# 法と統計学　第 4 回
# 確率の和の法則と積の法則 (3)
# 条件付確率

森　大輔

熊本大学法学部

2017 年

# 本日の流れ

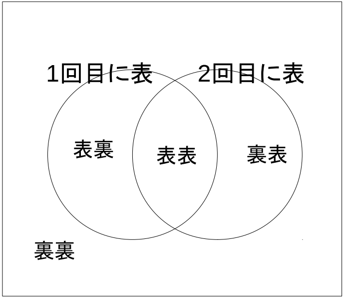
# 排反事象でない場合

## 排反事象でない場合

$A_1$ と $A_2$ が排反事象でないならば、
$\mathrm{P}(A_1 \text{または} A_2) = \mathrm{P}(A_1) + \mathrm{P}(A_2) - \mathrm{P}(A_1 \text{かつ} A_2)$

- 「$A_1$ または $A_2$」は、$A_1$ と $A_2$ のどちらか一方か、両方起こること
- 2 つの事象が排反事象でない場合、同時に起こりうる
- 単純に足し算しただけだと、2 つ同時に起こることを 2 回考えてしまうので、引き算する
- 例：コインを 2 回投げ、少なくとも 1 回表が出る確率
  - 1 回目に表が出る ($A_1$) と 2 回目に表が出る ($A_2$) は、排反事象ではない
  - $A_1$ には表表と表裏、$A_2$ には表表と裏表が入る

# ベン図




- 1 回目に表 ($A_1$) と 2 回目に表 ($A_2$) を単純に足し算すると、表表 ($A_1$ かつ $A_2$) が 2 回入るので、1 回引き算
- 2 回コインを投げた時の結果全部 (表表、表裏、裏表、裏裏) 中、1 回目に表が出るのは 2 通りなので $\mathrm{P}(A_1) = 2/4 = 1/2$
- 同様に考えて $\mathrm{P}(A_2) = 1/2$、$\mathrm{P}(A_1 かつ A_2) = 1/4$
- よって $\mathrm{P}(A_1 または A_2) = 1/2 + 1/2 - 1/4 = 3/4$

# ボンフェローニの不等式

- 以上は事象が 2 つの場合
- 3 つ以上の事象があり、それらが排反事象でない場合も同様に考えようとすると、かなり複雑
- そこで、だいたいの確率の大きさを知るための簡便法として、ボンフェローニの不等式が有用なことがある

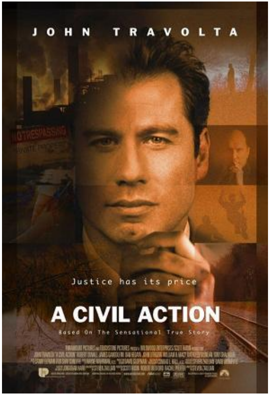
## ボンフェローニの不等式

$\mathrm{P}(A_1$ または $A_2$ または $\cdots$ または $A_n) \leq \mathrm{P}(A_1) + \mathrm{P}(A_2) + \cdots + \mathrm{P}(A_3)$

等号成立は $A_1$ から $A_n$ までがすべて排反事象のとき

- 例 : サイコロを 3 回投げ、少なくとも 1 回は 1 が出る確率
  - ボンフェローニの公式から、少なくとも $1/6 + 1/6 + 1/6 = 1/2$ より小さい確率だとわかる
    - 因みに正確には $1 - (5/6)^3 \fallingdotseq 0.42$
    - 1 から 3 回とも 1 でない確率を引く (積の法則使用)

