ベイズの基礎：ベイズの定理を導いて、その解釈を考えてみよう

ファイザーR&D合同会社
統計リサーチ・データサイエンス グループ
小宮山 靖

ベイズの定理

Pr(B|A)は事象Aが起きた上で事象Bが起こる条件付き確率。
英語では、”Probability of B given by A”と読む。

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)}$$

Pr(A)は事象Aが起きる確率

Pr(B)は事象Bが起きる確率

Pr(A|B)は事象Bが起きた上で事象Aが起こる確率で“条件付き確率”という。
英語では、”Probability of A given by B”と読む。

● ベイズの定理の２つの解釈

[解釈1]条件付き確率Pr(A|B) と 条件付き確率Pr(B|A) の間の関係を示している

[解釈2]エビデンスを考慮して主観的な信念の度合いが、どのように合理的に変化するはずかを表現している

今の時点でわかっていただかなくてOK。
最後に戻ってきます！

ベイズ統計の特徴は
不確実性(Uncertainty)を
確率で表現するところです

なので、「2つ以上の確率を
一緒に考える場合」を考えて
みましょう

ベイズの定理につながる確率の重要な話です

「偶数の目が出る」確率、「１または２の目が出る」確率…

- 「偶数の目が出る」確率は、1/6＋1/6＋1/6 =1/2
- 「１または２の目が出る」確率は、1/6+1/6＝1/3

…などと計算できます。

この単純な計算ができる背景には、サイコロの例の重要な性質があります。

Mutually Exclusive and Collectively Exhaustive （MECE; メーシー）

Mutually Exclusive → **互いに排反**

どの結果も同時に起こることはない

Collectively Exhaustive → **集合的に網羅されている**

すべての結果が考えられている

1つ目の確率：利き目は右目？左目？

あなたの利き目は右目か左目かご存じですか？

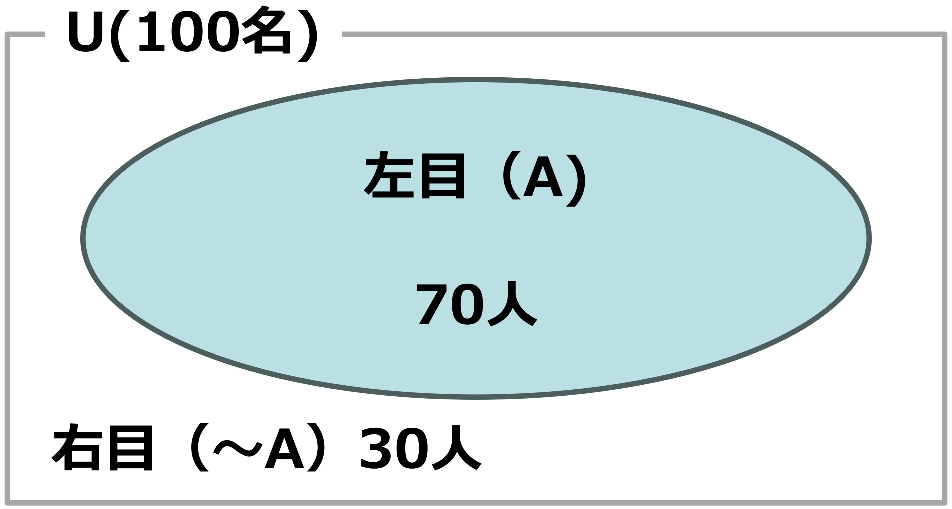
両手を合わせて三角形の輪を作り、
遠くの物をその輪を通して見ます。
その後、片目ずつ閉じ、対象物が
ずれない方（両眼で見た時と同じ方）
が利き目です。
【三角形法/ローゼンバッハ法】

利き目がどちらかは、眼鏡を作る時や、スポーツ（特に球技）で重要らしいです

100人に聞きました

事象(Event)	頻度（Frequency）
左目（A）	70
右目（～A）	30
全体(U)	100

これを集合のベン図で表現すると



この表では、「利き目が左目であるという事象」を A と表記し、
「利き目が右目であるという事象」を ～A と表記しています。
～○は、「○以外の全て(everything but A)」を表していて、
集合としては補集合(Complement)を意味します。

この例では、利き目を聞いた100人が情報源のすべてです。
情報源全体を U と表記しています。

Universeの U

補足：補集合の表記

- 集合Aの補集合には、いくつかの表記方法があります

$$\overline{A} \qquad A^c \qquad \sim A$$

- もっともよく使われているのは $\overline{A}$ ですが、以下の理由で本資料では $\sim A$ （←英米の教科書でよく使われている）を使っています

  - $\bar{A}$ はOffice 365で数式の挿入や特殊記号の挿入で書くことはできますが、
    - ✓挿入作業自体が煩雑
    - ✓他のフォント（特にゴシック系）とのバランスがよくない
    - ✓慣れていないと補集合であることを見落とす可能性がある

利き目を聞いた100人のうちの一人（Sさん）

- 100人中、70人が左目、30人が右目という情報があり、100人のうちの一人Sさんの利き目をあなたは知らないとします。
- この状況で、Sさんの利き目が左目である確率は、

$$\mathrm{Pr}(\text{Sさんが左目である；A})=\frac{\text{左目の}70\text{人}}{\text{利き目を聞いた}100\text{人}}=\frac{70}{100}=0.7$$

- Sさんの利き目が右目である確率は、

$$\mathrm{Pr}(\text{Sさんが右目である；}\sim\text{A})=\frac{\text{左目の}30\text{人}}{\text{利き目を聞いた}100\text{人}}=\frac{30}{100}=0.3$$

Aと～AはMECEなので
Pr(A)+Pr(～A)=1
が成り立ちます

注）この例では、左目か右目かを互いに排反（mutually exclusive）としています。
実際には、利き目を特定できない人が稀にいるようです。その場合には ～A という集合には利き目が右目の人と利き目を特定できない人が含まれることになります。

もう一つの確率：足指のタイプ

あなたの足の親指と人差し指の長さの関係なんて考えたことあります？
（今、ここで確認しなくていいですからね！）

Wikipedia
“Morton's toe”
より

【Morton’s toe （モートンズ·トー、ギリシャ足指）】
足の親指よりも人差し指が長い足の形状を指し、
一般的には生まれつきのものです。これにより、
通常よりも人差し指の付け根に過度な圧力がかかり、
痛みを引き起こすことがあります。

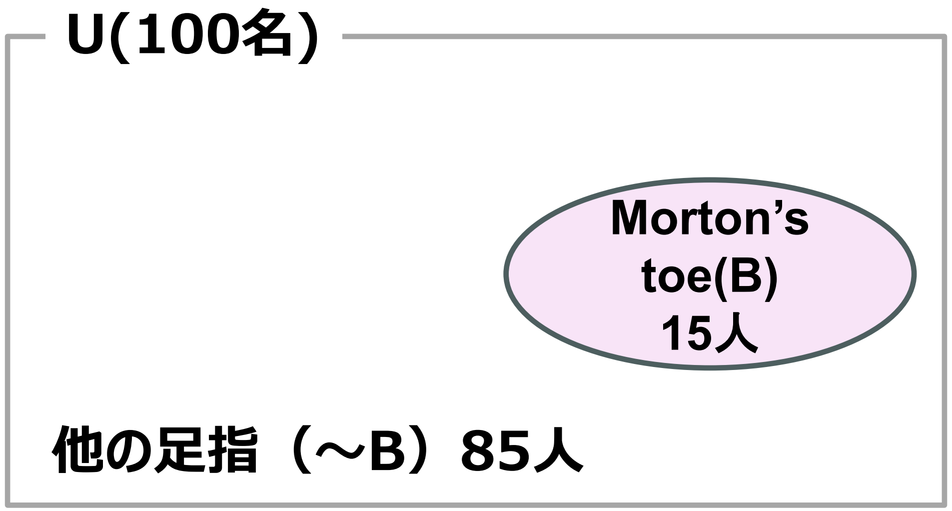
対処法としては、フィットした靴を選ぶ、インソール（足底板）を使う、
足底のパッドを試す、といった保存的療法が一般的なのだそうです。
全人類の20%未満（？）の人がMorton’s toeとWikipediaでは
紹介されています。

100人に聞きました

事象(Event)	頻度（Frequency）
Morton’s toe（B）	15
他の足指（～B）	85
全体(U)	100

この表では、「Morton’s toeという事象」を B と表記し、
「他の足指タイプであるという事象」を ～B と表記しています。

これを集合のベン図で表現すると



Bと～BもMECEなので
Pr(B)+Pr(～B)=1
が成り立ちます

2つの性質（利き目、足指）を同時に考えるとどうなる？

２つの性質はそれぞれ２つの結果があるので、全部で４つの組み合わせが考えられます

1. 左目 かつ Morton’s toe; A∩B と表記
2. 左目 かつ 他の足指タイプ; A∩～B と表記
3. 右目 かつ Morton’s toe; ～A∩B と表記
4. 右目 かつ 他の足指タイプ; ～A∩～B と表記

