情報学研究科 M1 下條朝也

2017/04/25

Optimal Predictions in Everyday Cognition

Thomas L. Griffiths and Joshua B. Tenenbaum

Psychological Science, Vol.17, No.9, 2006, Pp.767-773

スライド URL：https://www.slideshare.net/secret/rirtUxMRY42VxI

◆ SUMMARY

- 目的：人間の連続的な確率推論がベイズの定理に一致するか。
- 方法：既存データの確率分布を事前分布として事後分布を予測。その値が人間の判断した結果と一定以上一致するかを確かめた。
- 結論：ベイズの定理に基づく予測値と人間の推論結果とはおおよそ一致した。
- Tenenbaum が記した他の論文と異なる点（新規性）

  対象：因果帰納（causal induction）でなく，連続的な数の推論である。

  ex)

  - 因果帰納（Griffith & Tenenbaum, 2007）の例

    「自分は超能力者だ」と主張する友人がいたとする。それが本当かどうかを確かめるため，コイントスを 10 回行い，何回表が出たら友人を信用するか？

    確かに，偶然で 10 回連続表が出る確率は低いが，そもそも人間が超能力を持っている可能性はもっと低い。その結果，偶然であると判断する。

◆ INTRODUCTION

- 日常における連続量に対する推論

  ex)

  - 60 歳の男性と聞いて，残りの寿命はどれくらいだと予測するか？
  - 現在 4000 万ドルの興行収入を誇る映画では，最終的な収益はいくらになりそうか？

  上記のように，限られたデータしか得られていないため正しい答えがわからない状況に対しても，少なくとも最低限合理的な推測が行われる。

  →不確実状況下での認知的判断は事前確率に影響されない（ヒューリスティクスに依存する）とする Tversky & Kahneman,（1974）の考え方には否定的。

- ベイズの定理の例

  ex)

  - 式　$p(t_{total} | t) \propto p(t | t_{total}) * p(t_{total})$　の解説

    変数：t は現在の年齢。t_{total} は総寿命。

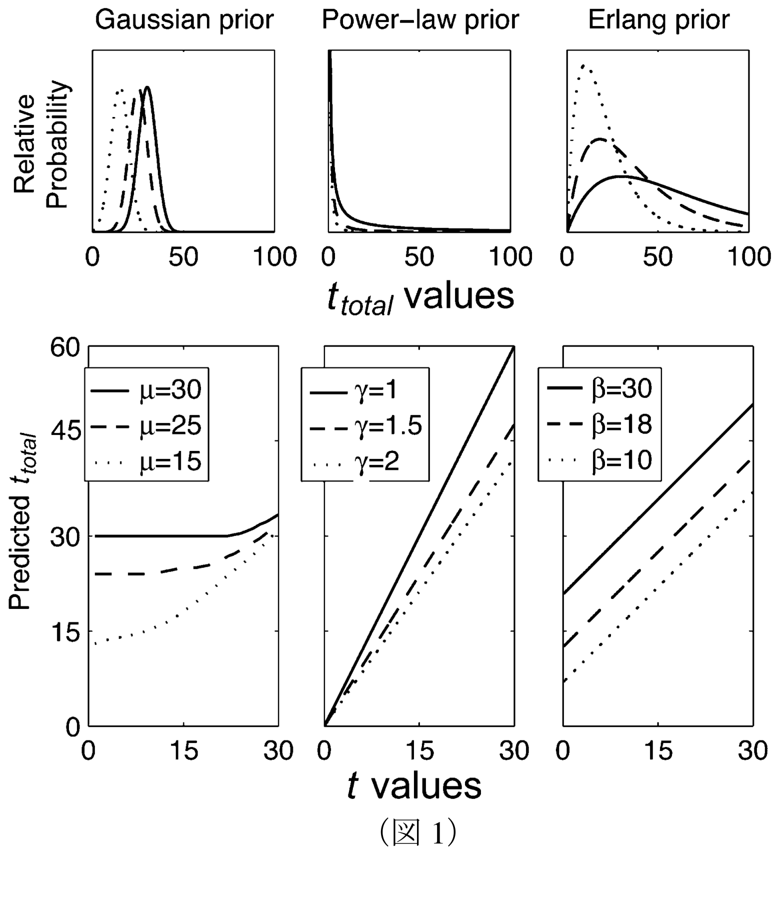
    $p(t_{total} | t)$：確率分布。現在年齢 t の下で総寿命が t_{total} である確率。

$p(t\,|\,t_{total})$：尤度。年齢 t の男性に出会う確率。今回は，どの年齢の人に出会う可能性もすべて等しいと仮定して，$p(t\,|\,t_{total}) = 1/\,t_{total}$ とする（実験では乱数を用いるため）。

