情報理論
第 3 回 確率分布

堀田 政二
工学部 情報工学科

確率変数と確率分布 (復習)

- 標本空間 Ω の中で定義される変数を X とし，この X がある具体的な値 x_k をとる確率が既知である場合，$X = x_k$ となる事象の確率を $P(X = x_k)$，あるいは $p(x_k)$ で表す
- このような変数 X を確率変数 (random variable) と呼ぶ
- 離散的確率変数: X の取りうる値が有限個，あるいは可算無限個の場合
- 連続的確率変数: X の取りうる値が連続で無限個の場合
- 確率変数 X とそれに対応する確率 $P(X = x_k)$ との対応関係を確率分布 (probability distribution) とよぶ

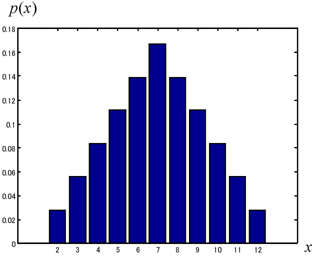
例：X を二つのサイコロを振った場合の和とした場合

X	2	3	4	5	6	7	8	9	10	11	12
$p(x_k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

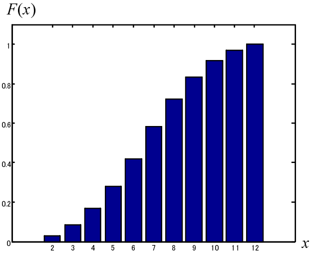
(2)

離散的確率変数の確率分布と分布関数

- **確率分布** $p(x_k)$ **の性質** $(x_1 < x_2 < \cdots < x_{n-1} < x_n)$
  - **全ての** x_k **に対して** $p(x_k) \geq 0$**，** $\sum_{k=1}^{n} p(x_k) = 1$
  - **確率変数** X **が** x_k **以下となる確率**
    $P(X \leq x_k) = \sum_{r=1}^{k} p(x_r) = F(x_k)$
  - $F(x_k)$ **を分布関数とよぶ**
- **一方，分布関数が与えられれば，確率が計算できる．具体的には，** $P(X = x_k) = P(X \leq x_k) - P(X \leq x_{k-1})$



確率分布



分布関数

期待値と分散

確率変数 X の期待値 (平均) と分散は次で定義される:

- 期待値

$$\mu = E(X) = \sum_{k=1}^{n} x_k p(x_k)$$

- 分散

$$\sigma^2 = V(X) = E((X-\mu)^2) =$$
$$\sum_{k=1}^{n} (x_k - \mu)^2 p(x_k) = E(X^2) - E(X)^2$$

確率変数の和と積の期待値・分散

二つの確率変数 X と Y の期待値と分散について考える

$E(X) = \sum_{k=1}^{n} x_k p(x_k)$, $E(Y) = \sum_{l=1}^{m} y_l p(y_l)$ **のとき**

- $E(X+Y) = E(X) + E(Y)$
- X **と** Y **が独立のとき**

$$E(XY) = E(X) \times E(Y)$$

- X **と** Y **が独立のとき**

$$V(X+Y) = V(X) + V(Y)$$

$E(X+Y)=E(X)+E(Y)$ の証明

- $P(X=x_k \cap Y=y_l)=p(x_k,y_l)$
- $\sum_{l=1}^{m} p(x_k,y_l)=p(x_k)$, $\sum_{k=1}^{n} p(x_k,y_l)=p(y_l)$, $\sum_{k=1}^{n}\sum_{l=1}^{m} p(x_k,y_l)=1$

より

$$
\begin{aligned}
E(X+Y) &= \sum_{k=1}^{n}\sum_{l=1}^{m}(x_k+y_l)p(x_k,y_l) \\
&= \sum_{k=1}^{n} x_k\Big(\sum_{l=1}^{m} p(x_k,y_l)\Big)+\sum_{l=1}^{m} y_l\Big(\sum_{k=1}^{n} p(x_k,y_l)\Big) \\
&= \sum_{k=1}^{n} x_k p(x_k)+\sum_{l=1}^{m} y_l p(y_l) \\
&= E(X)+E(Y)
\end{aligned}
$$

