

第三回ベイズ統計学・機械学習研究会

～条件付確率とベイズの定理～

もくじ

- ベイズの定理とは
- 条件付確率の復習
- ベイズの定理を使ってみよう

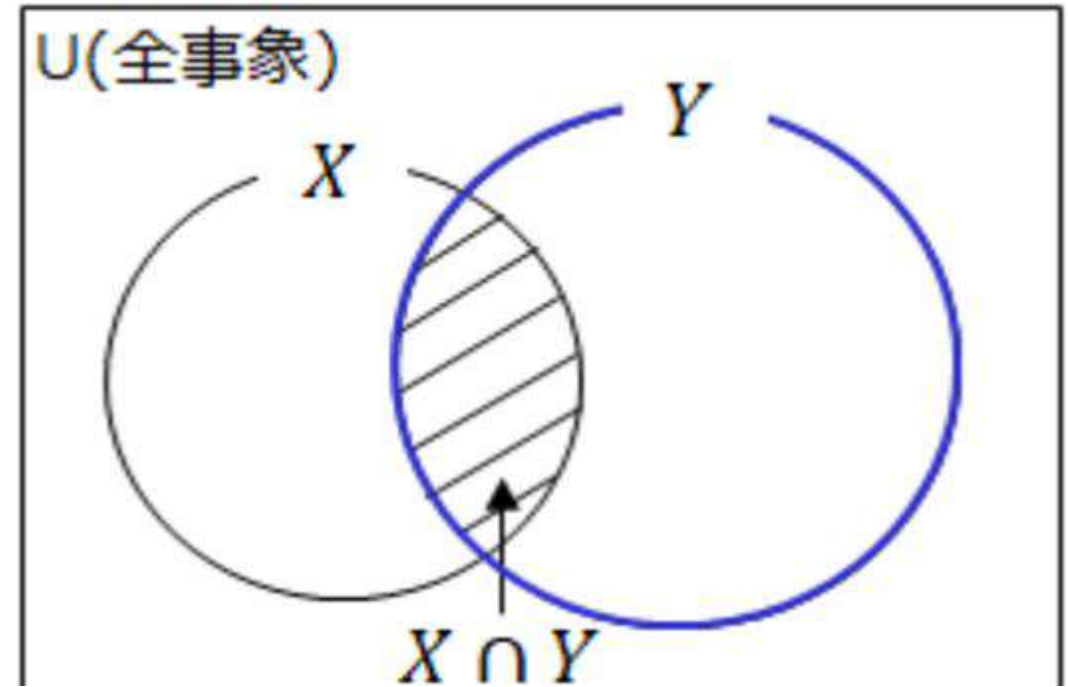
ベイズの定理とは

- 1740年ごろトーマス・ベイズが発見した条件付き確率に関して成り立つ定理
- $$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$
- ベイズ統計学はこの定理を応用して予測を行う

条件付確率の復習

- 条件付確率とはある事象が起こるという条件のもとで、別のある事象が起こる確率のこと
- $P(B|A)$ のような形で表される

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



例題

- XとYという二つの袋を用意する。Xには赤玉が3つ白玉が5つ、Yには赤玉が1つ白玉が3つ入っている。二分の一の確率でXかYを選び、その中の玉を一つとる。
- Xを選んだとして、赤玉を取る確率は？