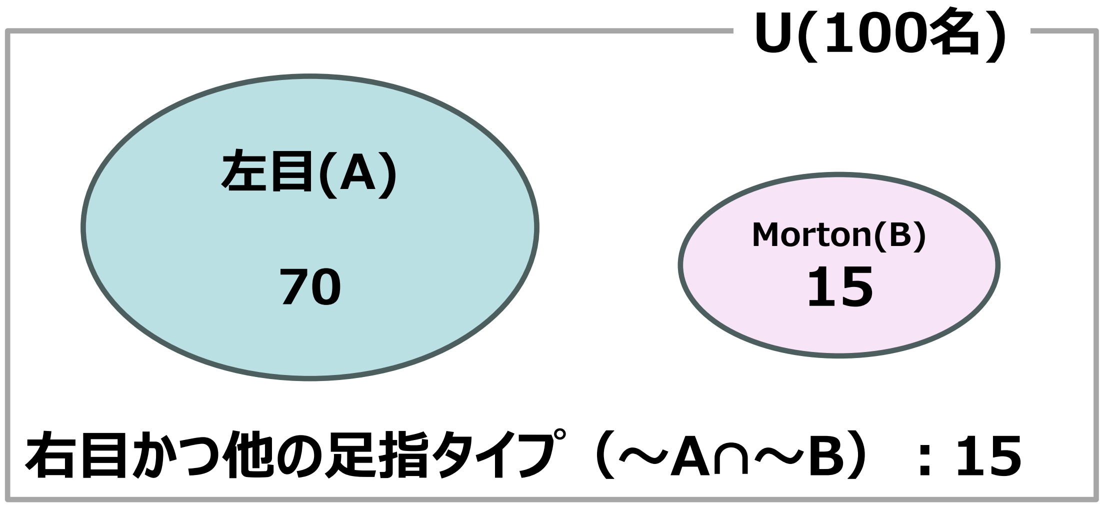
∩ は集合で使われる数学的表記で Intersection （共通部分）を意味します。
たとえば「 A∩B は“AかつB”（英語では”(both) A and B” ）」と読みます。

今までと同じ100人に聞きました

もし『AかつB』の人が5人だとしたら？
（次のスライド）

	左目(A)	右目(〜A)	合計
Morton's toe(B)	0 A∩B	15 〜A∩B	15
他の足指タイプ(〜B)	70 A∩〜B	15 〜A∩〜B	85
合計	70	30	100

これをベン図で表現すると



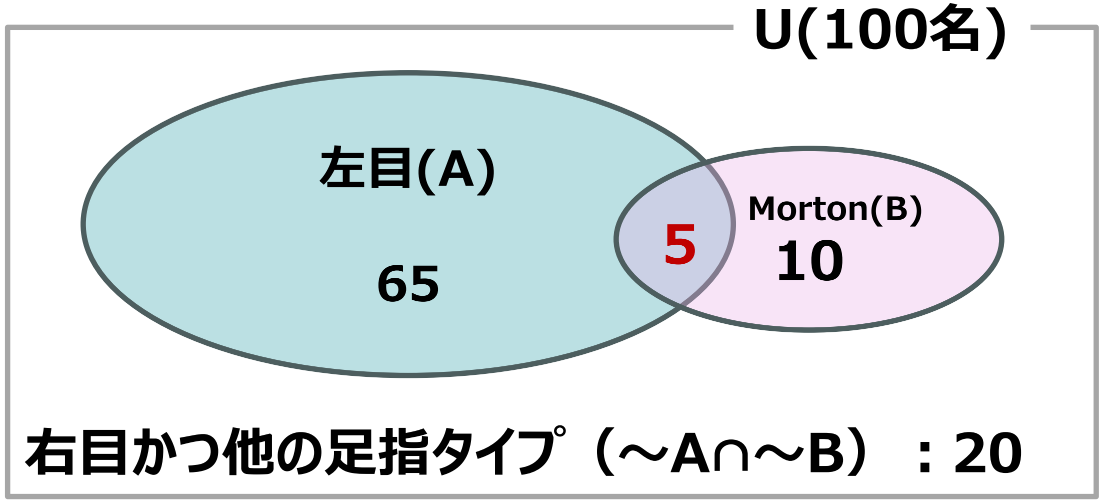
AかつBの
人がいない！
（AとBの
共通部分がない）

左目(A)とMorton(B)は
互いに排反
（両性質を同時に持つ人はいない）

左目かつMortonが5人だとしたら？

	左目(A)	右目(～A)	合計
Morton’s toe(B)	5	10	15
他の足指タイプ(～B)	65	20	85
合計	70	30	100

これをベン図で表現すると



AかつBの
人が5人いる！
（AとBの
共通部分がある）

左目(A)とMorton(B)は
互いに排反ではない
（両性質を同時に持つ人がいる）

分割表のマージン (Margin：周辺、縁、端の意)

の部分は マージン（Margin：周辺、縁、端の意）と言われます

	左目(A)	右目(～A)	合計
Morton's toe(B)	5	10	15
他の足指タイプ(～B)	65	20	85
合計	70	30	100

Morton(B)、そうじゃない(～B)の総人数を示しています（右目、左目にはよらない）

左目(A)、右目(～A)の総人数を示しています（Mortonかそうじゃないかにはよらない）

マージンの個々の人数がどうであれ総人数は100人

『この分割表と確率がどう結びつくのさ？』

『この分割表と確率がどう結びつくのさ？』

『この100人からランダムに選ばれたTさんが、この表のどこに属するか』の確率を考えます。あなたは、Tさんの利き目やMorton’s toeについて知りません。

	左目(A)	右目(～A)	合計
Morton’s toe(B)	5	10	15
他の足指タイプ(～B)	65	20	85
合計	70	30	100

$$\mathbf{Pr} = \frac{\text{注目する組み合わせのグループの人数}}{\text{全体の人数（}\leftarrow 100\text{人）}}$$

- **表全体を総合計人数（100人）で割ると確率の分割表になる**

	左目(A)	右目(～A)	合計
Morton’s toe(B)	0.05	0.10	0.15
他の足指タイプ(～B)	0.65	0.20	0.85
合計	0.70	0.30	1

ここで注意してほしいこと

- サイコロを振る賭けのときは、それぞれの目がどのくらいの確率で出やすいのか情報がなかったので、サイコロを多数回振る必要がありました。
- 今回の例では、各セルの人数が与えられているので、確率が簡単に計算できます。
- 「100人に聞きました」の100人：これが今考えてる情報源のすべてで私たちの Universe（全体集合であり、母集団）です。
- つまり、この100人は大きな集団からの標本ではない！
- もし、私たちがUniverseの構成（どのセルに何人か）を知らなければ、Universeから1人を連れてくるという試行を繰り返し、繰り返し行うことで確率が計算されます。

同時確率 (Joint Probability)

- ある１つのセルについての確率の計算を見ていきましょう。
- 100人からランダムに選ばれたTさんが、『左目(A) かつ Morton(B)』である確率を知りたいとします。
- 『左目(A) かつ Morton(B)』を A ∩ B という集合で表現し、この集合に含まれる人数を【 A ∩ B 】のように書くことにすると…

$$\Pr(A \cap B) = \frac{【A \cap B】}{【U】} = \frac{5}{100} = 0.05$$

事象Aと事象Bの同時確率という

に含まれる他の3つの確率も、同時確率です。

	左目(A)	右目(～A)	合計
Morton's toe(B)	0.05	0.10	0.15
他の足指タイプ(～B)	0.65	0.20	0.85
合計	0.70	0.30	1

同時確率 (Joint Probability)ーつづきー

- 確率の文脈で、“同時 (Joint)”という単語を聞いた時には、『かつ (AND)』という単語を連想してください。
集団の複数の性質を考えている（定量化しようとしている）と思ってください。

- 同時確率では、事象Aと事象Bの順序（どちらが先か）は区別されないので、以下が成立します。

$$\boxed{\Pr(A \cap B) = \Pr(B \cap A)}$$

同時確率の重要な性質の一つです

周辺確率 (Marginal Probability)

この表には、別の種類の確率も与えられています。
の部分はマージンと呼ぶんでしたね。マージンに含まれる確率は、
周辺確率(Marginal Probability)と呼ばれます。

	左目(A)	右目(～A)	合計
Morton’s toe(B)	0.05	0.10	0.15
他の足指タイプ(～B)	0.65	0.20	0.85
合計	0.70	0.30	1

Mortonの確率
他の足指の確率

Aか～Aかにはよらない
（利き目によらない）

左目の確率　右目の確率

Bか～Bかにはよらない
（Mortonかどうかによらない）

同時確率 と 周辺確率

	左目(A)	右目(〜A)	合計
Morton’s toe(B)	同時確率	同時確率	周辺確率
他の足指タイプ(〜B)	同時確率	同時確率	周辺確率
合計	周辺確率	周辺確率	1

