確 率 論 第６回
-- 確率変数と確率分布関数 --

情報工学科 山本修身

確率空間の復習

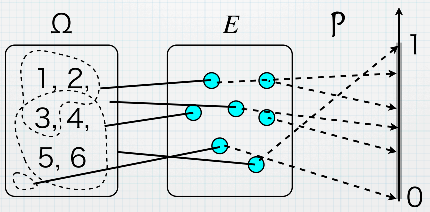
* 確率空間とは，標本空間，事象の集合，確率が組になったものである．すなわち，

$$(\Omega, E, \mathbb{P})$$

のことを**確率空間**と呼ぶ．確率空間があらかじめ与えられていないと確率を計算することはできない．

サイコロの確率空間（復習）

* たとえば，サイコロの確率空間はつぎのように図示される．

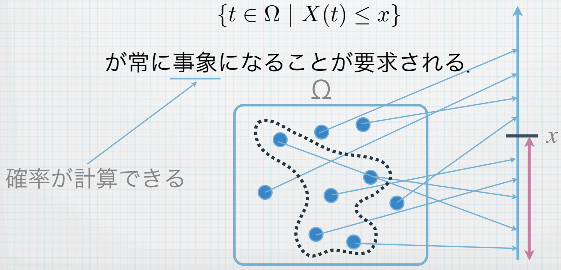


確率変数とは(1)

* 確率変数は厳密には変数ではなく，標本空間から実数への写像である．
* 任意の実数 x について，標本空間の部分集合

$$\{t \in \Omega \mid X(t) \le x\}$$

が常に事象になることが要求される．

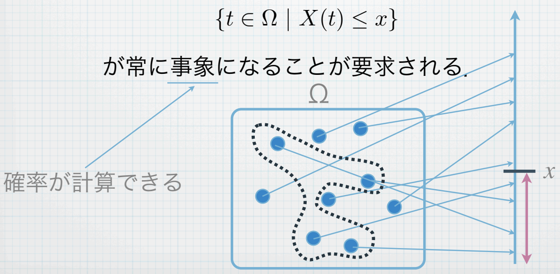


確率変数とは (2)

* 確率変数は厳密には変数ではなく，標本空間から実数への写像である．
* 任意の実数 x について，標本空間の部分集合

$$\{t \in \Omega \mid X(t) \le x\}$$

が常に事象になることが要求される．

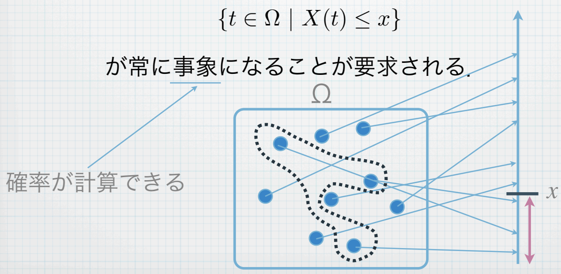


確率変数とは (3)

* 確率変数は厳密には変数ではなく，標本空間から実数への写像である．
* 任意の実数 x について，標本空間の部分集合

$$\{t \in \Omega \mid X(t) \le x\}$$

が常に事象になることが要求される．



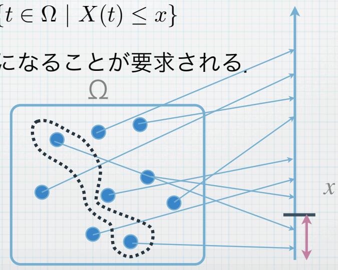
確率変数とは (4)

* 確率変数は厳密には変数ではなく，標本空間から実数への写像である.
* 任意の実数 x について，標本空間の部分集合

$$\{t \in \Omega \mid X(t) \le x\}$$

が常に事象になることが要求される.

確率が計算できる



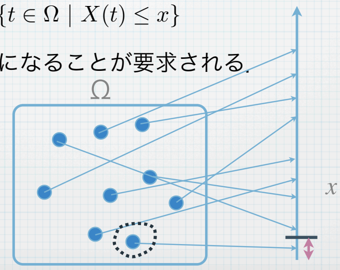
確率変数とは (5)

* 確率変数は厳密には変数ではなく，標本空間から実数への写像である.
* 任意の実数 x について，標本空間の部分集合

$$\{t \in \Omega \mid X(t) \le x\}$$

が常に事象になることが要求される.

確率が計算できる



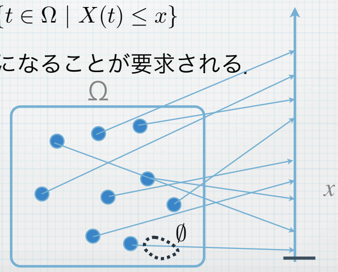
確率変数とは (6)

* 確率変数は厳密には変数ではなく，標本空間から実数への写像である.
* 任意の実数 x について，標本空間の部分集合

$$\{t \in \Omega \mid X(t) \le x\}$$

が常に事象になることが要求される.

