

第1回レポート課題

演習問題1～10の中から3題以上選んで解く。

演習問題 1. (X, Y) は2次元正規分布にしたがい、その期待値と分散共分散行列は以下で与えられる。

$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 2 & \rho \\ \rho & 1 \end{pmatrix}.$$

1. Σ が正定値行列であることから、 ρ が満たすべき(必要十分)条件を求める。
2. 条件付き確率密度関数 $p(y|x)$ の期待値と分散共分散行列を計算する。

演習問題 2. 確率変数 X の密度関数 $p(x)$ は、 a を定数として

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

で定義されている。

1. 定数 a を求める.
2. X の期待値と分散を求める.

演習問題 3. 確率の公理を参考にして以下の問に答える.

1. 加法定理を証明する.
2. 可算個の集合 A_1, A_2, \dots に対して $\Pr(\cup_i A_i) \leq \sum_i \Pr(A_i)$ を示す.

演習問題 4. X は $X \sim N(0, 1)$ であるような確率変数とする.

密度関数は $p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ である.

1. $Y = e^X$ の確率密度関数を求める.
2. Y の期待値を求める.

演習問題 5. $X_1, \dots, X_n \sim_{i.i.d.} P$, $\mu = E[X_i]$, $\sigma^2 = V[X_i]$ のとき, 以下が成立することを確認する.

$$Y = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{に対して} \quad E[Y] = \mu, \quad V[Y] = \frac{\sigma^2}{n}.$$

演習問題 6. 事象 B の条件の下で事象 A が起こる条件付き確率を $\Pr(A|B)$ と表す. 確率変数 X, Y は 0 または 1 を取るとし,

$$\Pr(Y = 1|X = 1) = p, \quad \Pr(Y = 0|X = 0) = q, \quad \Pr(X = 1) = r.$$

とおく. 以下の確率を p, q, r を用いて表す.

1. $\Pr(Y = 0|X = 1)$
2. $\Pr(Y = 1)$
3. $\Pr(X = 1|Y = 0)$

演習問題 7. 正の値をとるデータが独立に

$$X_1, \dots, X_n \sim_{i.i.d.} P$$

のように観測された. 統計モデルとしてパラメータ $\theta > 0$ をもつ密度関数

$$p(x; \theta) = \theta \exp(-\theta x), \quad x > 0$$

を仮定する. このとき最尤推定量 $\hat{\theta}$ を求める.

演習問題 8. 条件付き期待値 $E[X|y]$ の定義は以下で与えられる：

$$E[X|y] = E[X|Y = y] = \int xp(x|y)dx.$$

条件付き分散 $V[X|y]$ の定義も同様：

$$V[X|y] = \int (x - E[X|y])^2 p(x|y)dx$$

このとき、以下の式が成り立つことを示す。

$$E[V(X|Y)] = V(X) - \int (E[X|y] - E[X])^2 p(y)dy.$$

ここで $E[V(X|Y)] = \int V[X|y]p(y)dy.$

演習問題 9. X, Y はともに $\{-1, 0, 1\}$ に値をとる確率変数として, $\Pr(X = x, Y = y)$ を表のように定める. 以下の問に答えよ.

$y \setminus x$	-1	0	1
-1	0	$1/3$	0
0	$1/4$	0	$1/4$
1	0	$1/6$	0

1. X と Y は無相関か?
2. X と Y は独立か?

演習問題 10. 血液型とその遺伝子型の確率は次のように定められる.

● 血液型 : **A, B, AB, O.**

● 対立遺伝子 : **a, b, o.**

a,b,o の確率 : $\theta_a, \theta_b, \theta_o, \theta_a + \theta_b + \theta_o = 1, \theta_a, \theta_b, \theta_o > 0.$

$$\Pr(A) = \theta_a^2 + 2\theta_a\theta_o, \quad \Pr(B) = \theta_b^2 + 2\theta_b\theta_o$$

$$\Pr(AB) = 2\theta_a\theta_b, \quad \Pr(O) = \theta_o^2$$

表現型	遺伝子型	人数
A	aa, ao, oa	n_A
B	bb, bo, ob	n_B
AB	ab, ba	n_{AB}
O	oo	n_O

上の表にある n_A, n_B, n_{AB}, n_O を適当な値に設定して, $\theta_a, \theta_b, \theta_o$ の最尤推定値を数値計算によって求める. プログラミング言語は何を使ってもよい.

(C, R, python, matlab, mathematica など)