すべての確率の合計は１
（横方向の周辺確率の和も１）
（縦方向の周辺確率の和も１）
（同時確率の総和も１）

空欄になってるセルをうめる！

あなたは、「左目の確率」が0.7、「Mortonの確率」が0.15、「左目 かつ Morton」の確率が0.05であることだけを知っているとします。他の空欄の確率はうめられますか？

	左目(A)	右目(～A)	合計
Morton’s toe(B)	0.05	0.10	0.15
他の足指タイプ(～B)	0.65	0.20	0.85
合計	0.7	0.3	1

周辺確率の和も１
1−0.15=0.85

すべての確率の和は１

周辺確率の和も１
1−0.7=0.3

0.7−0.05=0.65

0.15−0.05=0.10

0.3−0.10=0.20

周辺確率が分かれば
他の空欄（同時確率）は
すべて引き算だけで計算できる

同時確率 と 周辺確率

前のスライドで示した計算をPr()表記を使って表現すると
「一般的に成り立つ関係」を表現できます

	左目(A)	右目(～A)	合計
Morton's toe(B)	Pr(A∩B)	Pr(～A∩B)	Pr(B)
他の足指タイプ(～B)	Pr(A∩～B)	Pr(～A∩～B)	Pr(～B)
合計	Pr(A)	Pr(～A)	1

Pr(B)
+Pr(～B) = 1

Bと～BはMECE
なのでこの式が
成り立つ

Pr(A) +Pr(～A) = 1

Aと～AはMECE
なのでこの式が
成り立つ

同時確率 と 周辺確率

前のスライドで示した計算をPr()表記を使って表現すると
「一般的に成り立つ関係」を表現できます

	左目(A)	右目(～A)	合計
Morton’s toe(B)	Pr(A∩B)	Pr(～A∩B)	Pr(B)
他の足指タイプ(～B)	Pr(A∩～B)	Pr(～A∩～B)	Pr(～B)
合計	Pr(A)	Pr(～A)	1

Pr(A∩B)+Pr(A∩～B)= Pr(A)　　Pr(～A∩B)+Pr(～A∩～B)= Pr(～A)

Aや～Aの周辺確率は、別の視点（B or ～B）で分割される

同時確率 と 周辺確率

横方向も同様で…

	左目(A)	右目(～A)	合計
Morton(B)	Pr(A∩B)	Pr(～A∩B)	Pr(B)
他の足指(～B)	Pr(A∩～B)	Pr(～A∩～B)	Pr(～B)
合計	Pr(A)	Pr(～A)	1

Pr(A∩B)+Pr(～A∩B)= Pr(B)

Pr(A∩～B)+Pr(～A∩～B)= Pr(～B)

Bや～Bの周辺確率は、
別の視点（A or ～A）で分割される

また、すべての同時確率の和も１になりますね。

1=Pr(B)+Pr(～B)= Pr(A∩B)+Pr(～A∩B)＋Pr(A∩～B)+Pr(～A∩～B)

「もし K さんがMorton’s toe だと、わかっている場合、Kさんが左目が利き目である確率は？」

- めちゃくちゃGood Questionです！
- その質問に答えるためには、条件付き確率（Conditional Probability）という別のタイプの確率を考える必要があります。

- これまでに学んだ同時確率や周辺確率は２つの確率（利き目はどっち？Morton’s toeかどうか？）を考えていましたが、利き目とMortonを知る順序については問いませんでした。
- 上の質問の場合、利き目とMortonについて知る順序が明確にあります
- つまり、『Mortonであることを知った上で（知った後で）、』利き目の確率を考えているからです。

条件付き確率（Conditional Probability）

条件付き確率とは、
ある事象が起きた条件下で、別の事象が起きる確率です。

KさんがMortonであることを知っている（確定している）という条件の下で

利き目が左目

| … は、given…
Mortonであることを先に知っている

条件付き確率は、以下のように書きます。

Pr(A|B) "the probability of A, given that B has occurred"; Pr(左目|Morton)

（Bが起きたということが与えられている条件の下でのAの確率）

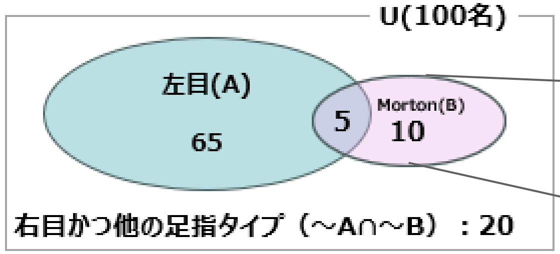
同様に他の条件付き確率も考えられます。
Pr(A|～B)＝Pr(左目|他の足指)；Kさんが他の足指であることを知った上での、左目が利き目の確率
Pr(B|～A)＝Pr(Morton|右目)；Kさんが右目が利き目であることを知った上での、Mortonである確率
などなど

条件付き確率についての超重要な公式

条件付き確率 $\Pr(A|B) = \dfrac{\Pr(A \cap B)}{\Pr(B)}$ 同時確率 / 周辺確率

given の事象

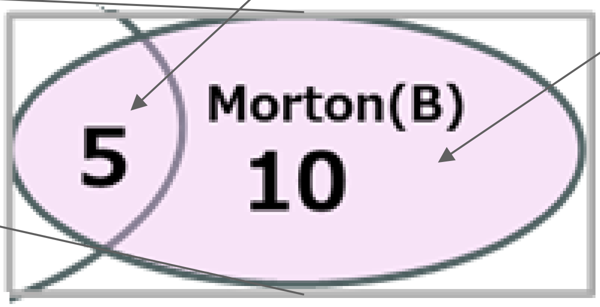
ベン図を使った説明



B に 焦点！

A∩B

～A∩B



↑先ほど考えていたベン図はこれで
私たちのUniverseは100人でした。

つまり、

$$\Pr(A|B) = \frac{5}{15} = 0.33 = \frac{\Pr(A \cap B)}{\Pr(B)}$$

$$\Pr(\sim A|B) = \frac{10}{15} = 0.66 = \frac{\Pr(\sim A \cap B)}{\Pr(B)}$$

これらの和も１です。Pr(A|B)+Pr(～A|B)＝１

B(Morton)であることを知った上で
A(左目)の条件付き確率を考える場合の
私たちのUniverseはMortonである15人。
その15人中、A(左目)と～A(右目)が占めている
割合を考えます。

条件付き確率についての超重要な公式

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

右辺の分数は、
Bのうち A∩Bが占める割合
を表しています

分割表を使った説明

この行(B)に
焦点を当てる

	左目(A)	右目(~A)	合計
Morton's toe(B)	5	10	15
他の足指タイプ(~B)	65	20	85
合計	70	30	100

これが分母の B の人数

これが分子の A∩B の人数

つまり $$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\frac{5}{100}}{\frac{15}{100}} = \frac{5}{15} = 0.33$$

条件付き確率についての超重要な公式（覚え方）

この公式は、条件付き確率と周辺確率と同時確率を結びつける
重要な関係式なので、はっきりとしたイメージを持っておきましょう

A|Bの表記の順序はそのままでA∩B

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

条件付き確率Pr(A|B)は
- ✓ AとBの同時確率
- ✓ Bの周辺確率

がわかれば計算できる

|の右の事象が(givenの事象)が分母に

Pr(A|B) ＝ Pr(B|A) ってことになるの？

さぁ、いよいよベイズの定理に辿り着く、最後の考察に入ります！

上の質問は
「Morton(B)が確定しているときの左目(A)の条件付き確率」と
「左目(A)が確定しているときのMorton(B)の条件付き確率」は
同じか？です。その答えはNO! です。計算してみましょう。

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\frac{5}{100}}{\frac{15}{100}} = \frac{5}{15} = 0.333$$

$$\Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} = \frac{\frac{5}{100}}{\frac{70}{100}} = \frac{5}{70} = 0.072$$

これらはぜんぜん別もの

（分母のPr(B)とPr(A)がたまたま同じ時だけ一致しますけどね）

前のスライドの計算を確率の分割表から見てみよう

	左目(A)	右目(～A)	合計
Morton’s toe(B)	0.05	0.10	0.15
他の足指タイプ(～B)	0.65	0.20	0.85
合計	0.70	0.30	1

Pr(A|B)

Bがgivenになってるので
Universe はココ

$$\Pr(A|B)$$

$$= \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0.05}{0.15} = 0.333$$

Pr(B|A)

Aがgivenになってるので
Universe はココ

$$\Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} = \frac{0.05}{0.70} = 0.072$$

条件付き確率と周辺確率から同時確率が計算できる？

条件付き確率についての超重要な公式はこうでした。

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

この等式の両辺に Pr(B) を掛ければ、このように変形できます。

$$\underbrace{\Pr(A \cap B)}_{\text{同時確率}} = \underbrace{\Pr(A|B)}_{\text{条件付き確率}} * \underbrace{\Pr(B)}_{\text{周辺確率}}$$