確率が計算できる



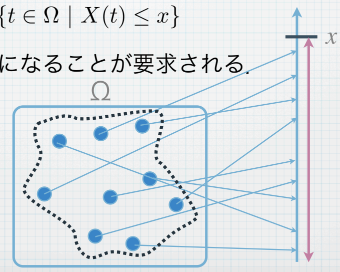
確率変数とは (7)

* 確率変数は厳密には変数ではなく，標本空間から実数への写像である.
* 任意の実数 x について，標本空間の部分集合

$$\{t \in \Omega \mid X(t) \le x\}$$

が常に事象になることが要求される.

確率が計算できる



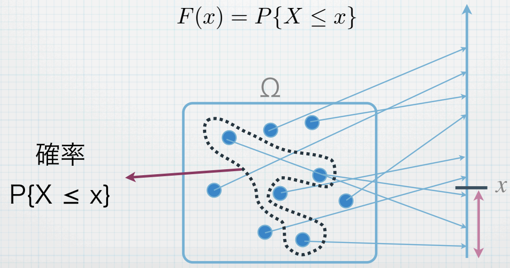
確率変数の効用

* 確率変数を導入することにより，すべての事象を対象するのではなく，{X ≤ x}という事象のみを対象とすることになり，議論が単純になる.
* 違う確率空間であっても確率変数としては同じ構造をもつ場合があり，良くでてくる確率変数を分類して研究することができる.

分布関数

* 分布関数はあるxを指定したとき，それに確率P{X ≤ x}を対応させる関数である.

$$F(x) = P\{X \le x\}$$



$F(x) = P\{X \leq x\}$

分布関数の形

- しきい値 x が十分大きくなれば，分布関数の値は，ほぼ全部の根元事象を含む事象の確率になる

分布関数の性質

- 分布関数の明らかに成り立つ性質としてつぎのようなものがある．

$$\lim_{x \to -\infty} F(x) = 0,$$
$$\lim_{x \to \infty} F(x) = 1$$

さらに，$F(x)$は単調増加関数である．

サイコロを振る試行の分布関数(1)

- サイコロを振って，その目を観測するという試行を考える．出る目は1から6までであり，それぞれが出るという根元事象をR_iと書こう．標本空間は

$$\Omega = \{R_1, R_2, R_3, R_4, R_5, R_6\}$$

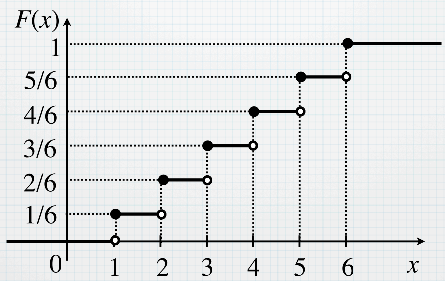
である．また，確率変数Xはつぎのように定義する：

$$X(R_i) = i$$

$F(x) = P\{X \leq x\}$

サイコロを振る試行の分布関数(2)

- 確率変数Xの分布関数はつぎのようになる．



連続的な確率変数の分布関数(1)

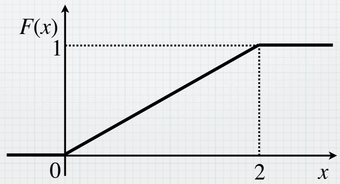
- 連続的な確率分布の例として，0から2までの数がすべて等確率で出現するような試行を考える．
- 連続的な分布であるので，たとえば1.0が出現する確率を考えれば0となる．区間$[a, b]$の値が出現する確率は$(b - a) / 2$と定義する．

連続的な確率変数の分布関数(2)

- 分布関数はF(x) = P{X ≤ x}と定義されるので，

$$F(x) = P\{X \leq x\} = \frac{x - 0}{2} \quad (x \leq 2)$$

となる．



確率密度関数とは

- **確率密度関数**とは，確率分布関数の導関数のことである．特に連続的な確率変数については分布関数を眺めるよりも密度関数を眺める方が確率分布の性質を直感的に理解しやすい．
- 定義より，

$$f(x) = \frac{dF(x)}{dx}$$

と書ける．

確率密度関数の性質

- 確率密度関数は分布関数の微分であり，確率分布（確率変数）を分布関数で眺めても密度関数で眺めても情報量に変わりはない．
- 分布関数を密度関数を用いて表現すれば，つぎのようになる：

$$F(y) = \int_{-\infty}^{y} f(x)dx$$

確率密度の解釈(1)

- 確率密度関数は微分の定義より，

$$f(x) = \lim_{\Delta x \to 0} \frac{P(X \leq x + \Delta x) - P(X \leq x)}{\Delta x}$$

が成り立ち，さらに分母の解釈から，

$$f(x) = \lim_{\Delta x \to 0} \frac{P(x < X \leq x + \Delta x)}{\Delta x}$$

と書ける．

確率密度の解釈(2)