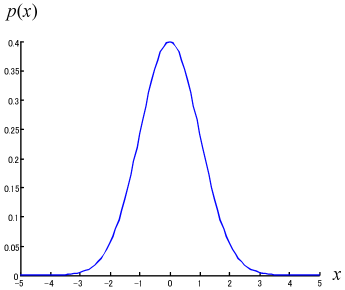
連続的確率変数と確率密度関数

連続値をとる確率変数 X **に対して，任意の定数** $a < b$ **により確率** $P(a \le X \le b)$ **が**

$$P(a \le X \le b) = \int_a^b p(x)dx$$

で定められるような連続関数 $p(x)$ **が** $(-\infty, \infty)$ **で存在するとき，この** $p(x)$ **を，確率密度関数** (probability density function) **と呼ぶ．また，確率密度関数は次の性質を持つ**:

- **任意の** x **に対して** $p(x) \ge 0, \quad \int_{-\infty}^{\infty} p(x)dx = 1$

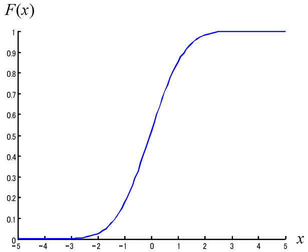


累積分布関数，平均と分散

確率変数 X が区間 $-\infty < X \leq x$ にある確率が

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

で定められる関数 $F(x)$ を，確率変数 X の累積分布関数 (cumulative distribution function) と呼ぶ．



連続量の確率変数 X の期待値と分散は次式で与えられる:

- $\mu = E(X) = \int_{-\infty}^{\infty} xp(x)dx$
- $\sigma^2 = V(X) = E\left((X-\mu)^2\right) = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)dx$

ベルヌーイ分布 (離散確率分布)

母集団が無数の 0 と 1 からなり，1 の割合が p であるとする．この母集団分布は，確率変数 X を使って，以下の表のように表せる．これをベルヌーイ分布とよぶ．

x_k	0	1
$P(X = x_k)$	$1 - p$	p

期待値: $E(X) = p$，**分散**: $V(X) = p(1 - p)$

- **母集団が全有権者からなり，内閣を支持する者を 1，その他のものに 0 が割り振られているとき，p は母集団における内閣支持率になる**

二項分布 (離散確率分布)

例：ある集団において，特性 A を持つ者の割合が p であり，持たない者の割合が $1-p$ であるとする．この時，集団から無作為に n 人を抽出したとき，特性 A を持つ者が x 人である確率を考える．

- **n 人のうち x 人が特性を持つ組合せは ${}_nC_x$ 通りある**
- **その各々に対して特性 A を持つ確率は p^x，残り $n-x$ 人が特性を持たない確率は $(1-p)^{n-x}$ であり，両者が共に起こるのは両者の積．したがって，求めたい確率分布は**

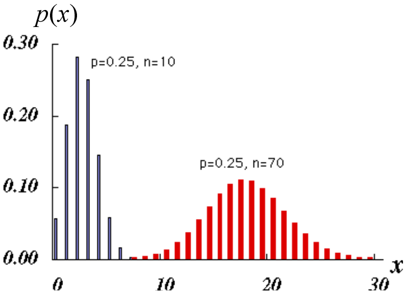
$$p(X=x) = {}_nC_x p^x (1-p)^{n-x},\ \ x = 0, 1, ..., n, p > 0$$

$$= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

であり，これを二項分布と呼ぶ (ベルヌーイ分布の一般化)．

二項分布の例と性質

- 中国では B 型の割合はおおよそ 25% $(p = 0.25)$ である．無作為に $n = 10$ 人と $n = 70$ 人を選んだとき，B 型の人が n 人のうち x 人含まれる確率
- n が大きくなると左右対称な正規分布に近づく
- $p = 0.5$ であれば左右対称になる



