Cognitive Science and Bayes

To Bayes or not to Bayes

Steven Rekké

Radboud University Nijmegen

07-04-'09

The main question

Are people 'Bayesian estimators'?

Or more precisely: should cognitive judgments be viewed as following optimal statistical inferences or following error prone heuristics that are insensitive to priors? (Like how human perception and memory are often explained.)

Several views

- The 'original' Bayesian view

  This view (approximately) states that people are conservative Bayesian estimators and their deviations from the norm are attributed to misperception of the impact of each datum, misaggregation of the joint impact of data, or to a response bias against extreme estimates. (Edwards, 1968)
  According to this approach, Ss' estimates are assumed to be qualitatively compatible with the normative model: they merely exhibit a conservative bias that needs to be explained.

- The heuristic view (Kahneman & Tverksy)
- The 'new' Bayesian view (Griffiths & Tenenbaum)

Example 1

1. A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50% of all babies are boys. The exact percentage of baby boys, however, varies from day to day. Sometimes it may be higher than 50%, sometimes lower.

For a period of 1 year, each hospital recorded the days on which (more/less) than 60% of the babies born were boys. Which hospital do you think recorded more such days?

	(More than 60%)	(Less than 60%)
The larger hospital	(12)	(9)
The smaller hospital	(10)	(11)
About the same (i.e., within 5% of each other)	(28)	(25)

Example 1 (answer)

1. A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50% of all babies are boys. The exact percentage of baby boys, however, varies from day to day. Sometimes it may be higher than 50%, sometimes lower.

For a period of 1 year, each hospital recorded the days on which (more/less) than 60% of the babies born were boys. Which hospital do you think recorded more such days?

	(More than 60%)	(Less than 60%)
The larger hospital	(12)	(9)*
The smaller hospital	(10)*	(11)
About the same (i.e., within 5% of each other)	(28)	(25)

Example 2

A well-known example:
In a class of 23 people, what is the probability that at least two of them have the same birthday? (i.e., same day and month)

Example 2 (answer)

A well-known example:
In a class of 23 people, what is the probability that at least two of them have the same birthday? (i.e., same day and month)

This probability exceeds .5 !

Example 3

But a more 'realistic' Question might be:
"How much longer do you expect a 60-year-old man to live?"

Heuristics view: Representativeness

Kahneman & Tverksy (1972) explore a heuristic called representativeness according to which the subjective probability of an event, or sample, is determined by:

- Similarity of Sample to Population
- Reflection of Randomness

Similarity of Sample to Population (1/2)

All families of six children in a city were surveyed. In 72 families the exact order of births of boys and girls was G B G B B G.

What is your estimate of the number of families surveyed in which the exact order of births was B G B B B B?

The sequences are about equally likely, but BGBBBB fails to reflect the proportion of boys and girls in the population. 75 of 92 Ss judged BGBBBB to be less likely ($p < .01$).

Similarity of Sample to Population (2/2)

There are two programs in a high school. Boys are a majority (65%) in program A, and a minority (45%) in program B. There is an equal number of classes in each of the two programs.

You enter a class at random, and observe that 55% of the students are boys. What is your best guess—does the class belong to program A or to program B?

Since the majority of students in the class are boys, the class is more representative of program A than of program B. Accordingly, 67 of 89 Ss guessed that the class belongs to program A ($p < .01$ by sign test). In fact, it is slightly more likely that the class belongs to program B (since the variance for $p = .45$ exceeds that for $p = .65$).

Reflection of Randomness

To be representative an uncertain event should also appear random. Regularities in order or distribution and are deemed relatively unlikely:

- HTHTHTHT or TTHHTTHH for a coin toss.
- Random distribution (of 20 marbles to 5 people) nr II below:

	I		II
Alan	4	*Alan*	4
Ben	4	*Ben*	4
Carl	5	*Carl*	4
Dan	4	*Dan*	4
Ed	3	*Ed*	4

Studies: Prediction

The main prediction Kahneman and Tversky discuss about subjective probabilities is that according to their Representativeness heuristic, sample size is not going to influence the subjective probabilities.

(Since the size of the sample does not reflect any property of the parent population, it does not affect representativeness.)

