大学院システム情報科学府 基礎共通科目　　2018 年 10 月 12 日

確率・統計特論 第 2 回

来嶋 秀治 (Shuji Kijima)

注意: 参照した文献等の情報を必ず記載すること.

1. 独立性，条件付き確率，ベイズの定理

定義 1. 2 つの事象 A と B に対して, $\Pr(A \cap B)$ を A と B の**同時確率** (*joint probability*) と言う. $\Pr(A \cap B)$ は $\Pr(A, B)$ とも書く.

定義 2. 事象 B の生起の下での事象 A の**条件付き確率** (*conditional probability*) は

$$\Pr(A \mid B) = \frac{\Pr(A, B)}{\Pr(B)}$$

と定義する.

定義 3. 2 つの事象 A と B が**独立** (*independent*) であるとは, $P(A, B) = P(A)P(B)$ が成り立つことをいう. 事象 $A_1, A_2, \ldots, A_k$ が**相互に独立** (*mutually independent*) とは, $P\left(\bigcap_{i \in I}^{n} A_i\right) = \prod_{i \in I} P(A_i)$ が任意の $I \subseteq \{1, \ldots, k\}$ について成り立つことをいう.

定理 4. 事象 $A_1, A_2, \ldots, A_n$ は相互排反とし, 標本空間 Ω に対して, $\bigcup_{i=1}^{n} A_i = \Omega$ が成り立つとする. このとき

$$\Pr(B) = \sum_{i=1}^{n} \Pr(A_i, B)$$

が成り立つ. この確率を**周辺確率** (*marginal probability*) という.

定理 5. (ベイズ (Bayes)) 2 つの事象 A と B に対して,

$$\Pr(A \mid B) = \frac{\Pr(B \mid A) \cdot \Pr(A)}{\Pr(B)}.$$

系 6. 事象 $A_1, A_2, \ldots, A_n$ は相互排反とし, 標本空間 Ω に対して, $\bigcup_{i=1}^{n} A_i = \Omega$ が成り立つとする. このとき

$$\Pr(A_i \mid B) = \frac{\Pr(B \mid A_i) \cdot \Pr(A_i)}{\sum_{j=1}^{n} \Pr(B \mid A_j) \cdot \Pr(A_j)}$$

が成り立つ.

2. 1 次元確率分布: 離散分布と連続分布

ここでは, 確率空間 $(\Omega, \mathcal{F}, P)$ が $\Omega \subseteq \mathbb{R}$ の場合を扱う.
離散分布: Ω が可算 (無限) 集合の場合.

- 確率関数 (probability function) $f(x) \stackrel{\text{def.}}{=} \Pr(X = x)$,
- 累積分布関数 (cumulative distribution function) $F(x) \stackrel{\text{def.}}{=} \Pr(X \leq x)$.

観察: $F(-\infty) = 0$, $F(+\infty) = 1$, F は単調非減少 (i.e., $(x \geq y \Rightarrow F(x) \geq F(y))$), F は右連続: $\lim_{\varepsilon \to +0} F(x + \varepsilon) = F(x)$. 左連続とは限らない.
※離散分布の場合, F は連続でない.
連続分布: Ω が非可算集合の場合.

- 確率密度関数 (probability distribution function) $f(x) \stackrel{\text{def.}}{=} \dfrac{\mathrm{d}}{\mathrm{d}x} F(x)$,
- 累積分布関数 (cumulative distribution function) $F(X) \stackrel{\text{def.}}{=} \Pr(X \leq x)$.

観察: 離散分布の場合の F の性質に加え, $F(x) = \int_{-\infty}^{x} f(z) \mathrm{d}z$ が成り立つ.
※連続分布の場合, F は微分可能(すなわち連続). f は F を通して確率空間と関連付けられることに注意.

3. 離散分布.

(i) 2 点分布 (ベルヌーイ分布; Bernoulli dist.) $\mathrm{B}(1;p)$ $(0<p<1)$

$$\Omega=\{0,1\},\quad f(k)=\begin{cases}1-p & (k=0),\\ p & (k=1).\end{cases}$$

2 点分布を生み出す試行のことを**ベルヌーイ試行** (*Bernoulli trial*) という.

(ii) 2 項分布 (binomial dist.) $\mathrm{B}(n;p)$ $(0<p<1)$

$$\Omega=\{0,1,2,\ldots,n\},\quad f(k)=\binom{n}{k}p^k(1-p)^{n-k}.$$

ベルヌーイ試行を n 回独立に行った際, "1" が k 回出現する確率.

(iii) 幾何分布 (geometric dist.) $\mathrm{Ge}(p)$ $(0<p<1)$

$$\Omega=\{0,1,2,\ldots\},\quad f(k)=(1-p)^k p.$$

ベルヌーイ試行を n 回行った際, "1" が初めて出現するまでに "0" の出現回数が k である確率.

(iv) ポアソン分布 (Poisson dist.) $\mathrm{Po}(\lambda)$ $(\lambda>0)$

$$\Omega=\{0,1,2,\ldots\},\quad f(k)=\frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}.$$

定数 λ が $\lambda=np$ を満たし, かつ n が十分大きいとき, 2 項分布 $B(n;p)$ はポアソン分布 $\mathrm{Po}(\lambda)$ に近似される. ⇒ 演習 5.

⇒ さまざまな**連続分布**は次回.

4. 独立同一分布

独立同一分布 (i.i.d.) 2 つの確率変数 X と Y が**独立同一分布に従う** (*i.i.d.*; independent and identically distributed) とは, X と Y が同一の分布に従い, かつ X と Y が独立であることをいう.

演習問題 (*は基礎的な問題.)

演習 1. 相互独立性. 確率変数 X は公正なさいころ A の出目を表し, 確率変数 Y は公正なさいころ B の出目を表し, もうひとつの確率変数を $Z=X+Y \pmod 6$ とする. X と Y の組, Y と Z の組, X と Z の組はいずれも独立である (すなわち pairwise independent) が, X,Y,Z は相互独立でないことを示せ.

***演習 2.** 定理 4, 5, 系 6 を示せ.

演習 3. いま赤と青の 2 枚のコインがあるものとする. 赤いコインの表の出る確率は $p_r=3/4$, 青いコインの表の出る確率は $p_b=1/2$ とする. 薄暗い部屋の中で 2 つのコインのうち 1 枚をランダムに選び, コイントスを行う. コインの表が出ているとき, トスしたコインが赤である確率はいくつか? ただしコインの表裏は触覚でわかるものとする.

***演習 4.** お菓子についているおまけのカードを集める. カードは全 12 種類で, 等しい確率で出現するものと仮定する.

(i)12 個お菓子を買う. ある種類のカード (鷹としよう) がちょうど x 枚出現する確率は? 特に $x=3$ については具体的な値を求めよ. 鷹の出現回数 X は何分布に従うか?

(ii) おかしをたくさん購入して全種類集めたい. いま 6 種類のカードをすでに集めたものとする. これから 7 種類目のカードが出てくるまでに購入するお菓子の個数を K で表す. $K=k$ となる確率は? 特に $k=4$ については具体的な値を求めよ. K は何分布にしたがうか?

演習 5. 2 項分布とポアソン分布. 以下を示せ;
定数 λ に対して, $\lambda=np$ が成り立つ n と p について,

$$\lim_{n\to+\infty}\binom{n}{z}p^z(1-p)^{n-z}=\mathrm{e}^{-\lambda}\frac{\lambda^z}{z!}$$

が成り立つ.