

ベイズ統計学とは？初心者向けのやさしい解説

甲第1225号証

統計学の基礎

ライター：[y0he1](#)[Twitter](#)[Facebook](#)

当ページにお越しいただき、ありがとうございます。
ここでは「ベイズ統計学とはなんたるか」について、初めての方にもわかるように解説していきます。

・頻度論との違いはこちら⇒ [ベイズ統計学の考え方～ベイズ論と頻度論の違い～](#)

・ベイズ統計学のトップページはこちら⇒ [ベイズ統計学のわかりやすいまとめ](#)

ベイズ統計学とは、ベイズの定理をもとにした統計的な考え方の一種です。

目次 [\[非表示にする\]](#)

1 ベイズ統計学の独特な考え方

1.1 ベイズでは主観確率を扱う

1.2 ベイズの定理(Bayes' theorem)の式

1.3 ベイズ統計学が推計統計学より優れている点

3 まとめ

2 なぜいまベイズ統計学は注目されているのか

2.1 ベイズ統計学の歴史

2.2 機械学習とベイズ統計学

2.3 ビッグデータとベイズ統計学

ベイズ統計学の独特な考え方

まず読者の皆さんにわかってほしいのは、「ベイズ統計学は特別！」だということです。

確率の考え方が他と違って特殊なのです。特に、[記述・推計統計学](#)とは大きく異なった考え方をします。そのため、ベイズ統計学を支持する人は特別

に"ベイジアン"と呼ばれ、推計統計学論者と今日に至るまで激しく対立しています。

では、どこがどう異なるのでしょうか。まずは、それぞれの特徴をざっと書いておきます。

記述統計学……**標本**に見られる特徴をわかりやすく表す。

推計統計学……**標本**を分析して、**母集団**について推測する。

ベイズ統計学……**標本**を必ずしも必要としない。データ不十分でも何とかして確率を導く。

ベイズでは主観確率を扱う

確率には、客観確率、主観確率があります。

例えば「Q：サイコロで6が出る確率は？」と聞かれたら $\frac{1}{6}$ と、誰もが即答できますよね。これを客観的な確率といいます。誰が見ても同じ答えになります。これが世に言う一般的な確率になります。

では、主観確率とはどのようなものなのでしょう
か？ここで分かりやすさのために、例を使って考えて見ます。

例

Q：電車で、隣に座っているサラリーマンの
おじさんが転職を考えている確率は？

「そんなのわかるわけじゃないか」と思った人、その通りです。この場合、人それぞれ違う答えになりますよね。 $\frac{1}{3}$ だと思う人もいれば $\frac{1}{100}$ と言う人もいるでしょう。

極端な例を上げましたが、これが、**主観的な確率**です。ベイズ確率は、このような主観確率を使っています。そして、何か新しい情報を得るたびにその確率を更新（アップデート）していきます。

今回の例では、「転職を考えている」と「転職を考えていない」の2通りが結果として考えられます。始めはなんの情報もないので、例えばどちらも50%と設定しておきましょう。これをベイズ統計学では**事前確率（prior probability）**と呼びます。

ここで、おじさんのことを観察していたら、電車内の転職広告をずっと見ていたので、70%だと考え直したとします。おじさんが転職広告を見ているとわかった事の後に設定した確率なので、これを事後確率(**posterior probability**)と呼びます。

さらに考えを進めてみましょう。今度は勇気を振り絞っておじさんに転職を考えているかどうかを質問して見ます。

私：「おじさんは転職を考えていますか？」

おじさん：「考えていません。」

本人否定により、確率を10%に設定し直します。ここでは質問という事を行ったので、その前の確率が事前確率、その後の確率が事後確率になります。

事前確率：70% 事後確率：10%

ということです。

このような具合に、ベイズ統計学では最初に確率を設定しておき、情報が入るたびに、"その時点での確率"を変更していきます。つまり、現時点でなんのデータもなかったとしても、とりあえず確率を設定しておくことができるのです。