Studies: Anecdote

People often remain skeptical in the face of solid evidence from a large sample, as in the case of the well-known politician who complained bitterly that the cost-of-living index is not based on the whole population, but only on a large sample, and added, 'Worse yet - a random sample.' (Kahneman & Tversky, 1972)

Study 1: Sampling Distributions (1/2)

On what percentage of days will the number of boys among 1000 babies be as follows:
Up to 50 boys
50 to 150 boys
150 to 250 boys
. . . .
850 to 950 boys
More than 950 boys
Note that the categories include all possibilities, so your answers should add up to about 100%.

Study 1: Sampling Distributions (2/2)

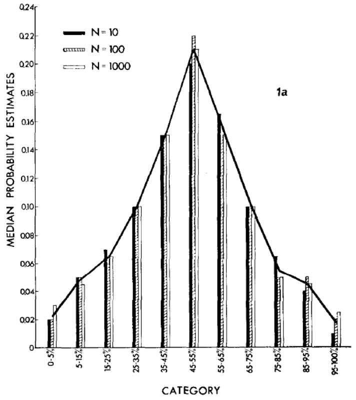


Figure 1a

PROBABILITY AND REPRESENTATIVENESS

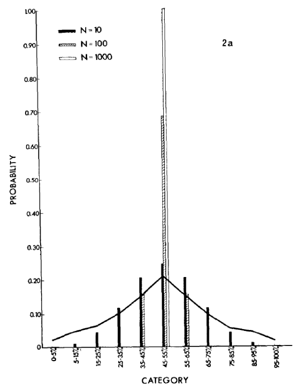


Figure 2a

Study 2: Posterior Probability

The average heights of adult males and females in the US are, respectively, 5 ft 10 in. and 5 ft 4 in. Both distributions are approximately normal with a standard deviation of about 2.5 in.

An investigator has selected one population by chance and has drawn from it a random sample.

What do you think are the odds that he has selected the male population if

(i) the sample consists of a single person whose height is 5 ft 10 in.?

(ii) the sample consists of 6 persons whose average height is 5 ft 8 in.?

The median subjective odds were 8 in case (i) and 2.5 in case (ii). Indeed, the significant majority of Ss (86 out of 115) assigned a higher value to the former case ($p < .01$ by a median test). The correct odds are 16 in case (i) and 29 in case (ii). The responses of the Ss, therefore, are not merely conservative—they violate the correct ordering of likelihoods. Here again, it appears that *Ss* base their judgments on sample mean with insufficient concern for sample size.

Conclusion

It appears from the results presented in the paper that:

- Sample size has no effect on subjective sampling distributions.
- Posterior binomial estimates are determined by sample proportion rather than by sample difference
- ..and that they do not depend on the population proportion.

From these results, Kahneman and Tversky conclude that in the evaluation of evidence, people are apparently no conservative Bayesians, they are no Bayesians at all.

The Bayesian view

Griffiths and Tenenbaum (2006) take a different standpoint than Kahneman and Tversky.

Their results suggest that everyday cognitive judgments follow the same optimal statistical principles as perception and memory, and reveal a close correspondence between people's implicit probabilistic models and the statistics of the world.

The Study (1/3)

In their study they asked subjects to predict life spans, movie runtimes, movie grosses, poem lengths, terms of US representatives, reigns of pharaohs, baking time for cakes and waiting times (on the phone).

An example question is:
If your friend read you her favorite line of poetry, and told you it was line 5 of a poem, what would you predict for the total length of the poem?

The Study (2/3)

Prior distributions were obtained from the Internet.

Sources of Data for Estimating Prior Distributions

Data set	Source (number of data points)
Movie grosses	http://www.worldwideboxoffice.com/ (5,302)
Poem lengths	http://www.emule.com/ (1,000)
Life spans	http://www.demog.berkeley.edu/wilmoth/mortality/states.html (complete life table)
Movie run times	http://www.imdb.com/charts/usboxarchive/ (233 top-10 movies from 1998 through 2003)
U.S. representatives' terms	http://www.bioguide.congress.gov/ (2,150 members since 1945)
Cake baking times	http://www.allrecipes.com/ (619)
Pharaohs' reigns	http://www.touregypt.com/ (126)

Note. Data were collected from these Web sites between July and December 2003.

The Study (3/3)

The total number of responses analyzed was 174 for movie grosses, 197 for poem lengths, 197 for life spans, 191 for reigns of pharaohs, 136 for movie run times, 130 for terms of U.S. representatives, 126 for baking times for cakes, and 158 for waiting times.