事象Aと事象Bの同時確率は、
「事象Bが起きたことを知った上で、事象Aが起きる条件付き確率」と
「事象Bの周辺確率（Aによらない）」の積である

A,B逆の条件付き確率 Pr(B|A)でも同様に考えられますね？

条件付き確率についての超重要な公式をPr(B|A)で考えると。

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

この等式の両辺に Pr(A) を掛ければ、このように変形できます。

$$\underbrace{\Pr(B \cap A)}_{\text{同時確率}} = \underbrace{\Pr(B|A)}_{\text{条件付き確率}} * \underbrace{\Pr(A)}_{\text{周辺確率}}$$

事象Bと事象Aの同時確率は、
「事象Aが起きたことを知った上で、事象Bが起きる条件付き確率」と
「事象Aの周辺確率（Bによらない）」の積である

同時確率って、AとBの順序は関係ないんでしたよね？

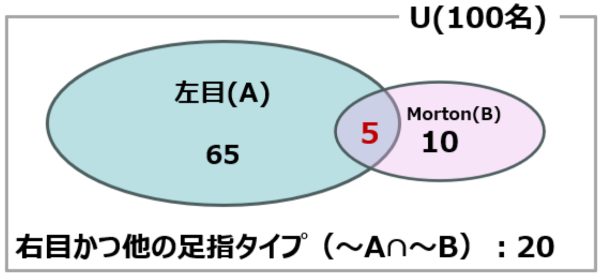
確かめてみましょう！

$$\Pr(A \cap B) = \Pr(A|\boldsymbol{B}) * \Pr(\boldsymbol{B}) = \frac{0.05}{0.15} * 0.15 = 0.05$$

$$\parallel$$

$$\Pr(B \cap A) = \Pr(B|\boldsymbol{A}) * \Pr(\boldsymbol{A}) = \frac{0.05}{0.70} * 0.70 = 0.05$$

思ったとおりですね。AとBの共通部分（青と赤の交わる部分）に
AとBが起こる順序は関係なかった。
ここで大事なことは、条件付き確率の
”given”以降の部分。これに対応する周辺確率を
条件付き確率に掛けると、同じ結果が得られるということ。



いよいよ、ベイズの定理を導きましょう！

$$\mathbf{Pr}(A \cap B) = \underline{\mathbf{Pr}(A|B) * \mathbf{Pr}(B)} \qquad \mathbf{Pr}(B \cap A) = \underline{\mathbf{Pr}(B|A) * \mathbf{Pr}(A)}$$

両者の左辺（Pr(A∩B) と Pr(B∩A)）は等しいのだから、両者の右辺も等しい

$$\mathbf{Pr}(A|B) * \mathbf{Pr}(B) = \mathbf{Pr}(B|A) * \mathbf{Pr}(A)$$

両辺を Pr(B) で割ると

ベイズの定理

$$\boxed{\mathbf{Pr}(A|B) = \frac{\mathbf{Pr}(B|A) * \mathbf{Pr}(A)}{\mathbf{Pr}(B)}}$$

で、その式のどこがスゴイのさ

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)}$$

この式は、
２つの周辺確率 Pr(A) と Pr(B)、２つの条件付き確率 Pr(A|B)、Pr(B|A)という
４つの確率について常に成り立つ関係を示しています。
４つのうち３つの確率がわかれば、残りの１つは簡単に計算できます。
特に２つの条件付き確率 Pr(A|B)、Pr(B|A)はAとBが逆(知る順序が逆)になっています。

どんな場合に使えるのか、有名な例を使って説明しましょう。
（この例は https://www.yudkowsky.net/rational/bayes で紹介されています）

ベイズの定理の適用例：乳がんとマンモグラフィ

Question：

40歳の女性のうち、定期健診に参加した女性の1%が乳がんを患っています。
乳がんを患っている　女性の80%はマンモグラフィ検査で陽性結果を示しますが、
乳がんを患っていない女性の9.6%も陽性結果を示します。
40歳のある女性が定期健診のマンモグラフィで陽性結果を受け取りました。
彼女が実際に乳がんを患っている確率はいくつでしょうか？

Pr（マンモ陽性|乳がん）

「乳がんを患っている女性がマンモグラフィで陽性となる確率」が与えられています。
「マンモグラフィ陽性の女性が乳がんを患っている確率」を聞いています。

Pr（乳がん|マンモ陽性）

おっ！これらは逆の条件付き確率じゃありませんか！

分割表に整理してみましょう

乳がんであるという事象を　A
乳がんでないという事象を～A
マンモ陽性という事象を　B
マンモ陰性という事象を～B

と表記することにしましょう

	乳がん(A)	No 乳がん(～A)	合計
マンモ陽性(B)	?	?	?
マンモ陰性(～B)	?	?	?
合計	0.01	0.99	1

1 − 0.01 = 0.99

「40歳の女性のうち、定期健診に参加した女性の1%が乳がんを患っています。」

Questionより

分割表に整理してみましょうーつづきー

「乳がんを患っている女性(A)の80%はマンモグラフィ検査で陽性結果(B)を示しますが、」

Questionより

この情報は、条件付き確率 Pr(マンモ陽性｜乳がん）を与えています。

Pr(B|A) = 0.8

条件付き確率は分割表にはありませんが、
Pr(B|A)を使うと「乳がん かつ マンモ陽性」の同時確率を計算できます。

Pr(B∩A)=Pr(B|A)＊Pr(A)=0.8＊0.01=0.008

Pr(～A∩B)か
Pr(～A∩～B)が
わかれば
分割表はすべて
うまる！

	乳がん(A)	No 乳がん(～A)	合計
マンモ陽性(B)	0.008	?	?
マンモ陰性(～B)	0.002	?	?
合計	0.01	0.99	1

0.01－0.008＝0.002

分割表が完成すればPr(A|B)はベイズの定理で計算できる

「乳がんを患っていない女性の9.6%も陽性結果を示します。」

Questionより

この情報は、条件付き確率 Pr(マンモ陽性 | No乳がん）を与えています。

Pr(B|～A) = 0.096

これを使うと「マンモ陽性 かつ No乳がん」の同時確率を計算できます。

Pr(B∩～A)=Pr(B|～A) ∗ Pr(～A)=0.096 ∗ 0.99=0.095

	乳がん(A)	No 乳がん(～A)	合計
マンモ陽性(B)	0.008	0.095	0.103
マンモ陰性(～B)	0.002	0.895	0.897
合計	0.01	0.99	1

Pr(B|A)=0.8は
重要な公式を用いて

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

$$= \frac{0.008}{0.01} = 0.8$$

でも求められる

Questionの中で与えられていた

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)} = \frac{0.8 * 0.01}{0.103} = 0.0776$$

Pr(A|B)以外の条件付き確率も計算できる

マンモの陽性/陰性の結果が得られたときの、乳がんの有無の確率を計算したいとすると、
以下の４とおりの条件付き確率が考えられます。

$\Pr(A\|B)$；マンモ陽性の時の、乳がんアリの確率
$\Pr(\sim A\|B)$；マンモ陽性の時の、乳がんナシの確率
$\Pr(A\|\sim B)$；マンモ陰性の時の、乳がんアリの確率
$\Pr(\sim A\|\sim B)$；マンモ陰性の時の、乳がんナシの確率

分割表の数値を拾いやすくするには、条件付き確率の重要な式

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

…AとBの同時確率

…Bの周辺確率

のほうが便利です。当然ベイズの定理からもこの重要な式は得られます。

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)} = \frac{\frac{\Pr(B \cap A)}{\cancel{\Pr(A)}} * \cancel{\Pr(A)}}{\Pr(B)} = \frac{\Pr(B \cap A)}{\Pr(B)}$$

Pr(A|B)以外の条件付き確率も計算できるーつづきー

	乳がん(A)	No 乳がん(～A)	合計
マンモ陽性(B)	0.008	0.095	0.103
マンモ陰性(～B)	0.002	0.895	0.897
合計	0.01	0.99	1

Pr(*A*|*B*) ; マンモ陽性の時の、乳がんアリの確率

$$\mathbf{Pr}(\boldsymbol{A}|\boldsymbol{B}) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0.008}{0.103} = 0.0776$$

Pr(A|B)以外の条件付き確率も計算できるーつづきー

	乳がん(A)	No 乳がん(～A)	合計
マンモ陽性(B)	0.008	0.095	0.103
マンモ陰性(～B)	0.002	0.895	0.897
合計	0.01	0.99	1

Pr(A|B)；マンモ陽性の時の、乳がんアリの確率 $\mathbf{Pr}(\boldsymbol{A}|\boldsymbol{B}) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0.008}{0.103} = 0.0776$

Pr(～A|B)；マンモ陽性の時の、乳がんナシの確率 $\mathbf{Pr}(\sim\boldsymbol{A}|\boldsymbol{B}) = \frac{\Pr(\sim A \cap B)}{\Pr(B)} = \frac{0.095}{0.103} = 0.9223$