$p(t_{total})$：事前確率。その事象に対する一般的な期待を表す。今回では，人間の平均寿命（75 歳，SD=16 年）に基づく。

つまり，出会ったときに t 歳である人が t_{total} 歳まで生きる可能性は，t 歳の人に出会う確率 ×（1/平均寿命）に比例する，という式。

t_{total} の良い予測値は，確率分布 $p(t_{total}\,|\,t)$の中央値に近いもの。t が平均よりずっと小さい場合，t_{total} の予測値はほぼ平均となる。t が平均値に近づくにつれて，t_{total} の予測値はゆっくりと増加し，t がさらに増加すると高く維持される（図 1）。また，事前分布として，左から順に正規分布，冪乗則分布，Erlang 分布を用いた（冪乗則分布と Erlang 分布は $p(t_{total} > t1\,|\,t) = 0.5$ となる値 t1 を中央値とした）。



（図 1）

- ◆ METHOD
  - ➢ Participants and Procedure

    学部生 350 人を 2 グループ（208 人と 142 人）に分けた。

    各グループは，5 つの異なる現象について予測を行った。

208 人の学部生からなる 1 グループは，映画の興行収入，詩の長さ，寿命，ファラオの統治期間，結婚期間の長さについての予測をした。142 人の学部生からなる第 2 のグループは，映画の上映時間，米国代表の任期期間，ケーキを焼く時間，行列の待ち時間，結婚期間の長さに関する予測を行った。

- Materials

各参加者に対して， グループごとに 5 つの事象が出題された。また，その際に提示される事前確率は 5 水準で設定され，被験者間でランダムに 1 水準提示した。

- 映画の興行収入：
  $ 100 万，$ 600 万，$ 1000 万，$ 4000 万，$ 1 億
- 詩の長さ：
  2 行，5 行，12 行，32 行，67 行
- 人間の寿命：
  18 歳、39 歳、61 歳、83 歳、96 歳
- ファラオの統治期間：
  1 年，3 年，7 年，11 年，23 年
- 結婚の期間（両グループに対して出題）：
  1 年，3 年，7 年，11 年，23 年
- 映画の上映時間：
  30 分，60 分，80 分，95 分，110 分
- 米国代表の連続任期期間：
  1 年，3 年，7 年，15 年，31 年
- ケーキを焼く時間：
  10 分，20 分，35 分，50 分，70 分
- 行列の待ち時間：
  1 分，3 分，7 分，11 分，23 分

いずれの場合も，参加者は文脈を確立するいくつかの文章を読んだ後，t_{total} を予測するよう求められた。

質問は調査形式で提示した。各質問は、期間または数量のどちらかを求めるように指示。

「以下の各質問を読んで，その下の行にあなたの予測を書いてください。私たちはあなたの直感に興味があるので，複雑な計算はしないでください。あなたの考えを教えてください！」

- 映画の興行収入：
  「1000 万ドルの興行収入を誇る映画について聞いたことがあるのですが，その映画がどれくらいの期間上映されているかは分かりません。その映画の興行収入の合計

金額を，あなたはいくらだと予測しますか？」

- 詩の長さ：

  「あなたの友人が，あなたの好きな詩を読んで，いま詩の 5 行目を読んでいると言われたら，詩の長さはどれくらいだと思いますか？」

- 人間の寿命：

  「ある保険会社が，人口統計情報に基づいて，人間の寿命に関する予測を行うためにアクチュアリーを採用しています。あなたがとある 18 歳の男性の保険事件を評価を担当する場合，彼の寿命はどれくらいだと予測しますか？」

◆ RESULTS

➢ 分析方法

  ✧ 回答の中央値と事後分布の予測値とを比較

➢ 各項目の回答総数と除外したもの

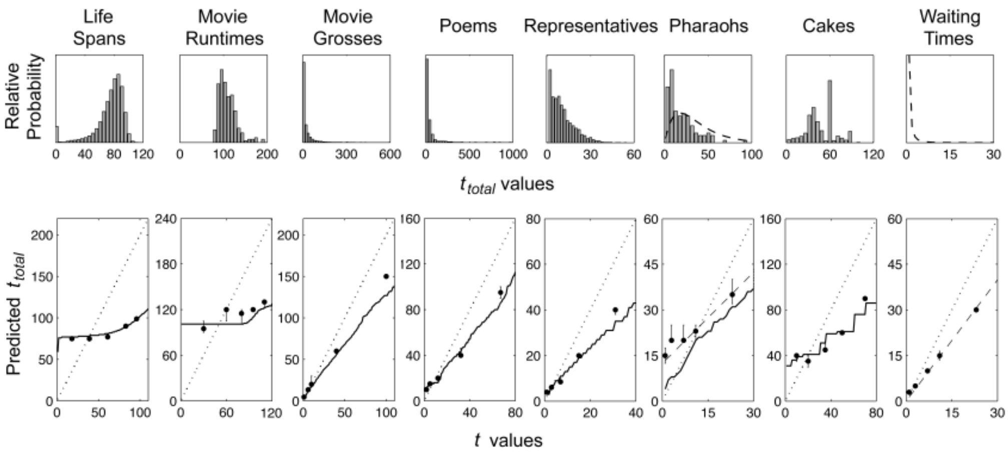
  ✧ 映画の興行収入 174 人，詩の長 197 人，寿命 197 人，ファラオ支配 191 人，映画の上映時間は 136 人，米国代表は 130 人，ケーキを焼く時間は 126 人

