

統計学 I の課題問題（多少の変更、加筆をしています）

問題 1（問 2.1）2つの事象 A, B の確率が $P(A \cup B) = \frac{7}{8}$, $P(A \cap B) = \frac{1}{4}$, $P(\bar{A}) = \frac{5}{8}$, と与えられている。つぎの確率をもとめよ。(i) $P(A)$ (ii) $P(\bar{B})$ (iii) $P(A \cap \bar{B})$ (iv) $P(A \cup \bar{B})$

問題 2（問 2.2）事象 A, B は互いに排反（同時に起ることはない）事象で $P(A) = 0.2$, $P(B) = 0.8$ で、さらにある事象 C の条件つき確率が $P(C|A) = 0.4$ と $P(C|B) = 0.5$ で与えられている。 $P(A|C)$ と $P(B|C)$ をもとめよ。答えは分数のままでよい。

問題 3（問 2.3）全事象 $S = \{1, 2, 3, 4\}$ とし、それぞれ 4 個の点に対して、同じ確率 $\frac{1}{4}$ ずつをもつとする。3つの事象 $A = \{1, 2\}$, $B = \{1, 3\}$, $C = \{1, 4\}$ において、(i) $P(AB) = P(A)P(B)$, $P(AC) = P(A)P(C)$, $P(BC) = P(B)P(C)$ を示せ。(ii) しかし、3つが同時に起こる積事象の確率 $A \cap B \cap C = ABC$ はそれぞれの確率の積にはならない。すなわち

$$P(A \cap B \cap C) \neq P(A)P(B)P(C)$$

問題 4（問 2.4）確率変数 X の平均（期待値）が $E(X) = m$, 分散が $V(X) = s^2$ であるという。つぎの確率変数 (a) $Y = 3X + 2$, (b) $Z = X - m$, (c) $W = \frac{X+b}{a}$ （ここで a, b は定数）について、それぞれの平均と分散をもとめよ。

📖 📖 📖 **まとめと補足事項** 📖 📖 📖

確率 (probability) :

全事象 Ω (ラージ オメガ), 空事象 ϕ (フィー、ファイ) と表し、任意の事象 (event) A の確率 $0 \leq P(A) \leq 1$,
 和事象 (union) $A \cup B$, 積事象 (intersection) $A \cap B = AB$, 補事象 (余事象) (complement) $A^c = \bar{A}$,
 確率の加法性 (additive) : $P(\Omega) = 1, P(\phi) = 0$, 事象 A, B に対して

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

補事象との関係式: $P(\bar{A}) = 1 - P(A)$, ($\because A \cup \bar{A} = \Omega, A \cap \bar{A} = \phi$)

$$\text{条件付き確率: } P(A|B) = \frac{P(A \cap B)}{P(B)}$$

場合分けの確率計算 : 3 個の互いに素な事象 A_1, A_2, A_3 があるとき、「どれかが必ず起こり、同時には起こることはない」 $A_1 \cup A_2 \cup A_3 = \Omega, A_i \cap A_j = \phi (i \neq j)$ ならば、

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + P(A_3)P(B|A_3)$$

それぞれの場合に分けて、条件付き確率 $P(B|A_i), i = 1, 2, 3$ をしてからまとめる。いいかえると、条件付確率の重み付き平均値となっている。

ベイズの定理 ; (逆確率の計算式) 原因事象の確率

$P(A), P(B), P(C)$ が与えられ、結果の事象 E に対する条件付き確率 $P(E|A), P(E|B), P(E|C)$ が分かるならば、結果 E からもとの原因 A, B, C に向けての条件付き確率 $P(A|E), P(B|E), P(C|E)$ が計算できる。たとえば、

$$P(A|E) = \frac{P(A)P(E|A)}{P(A)P(E|A) + P(B)P(E|B) + P(C)P(E|C)}$$

他の場合 $P(B|E), P(C|E)$ も同様。

2 つの事象の独立 (independent): A と B が互いに独立とは

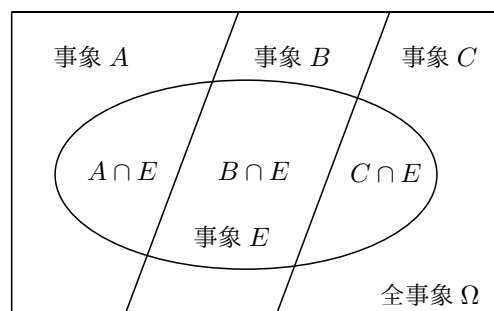
$$P(A \cap B) = P(AB) = P(A) \cdot P(B)$$

これと同値な命題: $P(\bar{A}B) = P(\bar{A}) \cdot P(B), P(A\bar{B}) = P(A) \cdot P(\bar{B}), P(\bar{A}\bar{B}) = P(\bar{A}) \cdot P(\bar{B})$. 独立であれば、事象の起こることと起こらないことが、お互いに影響しない。より一般に n 個の組が独立とは、 $A_1, A_2, \dots, A_n$ が、2 個ずつが”互いに独立”でさらにすべての積事象の確率もそれぞれの確率の積に等しくなるときをいう。

確率実験 (事象) の結果を全体の集まり (集合) Ω と表す。たとえば (i) 点集合 (離散的な点の集まり)、(ii) 数直線上の区間 (範囲) の集まりを (i) の場合では離散型、(ii) の場合では連続型とよぶ。離散的なときは、数列の和の計算で、連続的なときには、積分での面積計算により、確率の計算をおこなう。