事前確率を設定した段階から、何か新たな情報をゲットして、事後確率を更新していきます。そして、また新たな情報を手に入れたら、前回は事後確率だったものが事前確率となって、さらに確率の更新が行われるという仕組みです。

このようにベイズ統計学の考え方には、いわば**学習能力**があるということです。

さて、今の例ではテキトーに確率を変更していきましたが、実際にもこのように決めていくのでしょうか。もちろんそんなことはないのでご安心を。ただ、最初に決める確率だけは主観的なもので適当な値に決めておきます。（何の情報もないときは通常 $\frac{1}{2}$ とする）

(また、事前に確率分布を設定することも出来ます。これを事前分布と言います。事前分布は事後分布の計算を簡単にするために、共役事前分布というものを使うことがありますが、それについては[【ベイズ統計】共役事前分布とは？わかりやすく解説](#)で詳しくまとめたので、参考にさせていただきたいと思います。)

事前分布や事前確率を設定した後のアップデートによる変更はベイズの定理に基づいて行います。ではその定理について見てみましょう。

ベイズの定理(Bayes' theorem)の式

[ベイズの定理](#)の式は以下のようになります。

$$P(A|X) = \frac{P(X|A)P(A)}{P(X)} = P(A) \times \frac{P(X|A)}{P(X)}$$

$P(A|X)$:事象Xが起こった状況下で事象Aが起こる確率 (事後確率)

$P(A)$:事象Aが起こる確率 (事前確率)

$P(X)$:事象Xが起こる確率

$P(X|A)$:事象Aが起こった状況下で事象Xが起こる確率

このように非常にシンプルなものですが、これでは何が何やらという人もいますのでこの式を言葉で表します。

$$\text{事後確率} = \text{事前確率} \times \frac{\text{ある場合においての、そのデータが得られる確率}}{\text{そのデータが得られる確率}}$$

となります。今回のおじさんの転職の例でいうと転職広告をずっと見ていると感じた時には、

$$\text{事後確率} = \text{事前確率} \times \frac{\text{おじさんが転職を考えていた場合、転職広告をずっと一般的におじさんが転職広告をずっと見ている}}$$

ということになります。

このベイズの定理の式の導出(証明)と詳しい考え方はこちら。

⇒ [ベイズの定理の導出と考え方をわかりやすく解説](#)

ベイズ統計学が推計統計学より優れている点

ベイズ統計学の考え方は、人の考え方によく似ています。例えば、新しい情報を得たら、今までの考えを修正して、新しい考えを持ちますよね。これがベイズ統計学における事前確率、事後確率の関係にそっくりなのです。

推計統計学の考え方では、今ある標本データをまとめて解析します。そのため、新しい標本データが追加で得られた場合、再び解析は1からやり直します。

極端な例を挙げると、

1000万人の身長データを膨大な時間かけて解析しました。でもここで、新たに100人の身長データを手に入れたので、それも含めて解析したい。そうすると、また1000万100人のデータでもって、もう一度膨大な時間をかけて解析する必要がありますよね。

非常に、効率の悪い作業です。

対して、ベイズ統計学を使えば新たなデータを入手したら、その分の修正を加えればいいだけです。

なぜいまベイズ統計学は注目されているのか

ベイズ統計学ですが、その考え方の基である、ベイズの定理は1700年代からありました。では、なぜ近年急激に注目され始めたのでしょうか。

ベイズ統計学の歴史



その話をするために、まずは歴史を追っていきましょう。ベイズ統計学は、1700年中頃にトーマズ・ベイズによる、ベイズの定理の発表により、産声をあげました。その後、1800年代後半に再び現在のベイズ統計の考え方の基礎となる考え方をする人々が現れました。

しかし、推計統計学論者のフィッシャーらが、「主観確率を扱うのは科学的でない」とし、ベイズ統計学は闇に葬り去られてしまったのです。

ですが、科学的であるかないかは別として、ベイズ統計学は現実役に立つ学問です。そのことが徐々に分かってきて、1950年代に入り再び研究され注目を浴びるようになりました。