二項分布の平均

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{n} xp(x) = \sum_{x=0}^{n} x\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} \\
&= \sum_{x=1}^{n} x\frac{n(n-1)!}{x(x-1)!((n-1)-(x-1))!}pp^{x-1}(1-p)^{(n-1)-(x-1)} \\
&= np\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!((n-1)-(x-1))!}p^{x-1}(1-p)^{(n-1)-(x-1)}
\end{aligned}
$$

$x-1=x'$，$n-1=n'$ と置くと上式は

$$
E(X) = np\sum_{x'=0}^{n'} \frac{n'!}{x'!(n'-x')!}p^{x'}(1-p)^{n'-x'} = np
$$

二項分布の分散

$$
\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= \sum_{x=1}^{n} x^2 \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} - (np)^2 \\
&= np \sum_{x=1}^{n} x \frac{(n-1)!}{(x-1)!((n-1)-(x-1))!} p^{x-1} (1-p)^{(n-1)-(x-1)} - (np)^2
\end{aligned}
$$

$x-1 = x'$，$n-1 = n'$ と置くと上式は

$$
\begin{aligned}
V(X) &= np \sum_{x'=0}^{n'} (x'+1) \frac{n'!}{x'!(n'-x')!} p^{x'} (1-p)^{n'-x'} - (np)^2 \\
&= np((n-1)p+1) - (np)^2 = np(1-p)
\end{aligned}
$$

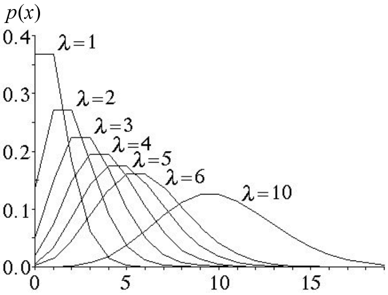
ポアソン分布 (離散確率分布)

二項分布において，n が十分大きく，p が小さいとき，np は適度な大きさとなるためパラメータ $\lambda = np$ を持つ以下のポアソン分布が二項分布の良好な近似となる:

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}$$

ただし，$x = 0, 1, ..., \lambda > 0$

- **期待値**: $E(X) = \lambda$，**分散**: $V(X) = \lambda$，$e = 2.71828...$
- λ **が大きくなると正規分布に近づく**



ポアソン分布の例

- $P(X=x)$ は，単位時間中に平均で λ 回発生する事象が，ちょうど x 回（$x=0,1,2,...$）発生する確率に相当する
- 例えば，事象が平均で 2 分間に 1 回発生する場合，10 分間の中で事象が発生する回数は $\lambda=5$ のポアソン分布モデルを使って求められる

交通事故件数，有感地震回数，カウンターへの到着客数等

【例題】あるデパートのオーダーメイド服売り場に，1 時間当たり平均 2 名の客が来る．来客数がポアソン分布に従うとき，次の確率を求めよ．

1. 1 時間に 4 名以上の客が来る確率
2. 1 時間に少なくとも 1 名の客が来る確率
3. 2 時間の間，まったく客の来ない確率
4. 次の客が来るまでの間隔が 30 分を超える確率

(15)

ポアソン分布の例

【解答】1 時間の来客数を X とすると，X はパラメータ 2 のポアソン分布に従う．

1. 1 時間に 4 名以上の客が来る確率

$$P(X \geq 4) = \sum_{x=4}^{\infty} \frac{e^{-2}2^x}{x!} = 1 - e^{-2} \sum_{x=0}^{3} \frac{2^x}{x!} = 0.143$$

2. 1 時間に少なくとも 1 名の客が来る確率

$$1 - P(X = 0) = 1 - e^{-2} = 0.865$$

3. 2 時間の間，まったく客の来ない確率

$$P(X = 0)^{\times} P(X = 0) = e^{-2} \times e^{-2} = 0.018$$

4. 30 分間の客の数を Y とすると，Y はパラメータ 1 のポアソン分布に従う．$Y = 0$ となる確率を求めればよいので

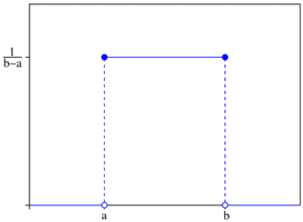
$$P(Y = 0) = e^{-1} = 0.368$$

一様分布 (連続確率分布)