- Xを選んだとして、赤玉を取る確率は？

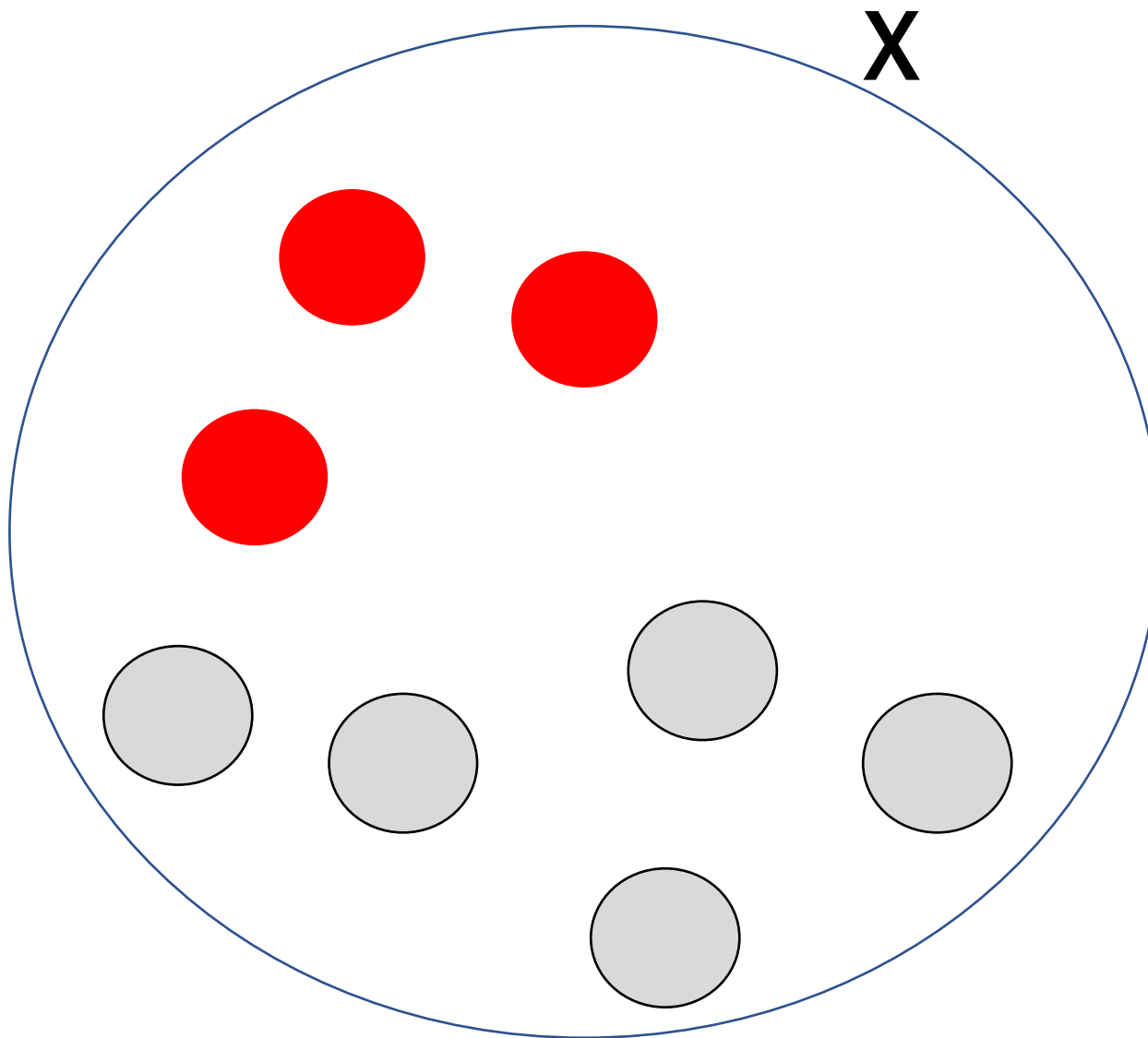
条件付確率

$$P(\text{赤玉}|X) = \frac{3}{8}$$

$$P(\text{赤玉}|X) = \frac{P(\text{赤玉} \cap X)}{P(X)}$$

$$= \frac{\frac{3}{16}}{\frac{1}{2}}$$

$$= \frac{3}{8}$$



練習問題

- 1から3の目が赤色で塗られており、4から6の目は青色で塗られているさいころがある。今、このさいころを投げて青色の目が出た時、この目が偶数である確率を求めよ。(統計学の時間 | 統計WEBより)

答え

- 青が出たという前提で、偶数の目が出る確率を問われているので条件付確率の問題です
- $P(B|A) = \frac{P(A \cap B)}{P(A)}$ より
- $P(A) = \frac{3}{6}$ (青が出る確率)
- $P(A \cap B) = \frac{2}{6}$ (青でかつ、偶数の目が出る確率)
- $P(B|A) = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{2}{3}$

ベイズの定理例題

- XとYという二つの袋を用意する。Xには赤玉が3つ白玉が5つ、Yには赤玉が1つ白玉が3つ入っている。二分の一の確率でXかYを選び、その中の玉を一つとる。
- 赤玉を取った時、それがXの袋に入っていたものである確率は？

続き

- 時間の流れが逆の確率なのでベイズの定理を使います

- $P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$ に値を一つずつ入れていく

続き

$$\bullet P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

- $P(\theta)$ はXの袋を選ぶ確率
- $P(D)$ は赤玉を選ぶ確率
- $P(D|\theta)$ はXの袋を選んだ時の、赤玉を選ぶ条件付確率なので

$$P(\theta) = \frac{1}{2}$$

$$P(D) = \left(\frac{1}{2} \times \frac{3}{8}\right) + \left(\frac{1}{2} \times \frac{1}{4}\right)$$

$$= \frac{3}{16} + \frac{1}{8}$$

$$= \frac{5}{16}$$

$$P(D|\theta) = \frac{3}{8}$$

$$P(\theta|D) = \frac{\frac{3}{8} \times \frac{1}{2}}{\frac{5}{16}} = \frac{3}{5}$$

練習問題

- ある映画の試写会を行い、満足度のアンケート調査を行った。試写会に参加したのは300人でそのうち女性が180人であり、満足したと回答したのは男性の50%、女性の75%であった。この映画を見て満足しなかったと答えた人が女性である確率はいくらか。(統計学の時間 | 統計WEBより)

- $$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

- 男性が満足していない確率は $P(D1) = 1 - 0.5 = 0.5$
- 女性が満足していない確率は $P(D2) = 1 - 0.75 = 0.25$
- 女性であるという前提で、満足していない確率は

$$\begin{aligned} P(D|\theta) &= \frac{\text{女性でかつ満足していない確率}}{\text{女性である確率}} \\ &= \frac{0.6 \times 0.25}{0.6} \\ &= 0.25 \end{aligned}$$

答え

- 映画に満足できなかったという前提で、その人が女性である確率を問われているので時間の流れが逆の条件付確率の問題です

$$\bullet P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

- $P(\theta)$ は女性である確率
- $P(D)$ は映画に満足できない確率
- $P(D|\theta)$ は女性であるという前提で、映画に満足していない確率

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \text{より}$$

$$P(\theta) = \frac{180}{300} = \frac{3}{5} = 0.6$$

$$\begin{aligned} P(D) &= (0.4 \times 0.5) + (0.6 \times 0.25) \\ &= 0.2 + 0.15 \\ &= 0.35 \end{aligned}$$

$$P(D|\theta) = 0.25$$

$$P(\theta|D) = \frac{0.25 \times 0.6}{0.35} = 0.429$$