Pr(A|B)以外の条件付き確率も計算できるーつづきー

	乳がん(A)	No 乳がん(～A)	合計
マンモ陽性(B)	0.008	0.095	0.103
マンモ陰性(～B)	0.002	0.895	0.897
合計	0.01	0.99	1

Pr(*A*|*B*) ; マンモ陽性の時の、乳がんアリの確率 $\mathbf{Pr}(\boldsymbol{A}|\boldsymbol{B}) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0.008}{0.103} = 0.0776$

Pr(～*A*|*B*) ; マンモ陽性の時の、乳がんナシの確率 $\mathbf{Pr}(\sim\boldsymbol{A}|\boldsymbol{B}) = \frac{\Pr(\sim A \cap B)}{\Pr(B)} = \frac{0.095}{0.103} = 0.9223$

Pr(*A*|～*B*) ; マンモ陰性の時の、乳がんアリの確率 $\mathbf{Pr}(\boldsymbol{A}|\sim\boldsymbol{B}) = \frac{\Pr(A \cap \sim B)}{\Pr(\sim B)} = \frac{0.002}{0.897} = 0.0022$

Pr(A|B)以外の条件付き確率も計算できるーつづきー

	乳がん(A)	No 乳がん(～A)	合計
マンモ陽性(B)	0.008	0.095	0.103
マンモ陰性(～B)	0.002	0.895	0.897
合計	0.01	0.99	1

Pr(*A*|*B*)；マンモ陽性の時の、乳がんアリの確率

$$\mathbf{Pr}(\boldsymbol{A}|\boldsymbol{B}) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0.008}{0.103} = 0.0776$$

Pr(～*A*|*B*)；マンモ陽性の時の、乳がんナシの確率

$$\mathbf{Pr}(\sim\boldsymbol{A}|\boldsymbol{B}) = \frac{\Pr(\sim A \cap B)}{\Pr(B)} = \frac{0.095}{0.103} = 0.9223$$

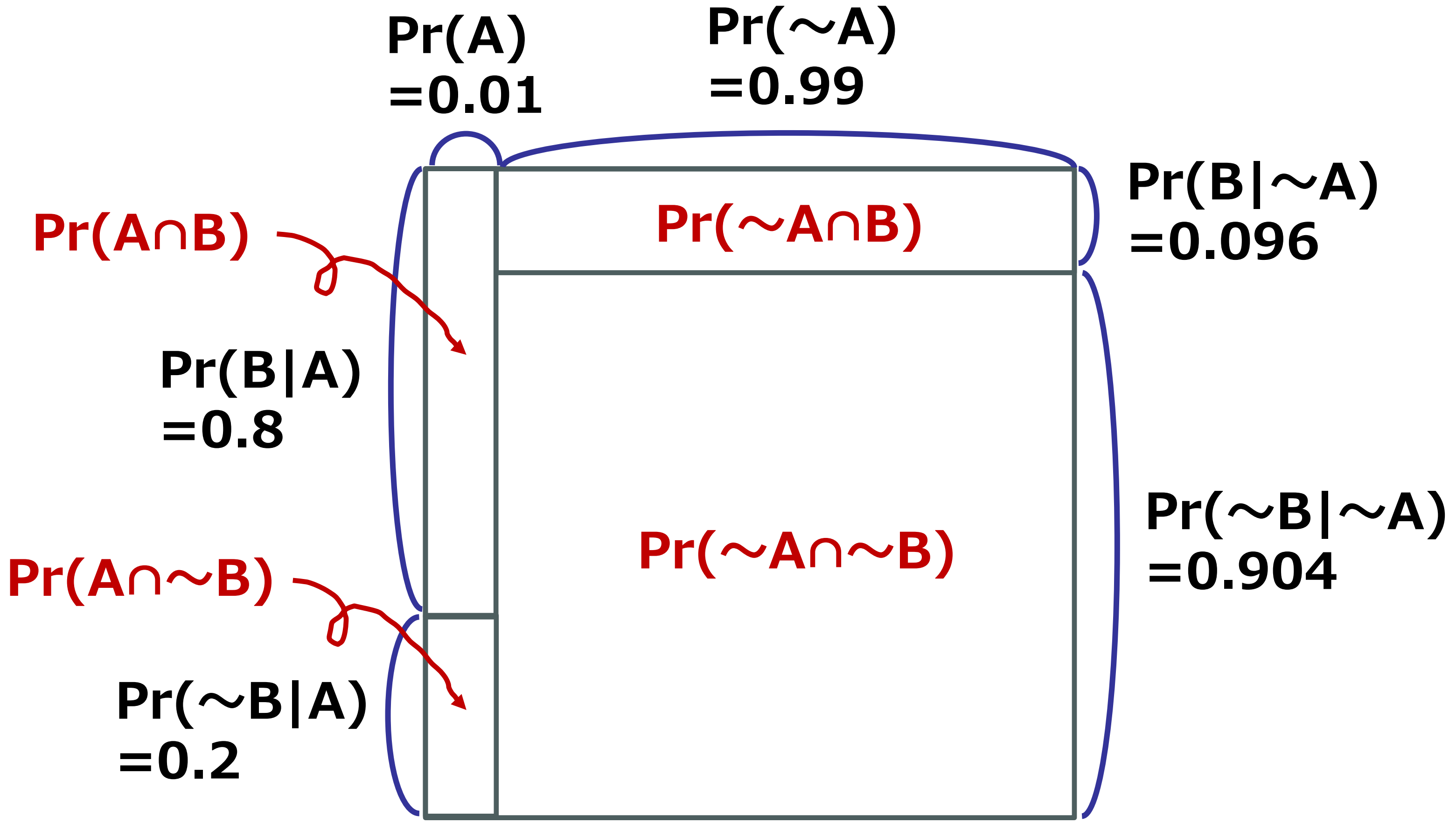
Pr(*A*|～*B*)；マンモ陰性の時の、乳がんアリの確率

$$\mathbf{Pr}(\boldsymbol{A}|\sim\boldsymbol{B}) = \frac{\Pr(A \cap \sim B)}{\Pr(\sim B)} = \frac{0.002}{0.897} = 0.0022$$

Pr(～*A*|～*B*)；マンモ陰性の時の、乳がんナシの確率

$$\mathbf{Pr}(\sim\boldsymbol{A}|\sim\boldsymbol{B}) = \frac{\Pr(\sim A \cap \sim B)}{\Pr(\sim B)} = \frac{0.895}{0.897} = 0.9978$$

Aの周辺確率と、条件付き確率と同時確率の関係



同時確率が
周辺確率と条件付き確率の積
なので、同時確率を面積として
表現した図です。
これまで出てきた式や数字と
にらめっこしてみて！

<<各辺の長さが１、面積が１の正方形>>(←１は確率の和が１であることに対応)

ベイズの定理　(本資料の冒頭のスライドより)

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)}$$

● ベイズの定理の２つの解釈

[解釈1]条件付き確率Pr(A|B) と
条件付き確率Pr(B|A) の間の関係を示している

[解釈2]エビデンスを考慮して主観的な信念の度合いが、
どのように合理的に変化するはずかを表現している

これまでの説明で、
[解釈1] について
理解が深まったと思います

ここからは、[解釈2] に
ついて説明します

違った視点で見てみよう

マンモグラフィ検査をする前は、この女性が乳がんであるかどうかは、
40歳の女性の乳がんの割合（確率）＝0.01（1%）←有病率（Prevalence）
という情報に頼るしかありませんでした。

マンモグラフィの
情報なし

	乳がん(A)	No 乳がん(～A)	合計
合計	0.01	0.99	1

この情報は分割表の周辺確率にありました。

この確率は、マンモ検査の結果（データ）を得る**前**の、事前確率(Prior Probability)と呼ばれ、
事前に持っていた仮説に対する確信の程度と解釈できます。

違った視点で見てみよう一つづき一

そこにマンモグラフィ検査の特性情報が追加されて、
「この女性がマンモグラフィ検査陽性（B）である」という情報（データ）が得られました。

	乳がん(A)	No 乳がん(∼A)	合計
マンモ陽性(B)	0.008	0.095	0.103
マンモ陰性(∼B)	0.002	0.895	0.897
合計	0.01	0.99	1