また、ここ最近、再びよく使われるようになったのは「機械学習」「ビッグデータ」の分野で有効利用出来たからです。

機械学習とベイズ統計学

機械学習とは、コンピューターが得られたデータから法則性を自ら見つけ出し、それを利用して将来を予測する機能のことを言います。

つまり、コンピューターが学習能力を持って、ますます頭が良くなるということです。映画などで見る、実に近未来的な話ですね。その、技術がベイズ統計学でもってますます発展しているのです。

現在、機械学習は迷惑メールやスパムメールの判別で役立っています。"ベイジアンフィルター"と呼ばれるメールフィルターを採用しています。

このフィルターは、日々変化する多種多様な迷惑メールにうまく対応していきます。

事前に迷惑メールについての定義を作っておき、迷惑メールとして分別されたメール（または、ユーザーが迷惑メールフォルダーに移動したメール）から法則を見出します。その見出された基準を元に、新

たに受信されるメールについて調べます。これにより、日々変化する迷惑メールに対応して、正しい判別をすることができるのです。

ビッグデータとベイズ統計学

皆さん、そもそもビッグデータという言葉をご存じでしょうか。ビッグデータとは、世の中のいろいろなデータの集まりを指します。これが、膨大なデータ量なのでビッグという言葉で表します。

イメージが沸きにくいかもしれないので、簡単な例をいくつか挙げてみます。

Snsのつぶやき 監視カメラの映像・音声 電力メーター 電子カルテ ネット通販の購入履歴 ICカード 利用情報 カーナビのGPS スマートフォン 環境気象データ 電車の運行情報

などです。

この、ビッグデータ解析の大きな目的の一つはマーケティングです。

均一性や普遍性のないマーケティングの世界では、多角的かつ柔軟な対応が望まれます。人はそれぞれが全く違った生活基盤を持っており、その生活の行動の中で消費が生まれます。そのなかで、マーケット調査をしようとするならば、人々が普段何気なく行動している事柄を分析しなくてはなりません。

それを分析するときに、従来のように「人々はもともと～のような性質をしていて、それはどのような確率分布に従い、何人から調査を得れば正しい結果を得られて～」というような議論をしてはきりがありません。

ベイズ統計学のように、人々がある行動をしたら、事前確率を事後確率に置き換えて分析するというような柔軟な分析の仕方がビッグデータを分析する上で非常に役に立つのです。

まとめ

ベイズ統計学についていかがでしたでしょうか。ベイズ統計学が認められ、広く知られるようになってからまだ50年ちょっとしか経っていません。にもかかわらず、様々な分野に応用され、まさに伸び盛りの分野だと言えます。

ベイズ統計学の有用性が認められている一方で、[推計統計学](#)論者とベイジアンの方の考え方の違いによる対立は今日でも続いています。互いに口もきかないという噂も、、、

さて、あなたはどっち派になるのでしょうか？笑

参考記事

・ベイズ統計学のトップページはこちら⇒[ベイズ統計学のわかりやすいまとめ](#)

・ベイズ論と頻度論の違いをこちらのページで、具体例を交えつつ、詳しくまとめました。

⇒ [ベイズ統計学の考え方～ベイズ論と頻度論の違い～](#)

・ベイズの定理についてさらに詳しく知りたい方はこちら。

⇒ [ベイズの定理の導出と考え方をわかりやすく解説](#)

・正規分布に従う母集団からデータを得るとき、事後分布の平均、分散の導出と、そこから得られる性質はこちら。

⇒ [正規分布の事後分布の平均と分散【ベイズ】](#)

・ベルヌーイ分布に従う分布からデータを取ってきたときの事後分布の平均、分散の導出とその性質こちら。（データがベルヌーイ分布に従うとき、共役事前分布はベータ分布になります。）

⇒ [ベータ分布の事後分布の平均と分散【ベイズ】](#)