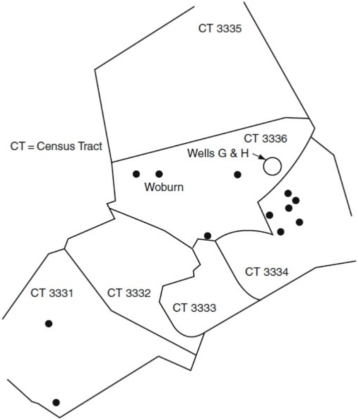
# 適用例：水道汚染訴訟




https://en.wikipedia.org/wiki/A_Civil_Action_(film) より

- ジョナサン・ハー (2000)『シビル・アクション：ある水道汚染訴訟 (上)(下)』新潮社.
  - ノンフィクション小説
- 映画「シビル・アクション」
  - 上記の小説をもとに制作、主演ジョン・トラヴォルタ
- マサチューセッツ州のウォーバンで小児白血病が集団発生
  - 原告の弁護士シュリクトマンが、被告グレース社とベアトリス・フーズ社の弁護士と戦う

# 適用例：水道汚染訴訟 (図は Finkelstein (2009:45) より)




- 1969 年から 1979 年の間に、ウォーバンでは小児白血病の発生が 12 件 (図の黒丸)
  - 全米の発生率に基づく予測では 5.3 件しか発生しないはず
- 図の G と H という井戸が汚染されていることが明らかに
- 原因は近くのゴミ捨て場の産業廃棄物
- グレース社とベアトリス・フーズ社が廃棄物の投棄に関与の疑い

# 適用例：水道汚染訴訟

- ウォーバンは、6 つの地区に分かれている
  - 図に CT3336 などと書かれているのが地区名
- CT3334 で 6 件の集団発生
  - 井戸に近い地区
- 今、井戸水が原因ではなく白血病はどの地区でも同じ確率で発生すると仮定
  - すると、白血病が 1 件中 1 件、ある特定の 1 つの地区で発生する確率は $1/6$
    - 地区は 6 つあるから
  - この仮定が正しいとしたときに、白血病が 12 件中 6 件以上、6 つの地区のうち少なくとも 1 つの地区で発生する確率を求める
  - この確率が異常に小さいとき、仮定を棄て去る (つまり井戸水が原因)
    - 前々回のハウランド遺言事件と同じ仮説検定の論理

# 練習問題 1：水道汚染訴訟

白血病が 1 件中 1 件、ある特定の 1 つの地区で発生する確率は 1/6 である。
(1) 白血病が 12 件中 6 件、ある特定の 1 つの地区で発生する確率を求めなさい。
(2) 白血病が 12 件中 6 件以上、ある特定の 1 つの地区で発生する確率を求める計算式を書きなさい。
(3) 白血病が 12 件中 6 件以上、6 つの地区のうち少なくとも 1 つの地区で発生する確率は 0.05 より大きいか小さいか、ボンフェローニの不等式を用いて判断しなさい。

# 適用例：水道汚染訴訟

- 白血病が 12 件中 6 件以上、ある特定の 1 つの地区で発生する確率は 0.0079
  - 白血病が 12 件中 12 件、ある特定の 1 つの地区で発生する確率 $(1/6)^{12}$
  - 白血病が 12 件中 11 件、ある特定の 1 つの地区で発生する確率 ${}_{12}C_{11} \times (1/6)^{11} \times (5/6)$
    - 前回の反復試行の確率
  - ……
  - 白血病が 12 件中 6 件、ある特定の 1 つの地区で発生する確率 ${}_{12}C_6 \times (1/6)^6 \times (5/6)^6$
  - これらを足すと 0.0079

# 適用例：水道汚染訴訟

- 白血病が 12 件中 6 件以上、6 つの地区のうち少なくとも 1 つの地区で発生する確率
  - ボンフェローニの不等式から、$0.0079 \times 6 = 0.047$ よりは小さいことがわかる
- 0.047 は「異常に小さい」か？
- 判断する基準が必要
  - この基準を有意水準と呼ぶ（第 12 回目ぐらいに詳細）
  - よく使われる基準が、0.05 より小さければ「異常に小さい」と判断するというもの
- 0.047 は 0.05 よりも小さいので、異常に小さいと判断
- 白血病はどの地区でも同じ確率で発生すると仮定して、白血病が 12 件中 6 件以上、6 つの地区のうち少なくとも 1 つの地区で発生する確率は異常に小さい
- よって、井戸水が原因で集団発生したと考えた方が適当