この女性が乳がんである（A）かどうかは、
条件付き確率 Pr(A|B) で表現できました。

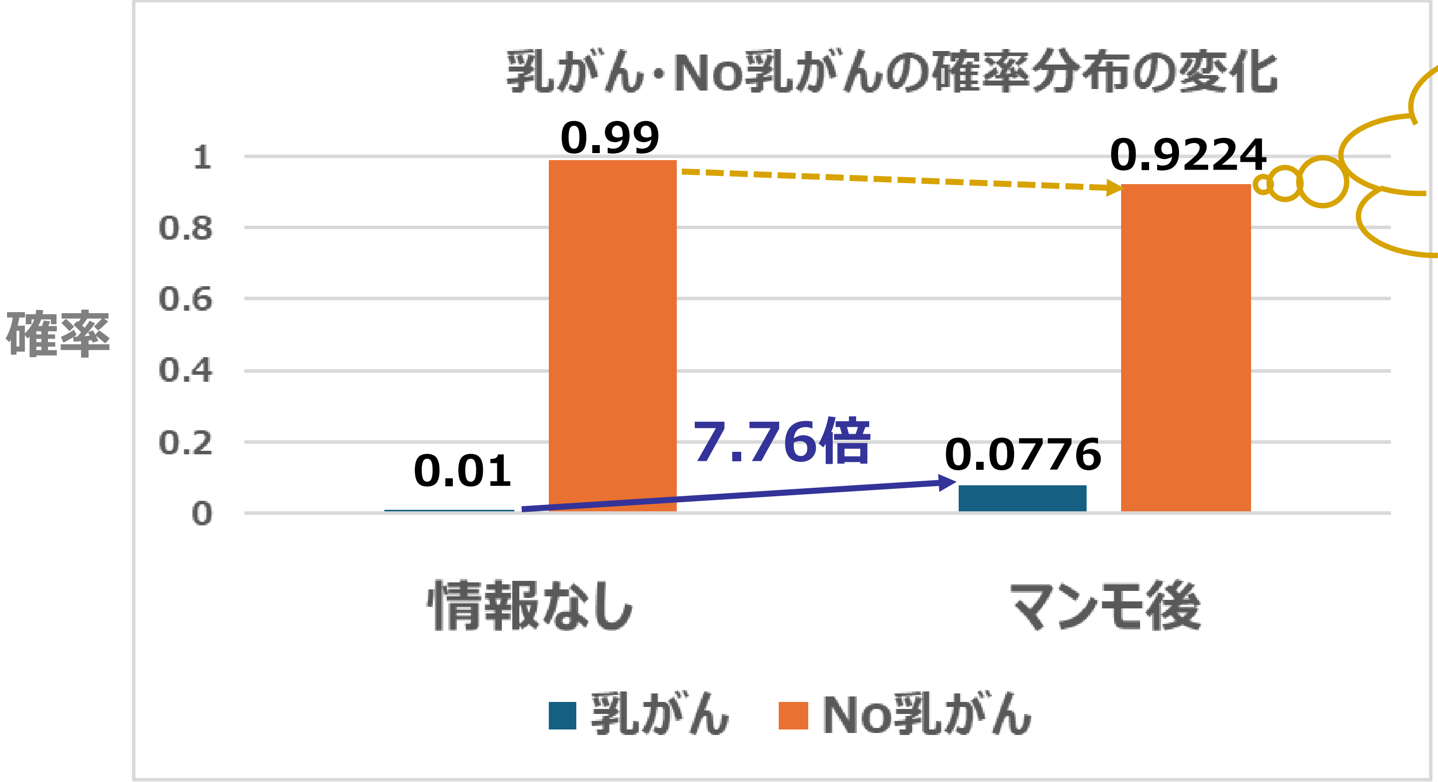
事後確率　　事前確率

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)} = \cdots = 0.0776$$

条件付き確率 Pr(A|B) は、マンモ検査の結果（データ）を得た**後**の、
事後確率(Posterior Probability)と呼ばれ、
得られた情報を加味してUpdateされた仮説に対する確信の程度と解釈できます。

違った視点で見てみようーつづきー

マンモ検査の前の事前確率（←周辺確率Pr(A)）＝0.01であった乳がんの確率は、
マンモ検査の後の乳がんの事後確率（←条件付き確率）として Pr(A|B)=0.0776と計算され、
彼女の乳がんの確率は一般的な統計情報よりも7.76倍に高まったのでした。



乳がんの確率が増えた分、
No乳がんの確率は減った

乳がんの可能性が7.76倍にもなった背景には
マンモグラフィ—の検査特性も影響している

マンモグラフィ検査が、乳がんのあり／なしの判断にまったく役立たない検査であるとすれば
同時確率は乳がんあり／なしの周辺確率を50:50で分配することになり、分割表はこうなるでしょう。

	乳がん(A)	No 乳がん(～A)	合計
マンモ陽性(B)	0.005	0.495	0.5
マンモ陰性(～B)	0.005	0.495	0.5
合計	0.01	0.99	1

Pr(*A*|*B*) ; マンモ陽性の時の、乳がんアリの確率

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0.005}{0.5} = 0.01$$

周辺確率 Pr(A) と同じで変化なし

議論を発展させるために
ベイズの定理の分母を変形しておきます

	乳がん(A)	No 乳がん(∼A)	マージン
マンモ陽性(B)	Pr(A∩B)	Pr(∼A∩B)	Pr(B)
マンモ陰性(∼B)	Pr(A∩∼B)	Pr(∼A∩∼B)	Pr(∼B)
マージン	Pr(A)	Pr(∼A)	Total=1

事象Bを、Aまたは∼Aという事象と同時に考えて
これらに対応する同時確率で分解したもの

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)}$$

$$= \frac{\Pr(B|A) * \Pr(A)}{\Pr(A \cap B) + \Pr(\sim A \cap B)}$$

$$= \frac{\Pr(B|A) * \Pr(A)}{\Pr(B|A) * \Pr(A) + \Pr(B|\sim A) * \Pr(\sim A)}$$

“重要な公式”を使って、以下が成り立つので

$$\Pr(A \cap B) = \Pr(B \cap A) = \Pr(B|A) * \Pr(A)$$

$$\Pr(\sim A \cap B) = \Pr(B \cap \sim A) = \Pr(B|\sim A) * \Pr(\sim A)$$

注意：この例ではマンモ陽性(B)のみ考えていて、
陰性(∼B)は経験していないので、∼Bが含まれていないことをご確認ください。

数式が大きくなって…くらくらしてきました？

まぁ、まぁ、落ち着いて！新バージョンのベイズの定理を見ていきましょう

$$\mathrm{Pr}(A|B) = \frac{\mathrm{Pr}(B|A) * \mathrm{Pr}(A)}{\underbrace{\mathrm{Pr}(B|A) * \mathrm{Pr}(A)} + \underbrace{\mathrm{Pr}(B|{\sim}A) * \mathrm{Pr}({\sim}A)}}$$

この分母全体は、
Bの周辺確率 Pr(B) でしたよね

この部分はBとAの
同時確率Pr(B∩A)
でしたよね

この部分はBと～Aの
同時確率Pr(B∩～A)
でしたよね

これらを足し合わせると

新バージョンのベイズの定理をじっくり見ると、注目すべきことが見えてきます。
（もうお気づきかしら？）

新バージョンのベイズの定理で注目すべきこと

そうです。青字の部分が共通していますね。

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B|A) * \Pr(A) + \Pr(B|{\sim}A) * \Pr({\sim}A)}$$

新バージョンの式で乳がん・マンモ問題を解いてみましょう

	乳がん(A)	No 乳がん(～A)	合計
マンモ陽性(B)	0.008	0.095	0.103
マンモ陰性(～B)	0.002	0.895	0.897
合計	0.01	0.99	1

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B|A) * \Pr(A) + \Pr(B|\sim A) * \Pr(\sim A)}$$

$$= \frac{\frac{0.008}{0.01} * 0.01}{\frac{0.008}{0.01} * 0.01 + \frac{0.095}{0.99} * 0.99} = 0.0776$$

式を変形した
だけですから
当たり前ですが、
さっきと同じです

検査の感度、特異度との関係

	乳がん(A)	No 乳がん(~A)	合計
マンモ陽性(B)	0.008	0.095	0.103
マンモ陰性(~B)	0.002	0.895	0.897
合計	0.01	0.99	1

陽性的中率

Positive Predictive Value

$$= \frac{\boldsymbol{Pr(A \cap B)}}{\boldsymbol{Pr(B)}} = \frac{0.008}{0.103} = 0.0778$$

これって、Pr(A|B)
ではないですか！
マンモ陽性を知った後の
事後確率と
解釈できます

これは周辺確率P(A)ですが、
マンモ検査の前のAの
事前確率と解釈できます

感度

Sensitivity

$$= \frac{\boldsymbol{Pr(A \cap B)}}{\boldsymbol{Pr(A)}} = \frac{0.008}{0.01} = 0.8$$

これって、
Pr(B|A)
ではないですか！

特異度

Specificity

$$= \frac{\boldsymbol{Pr(\sim A \cap \sim B)}}{\boldsymbol{Pr(\sim A)}} = \frac{0.895}{0.99} = 0.904$$

これって、
Pr(~B|~A)
ではないですか！

ベイズの定理 → 感度、特異度、事前確率、事後確率の関係

$$\Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} = \text{感度}$$

マンモ検査の前に持っていた仮説（乳がん）の事前確率

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)} = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B|A) * \Pr(A) + \Pr(B|\sim A) * \Pr(\sim A)}$$