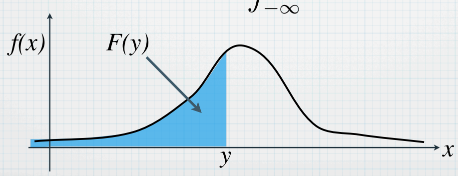
- 最後の式から，確率密度は確率変数Xが区間$[x, x + \Delta x]$に含まれる確率をΔxで割った値の極限値であることがわかる．これは，確率そのものではなく，この区間に含まれる確率がΔxが小さくなると共にどの程度小さくなるかという割合を示している．

$$f(x) = \lim_{\Delta x \to 0} \frac{P(x < X \leq x + \Delta x)}{\Delta x}$$

確率密度関数と分布関数の関係

- 密度関数は分布関数の導関数であるが，逆に考えれば分布関数は密度関数の積分である．すなわち，つぎのように書ける：

$$F(y) = \int_{-\infty}^{y} f(x)dx$$



密度関数の全域の積分は1である

- 密度関数を全域で積分すると1である．
- なぜならば，

$$\lim_{x \to \infty} F(x) = 1$$

より，

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

確率関数

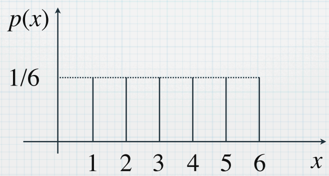
* 離散した確率変数（確率変数で離散した値しかとらないもの）については前述のように分布関数は階段関数になる．従って，分布関数を定理することができない．
* 密度関数に対応するものとして，確率関数を定義することができる．ただし，密度関数とは次元が異なることを注意する必要がある．

$$p(x) = P(X = x)$$

サイコロの確率関数は…

* サイコロの確率関数は以下のように定義できる：

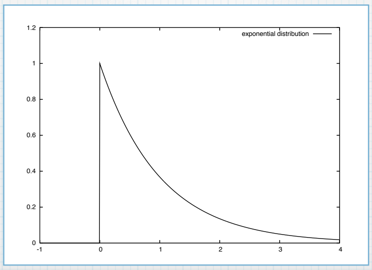
$$p(x) = \begin{cases} 1/6 & (x = 1, 2, 3, 4, 5, 6) \\ 0 & (\text{otherwise}) \end{cases}$$



確率分布の例：指数分布

* 指数分布はつぎのような密度関数を持つ確率分布である：

$$f(x) = \lambda \mathrm{e}^{-\lambda x} \quad (x > 0)$$



指数分布の性質 $\int_0^y \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x}\right]_0^y = 1 - e^{-\lambda y}$

* 密度関数を積分することにより，分布関数は，

$$F(y) = 1 - \mathrm{e}^{-\lambda y} \quad (y > 0)$$

である．また，確率変数Xがx以上であるという条件のもとで，$[x, x + \Delta x]$に入る確率は

$$\begin{aligned} P\left(x < X \le x + \Delta x \mid x < X\right) &= \frac{P(x < X \le x + \Delta x)}{P(x < X)} \\ &= \frac{F(x + \Delta x) - F(x)}{1 - F(x)} \\ &= \frac{f(x)\Delta x}{1/\lambda f(x)} = \lambda \Delta x \end{aligned}$$

と計算できる．

指数分布従う現象

* 色々な現象があるが，有名なのは，放射性の原子が崩壊するまでの時間の分布である．時間とともに崩壊していない確率は小さくなっていく．ちょうど崩壊していない確率が1/2になる時間を半減期と呼ぶ．これはこの分布の性質から ln2 / λ となる．

$$\boxed{F(y) = 1 - \mathrm{e}^{-\lambda y} \quad (y > 0)}$$

$$\begin{aligned} 1 - e^{-\lambda x} &= 1/2 \\ e^{-\lambda x} &= 1/2 \\ -\lambda x &= -\ln 2 \\ x &= \ln 2/\lambda \end{aligned}$$

幾何分布

* 幾何分布は離散分布でありながら，とる値が無限個存在する分布である．
* 1枚のコインを何回も振る試行を考える（これをベルヌーイ試行と呼ぶ）．このとき，表が最初にでるまでの振った回数の分布である．確率関数はつぎのようになる．

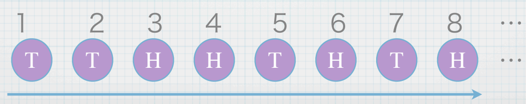
H: p

T: q = 1 - p

$$P(X = k) = q^{k-1}p \quad (k = 1, 2, 3, \ldots)$$

ただし，$q = 1 - p$

∞回振る



幾何分布の性質

* 確率変数は無限種類の値をとるが，確率を全部足すと1になる．

$$\sum_{i=1}^{\infty} P(X=k) = \sum_{i=1}^{\infty} q^{k-1} p = \frac{p}{1-q} = 1$$

$$q = 1 - p$$