確率変数 X が次の密度関数を持つとき，X は区間 $[a,b]$ 上の一様分布に従うという

$$p(x) = \begin{cases} 1/(b-a) & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

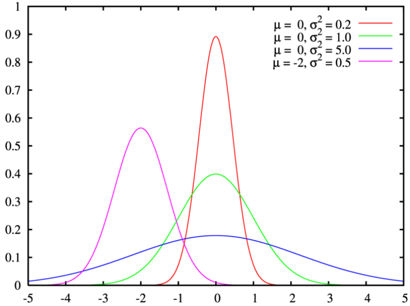
- **期待値**: $E(X) = (a+b)/2$，**分散**: $V(X) = (b-a)^2/12$



正規分布 normal distribution (連続確率分布)

確率変数 X が次の密度関数を持つとき，X は平均 μ，分散 σ^2 の正規分布 $\mathcal{N}(\mu,\sigma^2)$ に従うという

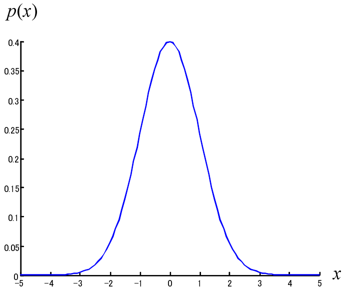
$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},\quad -\infty\le x\le\infty,\quad \sigma^2>0$$



標準正規分布 (連続確率分布)

変数変換 (標準化という) $x \leftarrow (x-\mu)/\sigma$ により得られた確率変数 x は $\mathcal{N}(0,1)$ の標準正規分布に従う

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



中心極限定理 (central limit theorem)

中心極限定理

期待値が μ**，分散が** σ^2 **である任意の確率分布に従う互いに独立な確率変数** $X_1, X_2, ..., X_n$ **の総和**

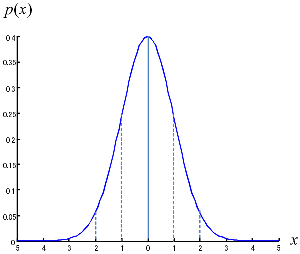
$$Z = X_1 + X_2 + \cdots X_n$$

は，やはり確率変数であり， n **が十分大きいとき，** Z **の確率分布は期待値** $n\mu$**，分散** $n\sigma^2$ **の正規分布であるとみなすことができる．特に** $n \to \infty$ **の極限では，** Z **の確率分布は正規分布に収束する．**

例: $[0, 1]$ **上の一様分布に従う** $n = 10$ **の確率変数の総和の分布** (500 **回**) **をプロットしてみる** (**デモ**)
$\Rightarrow$ **統計学や自然科学，社会科学の多くの分野で複雑な現象を簡単に表すモデルとして用いられている**

正規分布の性質

計算機を使うことなく正規分布に従った事象の確率を求める事ができる



- $P(X \leq 0) = 0.5$
- $P(\mu - \sigma \leq X \leq \mu + \sigma) = 0.682$
- $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.954$
- $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.998$

問題

- 【3.1】1 の目が出るまでサイコロを投げ続けるとき，投げる回数の確率分布を求めよ．
- 【3.2】電車が 24 時間，20 分間隔で走っているとする．適当な時刻にホームに着いたとき，10 分以上待つ確率を求めよ．
- 【3.3】次の関数

$$p(x) = \begin{cases} c(1-x^2) & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

が確率密度関数となるように c を求めよ．また，この分布の平均と分散を求めよ．
- 【3.4】確率変数 X が $\mathcal{N}(12, 3^2)$ の正規分布に従うとき，$P(9 \leq X \leq 18)$ を求めよ．