マンモ検査の後でUpdateされた仮説（乳がん）の事後確率

$$= \frac{Pr(B \cap \sim A)}{Pr(\sim A)}$$
$$= 1 - \frac{Pr(\sim B \cap \sim A)}{Pr(\sim A)}$$
$$= 1 - \text{特異度}$$

$$= 1 - \Pr(A)$$
$$= 1 - \text{事前確率}$$

	乳がん(A)	No 乳がん(~A)	合計	
マンモ陽性(B)	0.008	0.095	0.103	陽性的中率
マンモ陰性(~B)	0.002	0.895	0.897	
合計	0.01	0.99	1	

感度
Sensitivity
$= \frac{Pr(A \cap B)}{Pr(A)}$

特異度
Specificity
$= \frac{Pr(\sim A \cap \sim B)}{Pr(\sim A)}$

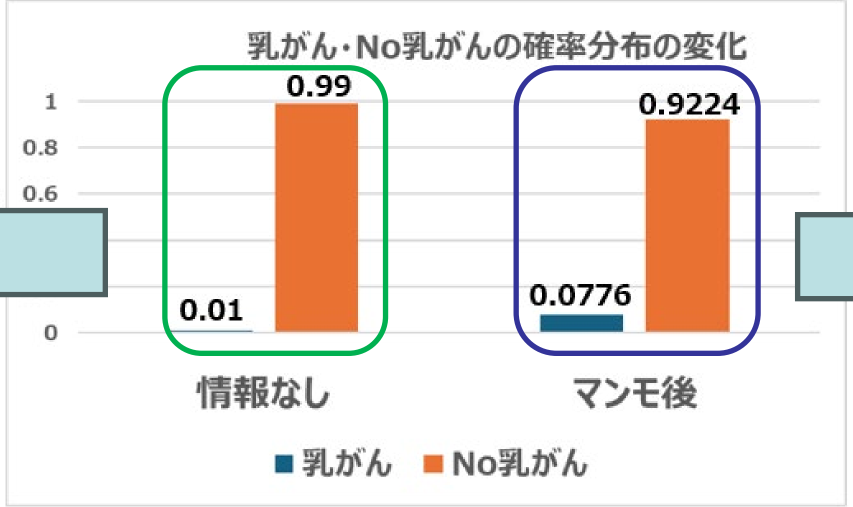
Positive Predictive Value
$= \frac{Pr(A \cap B)}{Pr(B)}$

つまり

$$\text{事後確率} = \frac{\text{感度} * \text{事前確率}}{\text{感度} * \text{事前確率} + (1 - \text{特異度}) * (1 - \text{事前確率})}$$

マンモ検査の後でUpdateされた仮説（乳がん）の事後確率

ベイズの定理の解釈２



データ
（マンモ検査の陽性結果 B）

Pr(A)

事前確率
Prior Probability
乳がんであるという仮説
に対する確信

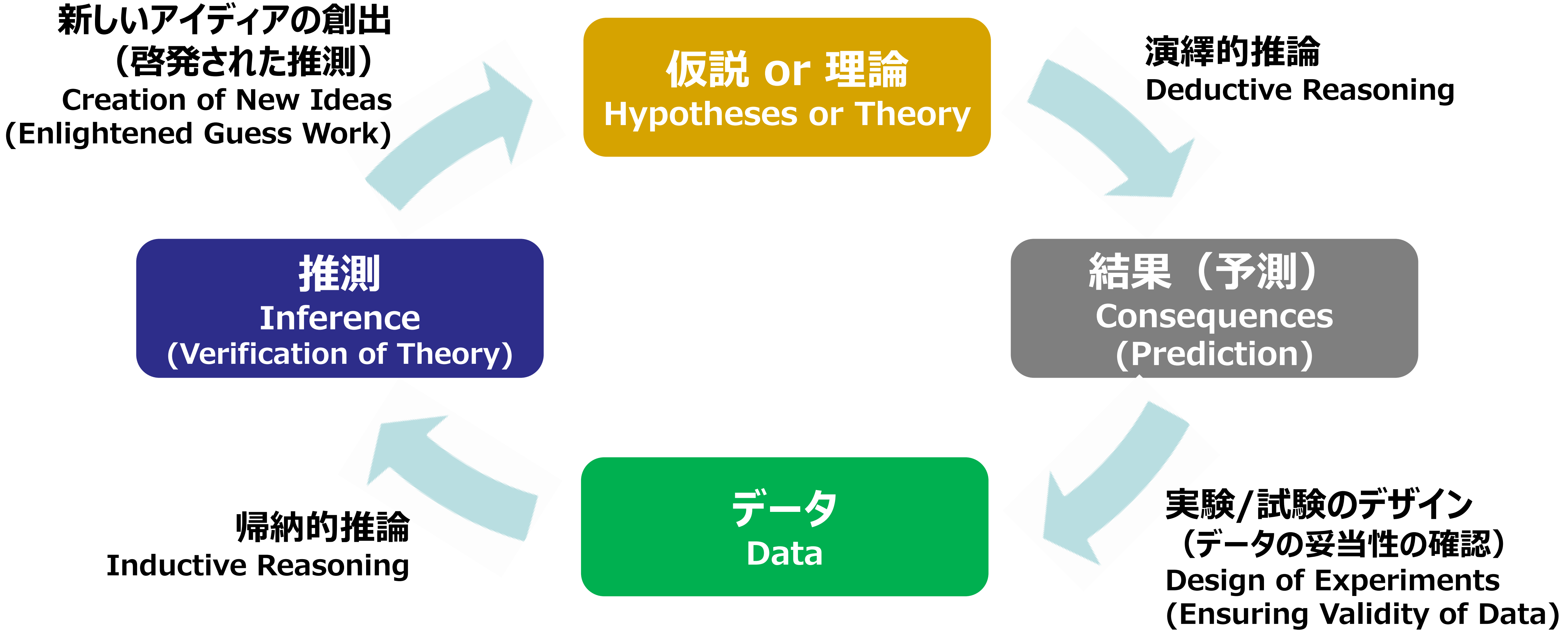
乳がんでないという仮説（～A）
に対する確信も計算されている

Pr(A|B)

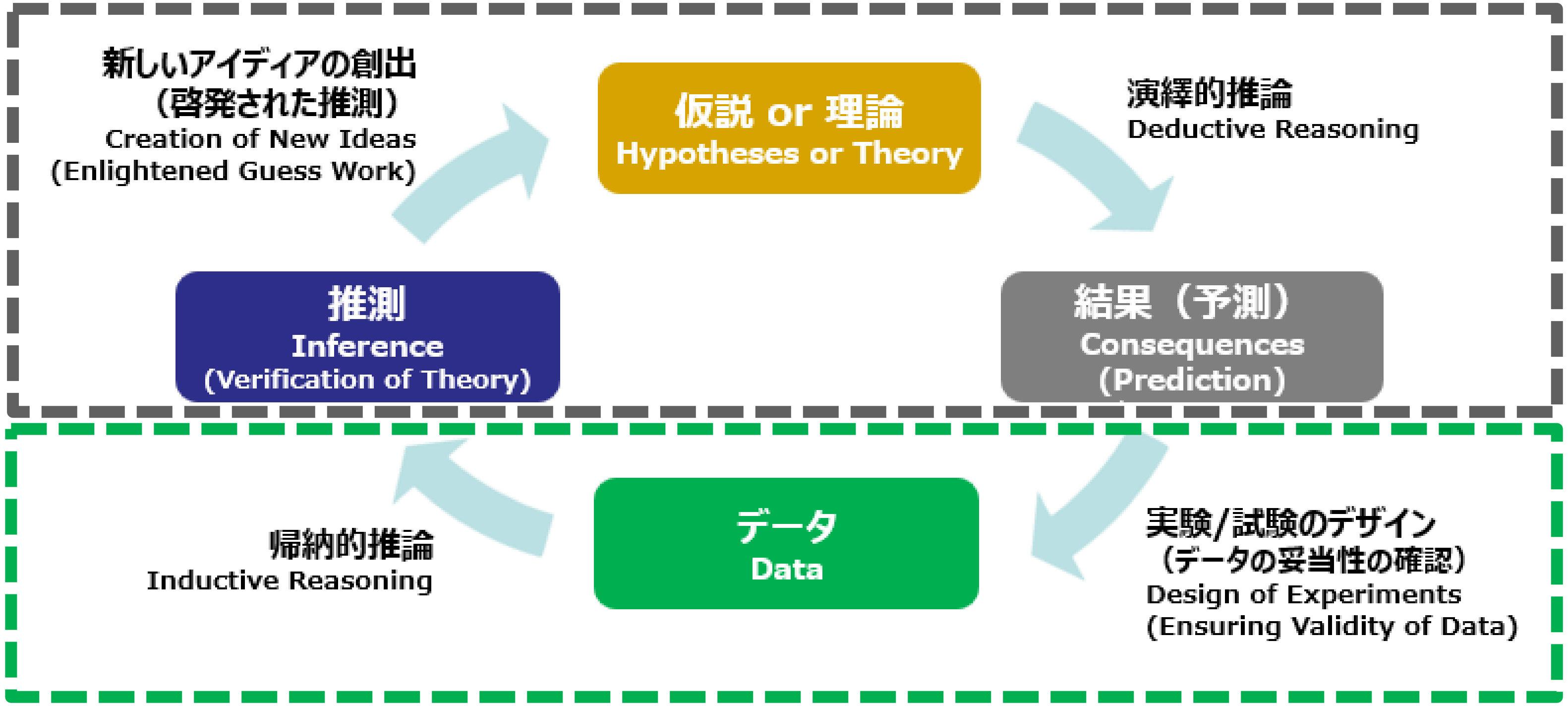
事後確率
Posterior Probability
乳がんであるという仮説
に対する更新された確信