確率変数 (random variable); $X(\omega) : \omega \in \Omega \rightarrow \text{数値}$

$$\begin{aligned} \{X = a\} &= \{\omega | X(\omega) = a\}, \\ \{X = a, X = b\} &= \{\omega | X(\omega) = a\} \cap \{\omega | X(\omega) = b\}, \\ \{X < b\} &= \{\omega | X(\omega) < b\}, \\ \{a \leq X < b\} &= \{\omega | a \leq X(\omega)\} \cap \{\omega | X(\omega) < b\} \quad \text{etc} \end{aligned}$$



$$\Omega = A \cup B \cup C$$

(i) X の取り得る値が n 個の場合: $P(X = x_k) = p_k, (k = 1, 2, \dots, n), 0 \leq p_k \leq 1, \sum_k p_k = 1$

確率分布 (probability distribution): r.v. X の確率分布とは F_X または X を省略して大文字で F と表し、

$$F_X(a) = P(X \leq a),$$

$$P(a < X \leq b) = P(\{X \leq b\} \cap \overline{\{X \leq a\}}) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a)$$

とくに和の記号で

$$\sum_{a < k \leq b} s_k = \{ \text{条件 } a < k \leq b \text{ を満たす } s_k \text{ の和をとる} \}$$

を意味するとする。

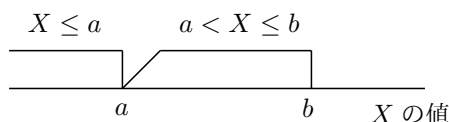
(ii) 連続型の場合では、区間での関数 (密度関数 f_X) の積分で表される。(小文字と大文字を使い分ける)

$$P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x) dx$$

このときには、積分の上限と下限が等しいときには $\int_a^a f_X(x) dx = 0$ となるから、 $P(X = a) = 0$ で

$$P(a < X \leq b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X < b) = F_X(b) - F_X(a)$$

が成り立ち、「離散型での等号がある場合とない場合には同じ確率値にはなるとは限らない」こととの違いに注意する。



確率変数 X の期待値 (expectation) を $E(X)$, 平均 (mean) $\mu_X = \mu$ ともいう、分散 (variance) を $V(X) = \sigma_X^2$ (2 乗をもちいる)。2 乗を付けなければ $\sigma(X) = \sqrt{\sigma_X^2} = \sigma$ を標準偏差とよぶ。記述統計でのデータの整理では、平均値とよんだ average をもちい、確率分布のばあいでは平均 mean をもちいる。

$$\text{期待値 (平均)} \quad \mu = E(X) = \begin{cases} \sum_k x_k P(X = x_k) = \sum_k x_k p_X(x_k) \\ \int_{-\infty}^{\infty} x f_X(x) dx \end{cases}$$

$$\text{分散} \quad \sigma^2 = V(X) = E[(X - \mu)^2] = E(X^2) - \mu^2$$

$$= \begin{cases} \sum_k (x_k - \mu)^2 P(X = x_k) = \sum_k x_k^2 p_X(x_k) - \mu^2 \\ \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x) dx = \int_{-\infty}^{\infty} x^2 f_X(x) dx - \mu^2 \end{cases}$$

同時分布, 結合分布 (joint distribution): 2 つ以上の確率変数を取り扱うときは、独立であれば取り扱いやすいが、一般には、同時に起こる事象を考えねばならない。すなわち、つぎのような積事象を考える。

$$F_{X,Y}(a,b) = P(X \leq a, Y \leq b) = P(\{X \leq a\} \cap \{Y \leq b\})$$

もし X, Y に関する 2 つの事象が独立であれば、確率の積に等しいから、

$$F_{X,Y}(a,b) = P(X \leq a, Y \leq b) = P(X \leq a) P(Y \leq b) = F_X(a) F_Y(b)$$

が成り立つ。これらを周辺 (marginal) 分布とよぶ。「周り」というよりは、他方の事象をすべてのばあい、全事象に尽くし出して、対象とするものだけに注視することです。たとえば $F_{X,Y}(a, \infty) = P(X \leq a, Y < \infty) = P(\{X \leq a\} \cap \Omega) = P(X \leq a) = F_X(a)$ 。

確率変数の変換 X, Y と定数について、 $Y = aX + b$ とすると、

1. $E(Y) = E(aX + b) = aE(X) + b$
2. $V(Y) = V(aX + b) = a^2V(X)$ (右辺は b によらない)
3. $\sigma(Y) = \sigma(aX + b) = |a|\sigma(X)$ (右辺は a の絶対値)

期待値の性質 X, Y と定数 a, b について、

1. $E(aX + bY) = aE(X) + bE(Y)$
2. もし X, Y が独立ならば、 X, Y に関する事象が独立とするから、 $E(XY) = E(X)E(Y)$ で積の期待値はそれぞれの積に等しい。 $V(aX + bY) = a^2V(X) + b^2V(Y)$
3. とくに $V(X - Y) = V(X) + V(Y)$ で、差の分散であっても、それぞれの和になる。

注意 独立でなくても、和の期待値は、それぞれの期待値になる。同一分布とは、同じ平均や分散をもつことであり、独立とは異なる。2つの確率変数 $E(X) = \mu_X$, $E(Y) = \mu_Y$ が独立であってもなくても、 $E(X + Y) = \mu_X + \mu_Y$ であり、和 $X + Y$ の分布を計算せずに期待値は計算できる。もし同一分布にしたがう、すなわち $\mu_X = \mu_Y = \mu$ であれば、 $E(X + Y) = 2\mu$ となる。しかし和の分散をするには、共分散 (covariance)

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X \mu_Y$$

を計算しなければならない。独立であれば、共分散がゼロとなるが、逆の命題が不成立、「共分散がゼロであっても、独立とは限らない」に注意する。