測統計学	記述統計学
立場	観測データを母集団から抽出された標本(サンプル)と見なす	観測データを母集団と見なす
母集団	その金融商品が理論上とりうる、すべてのレート・価格の変化率	実際に観測されたレート・価格の変化率