# 条件付確率の定義

- 事象 $A$ が起きたことを前提とした上で事象 $B$ が起きる確率のことを、条件付確率という
  - $\mathrm{P}(B|A)$ と書く
    - $\mathrm{P}_A(B)$ とも書く

**条件付確率の定義 1**

$$\mathrm{P}(B|A) = \frac{\mathrm{n}(A \text{ かつ } B)}{\mathrm{n}(A)}$$

$\mathrm{n}(A)$ は事象 $A$ が起こる場合の数を表すとする

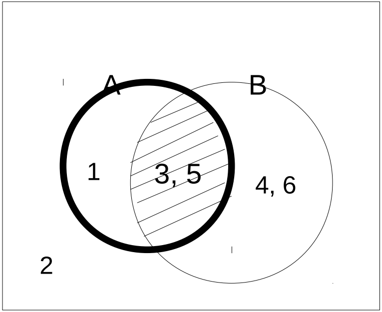
**条件付確率の定義 2**

$$\mathrm{P}(B|A) = \frac{\mathrm{P}(A \text{ かつ } B)}{\mathrm{P}(A)}$$

# 例題 1

- サイコロを 1 回振って出た目を記録しようと思ったが、目の数字を忘れてしまった。奇数であったことは覚えている。このとき出た目が 3 以上である確率は？
  - 事象 $A$ は「奇数の目が出た」、事象 $B$ は「3 以上の目が出た」
  - 条件付確率の定義 1 での解き方
    - 事象 $A$ は 1, 3, 5 の 3 通り、事象 $A$ かつ $B$ は 3,5 の 2 通り
    - よって答えは $\mathrm{P}(B|A) = 2/3$
  - 条件付確率の定義 2 での解き方
    - $\mathrm{P}(A \text{ かつ } B) = 2/6 = 1/3$、$\mathrm{P}(A) = 3/6 = 1/2$
    - よって答えは $\mathrm{P}(B|A) = \frac{1/3}{1/2} = 2/3$

# ベン図




- 条件付確率 $\mathrm{P}(B|A)$
  - 事象 $A$ が起きたことを前提とした上で事象 $B$ が起きる確率
  - 図の太丸 (事象 $A$) の中で、斜線部分 ($A$ かつ $B$) の占める割合

# 条件付確率の練習問題

## 条件付確率の練習問題 1-1

ある夫婦には子供が二人いる。二人のうち少なくとも一人は女の子であるということが分かった。このとき、二人とも女の子である確率はいくらか？ただし、男の子が生まれる確率、女の子が生まれる確率は半々だとする。

## 条件付確率の練習問題 1-2

ある夫婦には子供が二人いる。二人のうち年上の方は女の子であるということが分かった。このとき、二人とも女の子である確率はいくらか？ただし、男の子が生まれる確率，女の子が生まれる確率は半々だとする。

# 条件付確率の練習問題 1-1

- 事象 $A$ は「少なくとも一人は女の子」、事象 $B$ は「二人とも女の子」
- 条件付確率の定義 1 での解き方
  - 事象 $A$ は女女、女男、男女の 3 通り、事象 $A$ かつ $B$ は女女の 1 通り
  - よって答えは $\mathrm{P}(B|A) = 1/3$
- 条件付確率の定義 2 での解き方
  - $\mathrm{P}(A$ かつ $B) = 1/4$、$\mathrm{P}(A) = 3/4$
  - よって答えは $\mathrm{P}(B|A) = \frac{1/4}{3/4} = 1/3$