乳がんでないという仮説（～A）
に対する確信も更新されている

科学の学習過程 (C.R.Rao, 1997より)



科学の学習過程の主な役割分担　(C.R.Rao, 1997より)

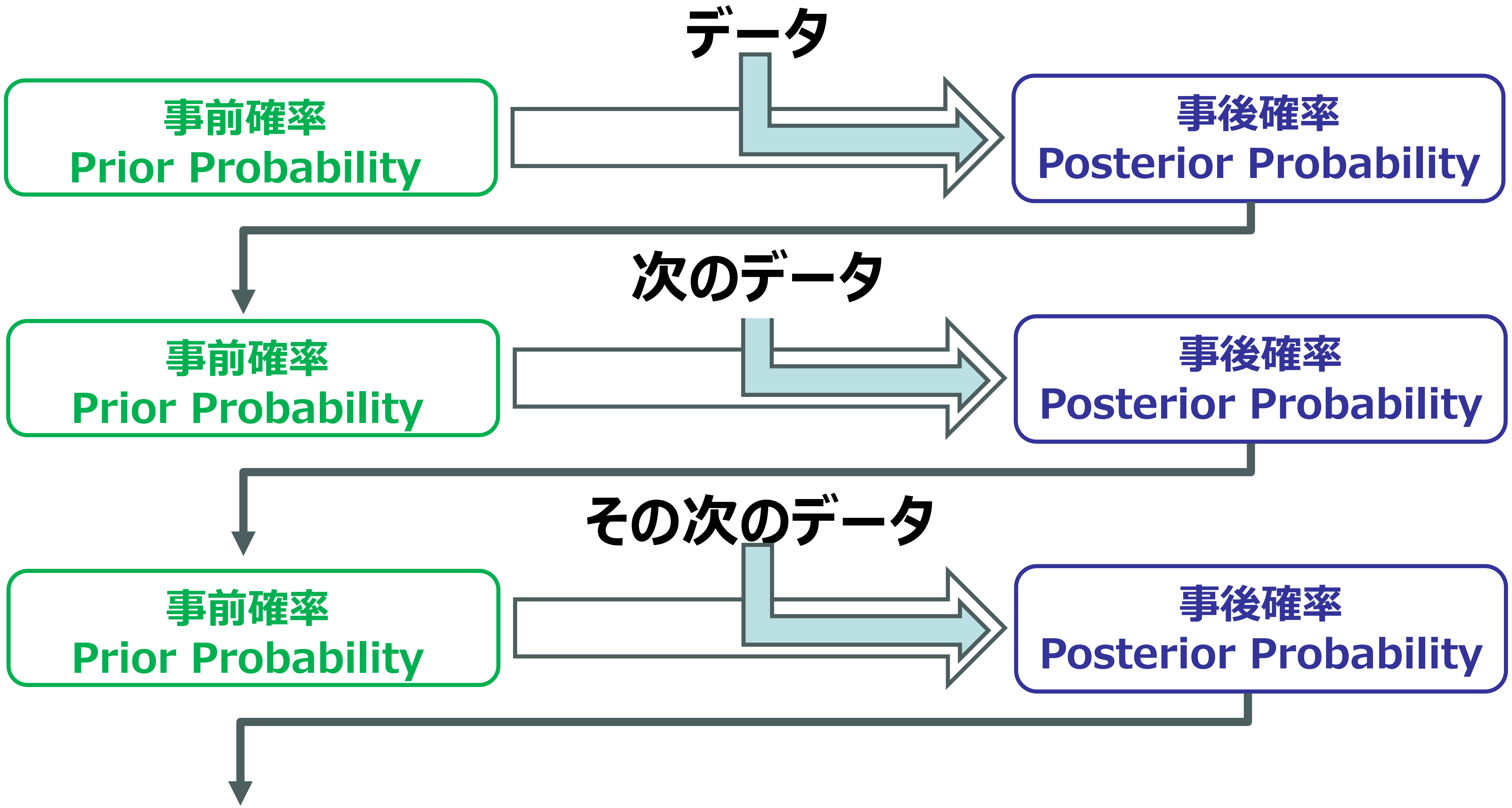


研究テーマと科学者が果たす
創造的役割に依存する部分

統計学の領域

ベイズの定理の解釈２

「今日の 事後 は、明日の 事前。」
“Today’s posterior is tomorrow’s prior.”



ベイズの定理

データBが与えられたときの仮説Aの事後仮説

仮説Aが与えられたときのデータBの尤度

仮説Aの事前確率

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B|A) * \Pr(A) + \Pr(B|{\sim}A) * \Pr({\sim}A)}$$

仮説～Aが与えられたときのデータBの尤度

仮説～Aの事前確率

乳がん·マンモ検査の例では…

仮説は２つでした：「乳がんである」(A)と「乳がんでない」(～A)の２つ

	乳がん(A)	No 乳がん(～A)	合計
マンモ陽性(B)	0.005	0.495	0.5
マンモ陰性(～B)	0.005	0.495	0.5
合計	0.01	0.99	1

仮説Aに関する項

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B|A) * \Pr(A) + \Pr(B|\sim A) * \Pr(\sim A)}$$

仮説～Aに関する項

仮説が２つより多い場合

仮説が *n* 個あって、*i* 番目の仮説の事後確率を求める場合

$$\Pr(H_i|data) = \frac{\Pr(data|H_i) * \Pr(H_i)}{\sum_{j=1}^{n} \Pr(data|H_j) * \Pr(H_j)}$$

$$= \frac{\Pr(data|H_i) * \Pr(H_i)}{\Pr(data|H_1) * \Pr(H_1) + \cdots + \Pr(data|H_i) * \Pr(H_i) + \cdots + \Pr(data|H_n) * \Pr(H_n)}$$

$H_1, H_2, \cdots, H_n$のすべてについて、
事後確率$= \Pr(H_1|data), \Pr(H_2|data), \cdots, \Pr(H_n|data)$
を求めれば data を得た後の事後確率の分布を求めることができる

ベイズの定理　(本資料の冒頭のスライドより)

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)}$$

● ベイズの定理の２つの解釈

[解釈1]条件付き確率Pr(A|B) と
条件付き確率Pr(B|A) の間の関係を示している

[解釈2]エビデンスを考慮して主観的な信念の度合いが、
どのように合理的に変化するはずかを表現している

[解釈2] に
ついてざっくり説明
しました

ベイズの定理

Pr(B|A)は事象Aが起きた上で事象Bが起こる条件付き確率。
英語では、"Probability of B given by A"と読む。

$$\Pr(A|B) = \frac{\Pr(B|A) * \Pr(A)}{\Pr(B)}$$

Pr(A)は事象Aが起きる確率

Pr(B)は事象Bが起きる確率

Pr(A|B)は事象Bが起きた上で事象Aが起こる確率で"条件付き確率"という。
英語では、"Probability of A given by B"と読む。

● ベイズの定理の２つの解釈

[解釈1]条件付き確率Pr(A|B) と 条件付き確率Pr(B|A) の間の関係を示している

[解釈2]エビデンスを考慮して主観的な信念の度合いが、どのように合理的に変化するはずかを表現している

さいごに

- ベイズ統計が頻度論と大きく異なるのが不確実性の扱い方
  - ベイズ流アプローチではあらゆる不確実性が確率によって表現される。
  - 未知の事象には確率分布が存在する。
  - 既知の事象は全て与えられた条件として扱われ、全ての確率は既知の値を条件として計算される。
  - 実験や試験の結果が未知の場合（実験や試験の前）には、それらにも確率が存在する。しかし実験や試験の結果が判明した時点で、それらは既知の事実として扱われ、確率の対象ではなくなる。
- 本日は最も簡単な事例（２Ｘ２の分割表）で説明しました。
  - もっと複雑な、仮説が連続量の確率分布をもつ場合についての説明は時間の都合上、無理でした。この世にたくさん存在するベイズ流統計の入門書で勉強してみてください。きっと今日の話が理解を助けてくれます。