  ✧ 参加者の過半数（52%）が，結婚の期間（この夫婦の結婚生活はいつまで続くか？）の問に対して「永遠に続く」と回答したため，結婚の刺激は不適切であったと言えるので分析は行わなかった。また，この「永遠に続く」という回答は，離婚しない夫婦の割合を正確に反映している（Kreider＆Fields, 2002）。

➢ 利用した関数

  人間の寿命と映画の上映時間はおよそ正規分布であり，映画の総興行収入と詩の長さはほぼ冪乗則であり，米国代表の任期やファラオの統治期間については Erlang を利用した。この実験では，これらの異なる設定における最適な統計的推測（実際のデータに一番近くなる関数による分布）に，人々の予測がどれほど対応しているかを調べた（人間のデータに合うように調整されたパラメータは設定しなかった)。その結果が図 2 である。

➢ 実線はベイズの定理による予測値，黒点は被験者の t_{total} の中央値，エラーバーは 68%信頼区間（標本平均±1*標準誤差）で描画した。



（図２）

- これらの結果は，認知的判断が事前確率に鈍感な非ベイジアンヒューリスティクスに基づいているという主張（Kahneman et al, 1982; Tversky＆Kahneman, 1974）と一致しなかった。
- また，現象の種類にかかわらず，単一の情報がない過去の $p(t_{total}) \propto 1 / t_{total}$ を採用する単純なベイジアン予測モデルと一致しなかった（Gott, 1993, 1994; Jaynes, 2003; Jeffreys, 1961; Ledford et, 2001）。
- ファラオの統治期間，ケーキを焼く時間，行列の待ち時間の結果を見ると，人々の予測能力の限界を知ることができる。
  - ファラオの治世に関する人々の予測は，適切な事前確率（Erlang 分布）と一致する形をしていたが，やや高すぎる結果となった（図 2)。
    - 私たちは追実験として，ファラオの治世の典型的な期間を述べるように 35 人の学部生に求め，ファラオ支配のための人々の主観的な事前確率を調べた。すると，平均値は 30 年となり，正しい値である 9.34 年からは大きく離れる結果となった（平均 17.9 である Erlang の方がまだ近い)。人間の主観と近い Erlang を使用することは，人間の判断と密接に対応していると言える。
    - ファラオの刺激は，人々が事前分布の適切な形式（どの程度の長さが妥当か）を知っているが，その詳細については知らない場合，人々が不正確な予測をすることになることを示している（結果的に，ベイズによる予測を全体的に上回る結果となった)。対照的に，ケーキの刺激に対する反応は，人々が簡単な形を欠いている状況においても正確な予測をすることができることを示している。ケーキを焼く時間は，図 2 であるように，不規則な分布に従っている（段階ずつ向上している)。しかし，経験的な事前分布があるにもかかわらず，人々の判断はベイジアン予測と近い。

◆ DISCUSSION

- 日常的な出来事に対する人々の予測は全体的にはきわめて正確。しかし，被験者の回答は，問題によって大きく変化し，その結果はベイズの定理を用いて予測していた事前分布の形と対応

していた。

- また，事前確率の予測から逸脱した場合も，直感的な判断を成功させる例（ex.ファラオの統治期間）も見られた。
- ベイジアンモデルは，対象の一般化を行う単語や概念の学習（Tenenbaum, 1999）や，論理的または因果関係を発見するために提案されてきた（Anderson, 1990; Shepard, 1987; Tenenbaum＆Griffiths, 2001; Griffiths＆Tenenbaum, 2006; Oaksford＆Chater, 1994）。しかし，これらの取り組みは実環境において最適なものではなかった。一方，今回の結果は，少なくとも日常的な予測課題の範囲で，人々は効果的に，世界の関連する出来事の統計に合わせて事前分布を適用する（できる）ことを実証している。