# 条件付確率の練習問題 1-2

- 事象 $A$ は「年上の方は女の子」、事象 $B$ は「二人とも女の子」
- 条件付確率の定義 1 での解き方
  - 事象 $A$ は女女、女男の 2 通り、事象 $A$ かつ $B$ は女女の 1 通り
  - よって答えは $\mathrm{P}(B|A) = 1/2$
- 条件付確率の定義 2 での解き方
  - $\mathrm{P}(A \text{ かつ } B) = 1/4$、$\mathrm{P}(A) = 2/4 = 1/2$
  - よって答えは $\mathrm{P}(B|A) = \frac{1/4}{1/2} = 1/2$

# 条件付確率の積の法則

## 条件付確率の積の法則

事象 $A$ と $B$ が独立でないとき、
$\mathrm{P}(A \text{ かつ } B) = \mathrm{P}(A) \times \mathrm{P}(B|A)$

- 求め方
  - 条件付確率の定義から $\mathrm{P}(B|A) = \frac{\mathrm{P}(A \text{ かつ } B)}{\mathrm{P}(A)}$
  - この両辺に $\mathrm{P}(A)$ を掛ければ求まる
- 数字 1 のカードが 2 枚、数字 2 のカードが 4 枚ある。計 6 枚のカードから 1 枚のカードを引き、カードを戻さず 2 枚目のカードを引く。1 枚目に数字 1 のカード、2 枚目に数字 2 のカードを引く確率は？
  - 事象 $A$ は「1 枚目に数字 1 のカードを引く」、事象 $B$ は「2 枚目に数字 2 のカードを引く」
  - $\mathrm{P}(A \text{ かつ } B) = \mathrm{P}(A) \times \mathrm{P}(B|A) = 2/6 \times 4/5 = 4/15$

# 独立と確率の積の法則

## 独立と条件付確率

事象 $A$ と $B$ が独立のとき、$\mathrm{P}(A|B) = \mathrm{P}(A)$ で、$\mathrm{P}(B|A) = \mathrm{P}(B)$
逆に $\mathrm{P}(A|B) = \mathrm{P}(A)$ で、$\mathrm{P}(B|A) = \mathrm{P}(B)$ のとき、事象 $A$ と $B$ は独立

- 独立ならば、$A$ の起こる確率に $B$ は影響を与えないし、$B$ の起こる確率に $A$ は影響を与えないということ
- さらに、これを条件付確率の積の法則に代入すると次のことも言える
    - 事象 $A$ と $B$ が独立のとき、$\mathrm{P}(A \text{ かつ } B) = \mathrm{P}(A) \times \mathrm{P}(B)$
    - これが前回までで見た確率の積の法則だった

# ベイズの定理

https://en.wikipedia.org/wiki/Thomas_Bayes より

## ベイズの定理

$$\mathrm{P}(B|A) = \frac{\mathrm{P}(B)\times\mathrm{P}(A|B)}{\mathrm{P}(B)\times\mathrm{P}(A|B)+\{1-\mathrm{P}(B)\}\times\mathrm{P}(A|B \text{ でない})}$$

# ベイズの定理

- $\mathrm{P}(B|A)$ を、$\mathrm{P}(A|B)$ や $\mathrm{P}(A|B$ でない$)$ などから求めるのがベイズの定理
  - $\mathrm{P}(B|A)$ の値はわからないが、$\mathrm{P}(A|B)$ や $\mathrm{P}(A|B$ でない$)$ の値はわかる場合がある
- ベイズの定理やそれを基礎にしたベイズ統計学の応用範囲は広い
  - ビル・ゲイツの言葉「21 世紀のマイクロソフトの基本戦略はベイズテクノロジーだ」
- しかし、理解が難しい部分もある
  - 人間の直感と違った答えが出ることがしばしばあるため
  - 自然頻度や樹形図の補助で解く方法を後に紹介

