2020.4.21(火)

情報統計力学 事前学習3

4月28日 (火) 午後１時ー２時半

マンモグラフィーによる乳がんの検診　の続き

前回、マンモグラフィーによる検診で陽性と判定されたときに、実際に乳がんである確率を計算し、$P(\text{乳がん}\mid\text{陽性}) \simeq 0.09$、つまり、約９％となることを示した。では、もう一度、同じ検査を受けて陽性と判定された場合に、実際に乳がんである確率はいくらになるだろうか？
以下、最初の検査で乳がんであると判定されたことを前提として計算する。検査の精度は同じであるので、$P(\text{陽性}\mid\text{乳がん}) = 0.9$、$P(\text{陽性}\mid\text{乳がんでない}) = 0.07$ である。乳がんである確率を、あらためて $P(\text{乳がん})$ と書くと、これは、前回求めた値、約９％である。従って、$P(\text{乳がん}) = 0.09$ とする。前回と同様に陽性である確率 $P(\text{陽性})$ を求めよう。

$$P(\text{乳がんでない}) = 1 - P(\text{乳がん}) = 1 - 0.09 = 0.91 \tag{1}$$

であり、前回と同様にして

$$\begin{aligned} P(\text{陽性}) &= P(\text{陽性}\mid\text{乳がん})P(\text{乳がん}) + P(\text{陽性}\mid\text{乳がんでない})P(\text{乳がんでない}) \\ &= 0.9 \times 0.09 + 0.07 \times 0.91 = 0.081 + 0.0637 = 0.1447 \simeq 0.14. \end{aligned} \tag{2}$$

ベイズの定理により

$$P(\text{乳がん}\mid\text{陽性}) = \frac{P(\text{陽性}\mid\text{乳がん})P(\text{乳がん})}{P(\text{陽性})} = 0.9 \times 0.09/0.14 = 0.578.. \tag{3}$$

よって、検診で２回とも乳がんと判定されたとき、実際に乳がんである確率は 58% 程度で、かなり大きくなることが分かる！

問題

(1) 検診２回で乳がんと判定されたとき、乳がんでない確率を求めよ。
(2) 検診で３回とも乳がんと判定されたとき、乳がんである確率を求めよ。

独立性

事象 $A, B (\in \mathcal{F})$ に対し、

$$P(A \cap B) = P(A)P(B) \tag{4}$$

となるとき、事象 A と B は**互いに独立である**といい、$A \perp\!\!\!\perp B$ と表す。

事象が３つ有る場合の独立性は以下のように定義する。

事象 $A, B, C (\in \mathcal{F})$ に対し、

$$\begin{aligned} &P(A \cap B) = P(A)P(B),\ P(A \cap C) = P(A)P(C),\ P(B \cap C) = P(B)P(C), \\ &P(A \cap B \cap C) = P(A)P(B)P(C) \end{aligned} \tag{5}$$

となるとき、事象 A, B, C は**互いに独立である**という。

より一般には、次のように定義する。

事象 $A_1, A_2, \cdots, A_n (\in \mathcal{F})$ の任意の部分列 $A_{i_1}, A_{i_2}, \cdots, A_{i_k} (2 \leq k \leq n)$ が、

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}) \tag{6}$$

となるとき、事象 $A_1, A_2, \cdots, A_n$ は**互いに独立である**という。

３つの事象の独立性について、$P(A \cap B \cap C) = P(A)P(B)P(C)$ が成り立つだけで良いような気がするだろう。しかし、$A \perp\!\!\!\perp B,\ A \perp\!\!\!\perp C,\ B \perp\!\!\!\perp C$ であっても、必ずしも $P(A \cap B \cap C) = P(A)P(B)P(C)$ とはならない。逆に、$P(A \cap B \cap C) = P(A)P(B)P(C)$ でも $A \perp\!\!\!\perp B,\ A \perp\!\!\!\perp C,\ B \perp\!\!\!\perp C$ は、必ずしも成り立たない。以下の問題を解いて、確認せよ。

問題

$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ とし、$\mathcal{F}$ を Ω の全ての部分集合とする。また、$P(\{\omega_k\}) = \frac{1}{4}\ (k = 1, 2, 3, 4)$ とする。

$A = \{\omega_1, \omega_2\},\ B = \{\omega_2, \omega_3\},\ C = \{\omega_3, \omega_1\}$ として、以下を示せ。

(1) $P(A \cap B) = P(\{\omega_2\}) = \frac{1}{4}$.
(2) $A \perp\!\!\!\perp B,\ A \perp\!\!\!\perp C,\ B \perp\!\!\!\perp C$.
(3) $P(A \cap B \cap C) = 0$.