2004 年 6 月 1 日

総合科目　数理・自然 6「統計学入門」講義レジュメ　(今野)

標本空間と事象

1. 標本点 — 可能な結果
2. 標本空間 — 可能な結果の全体の集合
3. 事象 — 標本空間の部分集合．標本空間自身も事象となり，全事象とよぶ．
   (a) 根元事象 — 標本点ひとつによって構成される事象
   (b) 和事象 — ふたつの事象のどちらかに含まれる根元事象のあつまり
   (c) 積事象 — ふたつの事象の両方に含まれる根元事象のあつまり
   (d) 補事象 — その事象に含まれない根元事象のあつまり
   (e) 空事象 — 標本点を一つも含まず決して起こらない事象
   (f) 排反事象 — 共通部分を持たない事象．すなわち，一方が起これば，他方が起こらないこと

確率の定義

ある事象 A が起こる確率を $P(A)$ と書く．

確率の公理

1. すべての事象 A に対して，$0 \leq P(A) \leq 1$
2. $P(\text{標本空間}) = 1$
3. 互いに排反な事象 $A_1, A_2, A_3, \ldots,$ に対して

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

   が成立

加法定理

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

条件付き確率

B が起こった場合にそのうちにさらに A が起こる確率

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

乗法の定理

$$P(A \cap B) = P(B) \cdot P(A|B)$$

独立性

事象 A の起こる確率が他の事象 B に影響されないこと

$$P(A) = P(A|B)$$

乗法の定理からこれは

$$P(A \cap B) = P(A) \cdot P(B)$$

と同値

ベイズの公式

A を事象とする．また，H_1 と H_2 も事象とし，$H_1 \cup H_2 = \Omega$（全事象）かつ $H_1 \cap H_2 = \emptyset$（空事象）とする．このとき，

$$P(H_i|\,A) = \frac{P(H_i)P(A|\,H_i)}{P(H_1)P(A|\,H_1) + P(H_2)P(A|\,H_2)}, \qquad i = 1,\,2$$

が成立する．

問題 5〔ベイズの公式の応用〕

$P(C) = 0.005$ とする．また，

	検査の結果		
被験者の状態	陽性　(A)	陰性　(A^c)	合計
ガンである　(C)	0.95	0.05	1.00
ガンでない　(C^c)	0.05	0.95	1.00

であったとする．すなわち，

$$P(A|\,C) = 0.95,\ P(A^c|\,C) = 0.05,\ P(A|\,C^c) = 0.05,\ \ P(A^c|\,C^c) = 0.95$$

である．

(i) ベイズの公式を用いて $P(C|\,A)$ を求めよ．

(ii) ベイズの公式を用いて $P(C|\,A^c)$ を求めよ．

(iii) $P(A|\,C) = P(A^c|\,C^c) = r\,(0 < r < 1)$ とする．$P(C) = 0.005$ は変わらないとき，$P(C|\,A) \geq 0.90$ となるための r の範囲を求めよ．

2004 年 6 月 1 日

総合科目　数理・自然 6 「統計学入門」講義レジュメ　(今野)

確率変数とは

それがとる値 (とる値の範囲) に対しそれぞれ確率が与えれる変数である．確率変数には，離散型確率変数と連続型確率変数がある．

離散型確率変数

1. 一般に，可算集合 $\{x_1, x_2, \dots,\}$ の中に値を取る確率変数 X を離散型確率変数とよぶ．それぞれの値の確率

$$P(X = x_i) = p_i, i = 1, 2, \dots$$

を確率変数 X の確率分布という．である．ここで，

$$p_i \geq 0,\ i = 1, 2, \dots, \sum_{i=1}^{\infty} p_i = 1$$

である．

2. 例〔さいころ〕: さいころをふって出る目に対応する確率変数を X とする．さいころが正しいならば，

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

である．これを表にまとめたもの

X の値	1	2	3	4	5	6	合計
確率	1 / 6	1 / 6	1 / 6	1 / 6	1 / 6	1 / 6	1

を確率分布表という．

3. 例（コインを 3 回投げる）: コインを 3 回投げたとき，3 回中に表の出た回数に対応する確率変数を X とする．コインが正しいならば，

$$P(X=0) = \frac{1}{8}, P(X=1) = \frac{3}{8}, P(X=2) = \frac{3}{8}, P(X=3) = \frac{1}{8},$$

となる．

X の値	0	1	2	3	合計
確率	1 / 8	3 / 8	3 / 8	1 / 8	1

4. 例（さいころを 2 回投げる）：さいころを 2 回投げたとき，ふたつのさいころの目の和に対応する確率変数を X とする．このとき，

和	2	3	4	5	6	7	8	9	10	11	12	合計
確率	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

となる．たとえば，$P(X = 4)$ の確率は

$$\begin{aligned} P(X=4) &= P\{(1,3),(2,2),(3,1)\} = P\{(1,3)\} + P\{(2,2)\} + P\{(3,1)\} \\ &= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{3}{36} \end{aligned}$$

よりわかる．

連続型確率変数

1. $P(X = x) = 0$ となり，確率変数 X の取る値が関数 $f(x)$ によって

$$P(a < X \leq b) = \int_a^b f(x)\,dx$$

とあらわされる場合，X は連続型確率変数といい，その分布を連続型という．また，$f(x)$ のことを確率変数 X の確率密度関数という．ただし，

$$\text{すべての } x \text{ に対して} \quad f(x) \geq 0, \quad \text{かつ} \quad \int_{-\infty}^{\infty} f(x)\,dx = 1$$

となる．
ここで，Δx を極めて小さい数とすれば，

$$P(x < X \leq x + \Delta x) \approx f(x)\Delta x$$

である．

2. 例（指数分布）：$\lambda > 0$ として，

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

電球が切れるまでの時間〔故障時間）に対応する確率変数 X の確率分布とし利用される．

3. 例（一様分布）：

$$f(x) = \begin{cases} 1 & (0 \leq x \leq 1) \\ 0 & (\text{その他}) \end{cases}$$

$\boxed{\text{累積分布関数}}$

1. 定義

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

2. 離散型分布

$$F(x) = \sum u \le xp(u)$$

3. 連続型分布

$$F(x) = \int_{-\infty}^{x} f(u)\,du$$

4. 性質

(a) 広義単調増加

$$x_1 < x_2 \qquad \text{ならば} \qquad F(x_1) \le F(x_2)$$

(b) 範囲

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

(c) 右側連続

$$\lim_{\epsilon\downarrow 0} F(x+\epsilon) = F(x), \qquad \text{各 } x \text{ に対して}$$

$\boxed{\text{モードとメディアン}}$

モードとは確率（密度）関数 $f(x)$ を最大にする点である．メディアンは $P(X \le m) \ge 1/2$ と $P(X \ge m) \ge 1/2$ をみたす点である．例で与えられた確率分布について，累積分布関数をもとめ，そのグラフを描け．

$\boxed{\text{問題 6}}$ 連続型確率変数 X の確率密度関数は

$$f(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{2}, & 0 \le x \le 2, \\ 0, & x > 2 \end{cases}$$

で与えれれるとする．

(i) 累積分布関数 $F(x)$ を求め，そのグラフを描け．

(ii) メディアンを求めよ．すなわち，$F(m) = 1/2$ をみたす点を求めること．

$\boxed{\text{問題 7}}$ 離散型確率変数 X の確率関数は

$$P(X=0) = \frac{1}{8}, P(X=1) = \frac{3}{8}, P(X=2) = \frac{3}{8}, P(X=3) = \frac{1}{8},$$

で与えられるとする．このとき，X の累積分布関数を求め，そのグラフを描け．