# ベイズの定理の求め方

- 条件付確率の定義
  - $\mathrm{P}(B|A) = \frac{\mathrm{P}(A \text{ かつ } B)}{\mathrm{P}(A)}$
- ここで確率の和の法則より
  - $\mathrm{P}(A) = \mathrm{P}(A \text{ かつ } B) + \mathrm{P}(A \text{ かつ } B \text{ でない})$
- これを条件付確率の式の分母に代入すると
  - $\mathrm{P}(B|A) = \frac{\mathrm{P}(A \text{ かつ } B)}{\mathrm{P}(A \text{ かつ } B) + \mathrm{P}(A \text{ かつ } B \text{ でない})}$
- 条件付確率の積の法則より
  - $\mathrm{P}(A \text{ かつ } B) = \mathrm{P}(B) \times \mathrm{P}(A|B)$
  - $\mathrm{P}(A \text{ かつ } B \text{ でない}) = \mathrm{P}(B \text{ でない}) \times \mathrm{P}(A|B \text{ でない}) = \{1 - \mathrm{P}(B)\} \times \mathrm{P}(A|B \text{ でない})$
- これらを分子と分母に代入して
  - $\mathrm{P}(B|A) = \frac{\mathrm{P}(B) \times \mathrm{P}(A|B)}{\mathrm{P}(B) \times \mathrm{P}(A|B) + \{1 - \mathrm{P}(B)\} \times \mathrm{P}(A|B \text{ でない})}$

# 例題 2：乳がん検査 (ベイズの定理)

　乳がんの早期発見のため、特定の年齢に達した女性は自覚症状がなくても定期的に検査を受けることが推奨されている。あなたが医師として、ある地域でこのような乳がんの検査を実施するとする。この地域で乳がんの検査に参加する 40 歳から 50 歳までの自覚症状のない女性について、以下のことがわかっている。

　これらの女性の 1 人が乳がんである確率は 0.8%である。また乳がんであれば、検査結果が陽性になる確率は 90%である。乳がんでなくても、陽性と出る確率は 5%ある。ある女性の検査結果が陽性と出た。

　この女性が実際に乳がんである確率はどれくらいか？

# 例題 2 の解き方

- $\mathrm{P}(\text{乳がん}) = 0.008$
  - 「これらの女性の 1 人が乳がんである確率は 0.8%」
- $\mathrm{P}(\text{陽性} \mid \text{乳がん}) = 0.90$
  - 「乳がんであれば、検査結果が陽性になる確率は 90%」
- $\mathrm{P}(\text{陽性} \mid \text{乳がんでない}) = 0.05$
  - 「乳がんでなくても、陽性と出る確率は 5%あります」
- $\mathrm{P}(\text{乳がん} \mid \text{陽性}) = \frac{\mathrm{P}(\text{乳がん}) \times \mathrm{P}(\text{陽性} \mid \text{乳がん})}{\mathrm{P}(\text{乳がん}) \times \mathrm{P}(\text{陽性} \mid \text{乳がん}) + \mathrm{P}(\text{乳がんでない}) \times \mathrm{P}(\text{陽性} \mid \text{乳がんでない})} = \frac{0.008 \times 0.90}{0.008 \times 0.90 + 0.992 \times 0.05} \fallingdotseq 0.13$
  - 「ある女性の検査結果が陽性と出ました。この女性が実際に乳がんである確率」

# 参考文献


- マイクル O. フィンケルスタイン (2014)『法統計学入門：法律家のための確率統計の初歩』木鐸社の第 2 章 pp.68–69、第 3 章 pp.70–77.
- Finkelstein, M. O. (2009) *Basic Concepts of Probability and Statistics in the Law*, Springer（上の原著）.
- Finkelstein, M. O. & B. Levin (2001) *Statisticf for Lawyers* (2nd ed.), Springer の第 1 章、第 2 章、第 3 章.
- ゲルト・ギーゲレンツァー (2010)『リスク・リテラシーが身につく統計的思考法：初歩からベイズ推定まで』早川書房.
- 草野耕一 (2014)「『検察官の誤謬』と『弁護人の誤謬』—ベイズの定理の実践的活用法」論究ジュリスト 10、20–29.
- 市川伸一 (1997)『考えることの科学』中公新書.