Bayesian model: Basic principles

We're trying to predict the total life span of a man we have just met, on the basis of the mans current age. So:

- t_{total} indicates the total amount of time the man will live
- t indicates his current age
- Task: estimate t_{total} from t

Bayes' rule

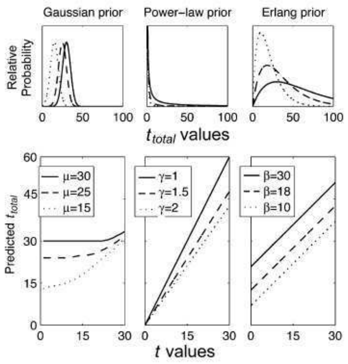
$$p(t_{total}|t) = \frac{p(t|t_{total}) * p(t_{total})}{p(t)}$$

We assume for simplicity that we are equally likely to meet a man at any point in his life, this probability is uniform, $p(t|t_{total}) = \frac{1}{t_{total}}$, for all possible values of t between 0 and t_{total} (and 0 for values outside that range).

Prediction

This yields a probability distribution $p(t_{total}|t)$ over all possible total life spans t_{total} for a man encountered at age t. A good guess for t_{total} is the median of this distribution.

Taking the median of $p(t_{total}|t)$ defines a Bayesian prediction function.



Simplification (1/2)

$$p(t_{total}|t) = \frac{p(t|t_{total})p(t_{total})}{p(t)} \qquad (A1)$$

where

$$p(t) = \int_0^\infty p(t|t_{total})p(t_{total})dt_{total} \qquad (A2)$$

By the assumption that t is sampled uniformly at random, $p(t|t_{total}) = 1/t_{total}$ for $t_{total} \geq t$ and 0 otherwise. Equation A2 thus simplifies to

$$p(t) = \int_t^\infty \frac{p(t_{total})}{t_{total}} dt_{total} \qquad (A3)$$

So the shape of the posterior distribution for any given value of t is determined entirely by the prior $p(t_{total})$.

Simplification (2/2)

Now we can derive an analytic form of the posterior distribution (except for the Gaussian) to get:

$$p(t_{total}|t) = \frac{t_{total}^{-(\gamma+1)}}{\frac{1}{\gamma}t^{-\gamma}} = \frac{\gamma t^{\gamma}}{t_{total}^{\gamma+1}},$$

for power-law and:

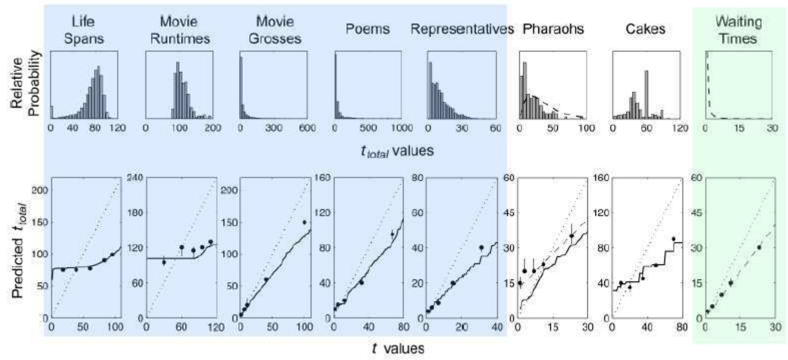
$$p(t_{total}|t) = \frac{\exp\{-t_{total}/\beta\}}{\beta\exp\{-t/\beta\}} = \frac{1}{\beta}\exp\{-(t_{total}-t)/\beta\},$$

for erlang.

Finally

As mentioned before, t_{total} is now predicted by taking the posterior median. (Solving $p(t_{total} > t_{predicted}|t) = 0.5$)

Results (1/2)



Results (2/2)

These results are inconsistent with claims that cognitive judgments are based on non-Bayesian heuristics that are insensitive to priors (Kahneman et al.).

The results are also inconsistent with simpler Bayesian prediction models that adopt a single uninformative prior:

$$p(t_{total}) \propto \frac{1}{t_{total}}$$

regardless of the phenomenon to be predicted (e.g. Gott, 1993, 1994; Jaynes, 2003).

Conclusion

Griffiths and Tenenbaum conclude that the results reveal a far closer correspondence between optimal statistical inference an everyday cognition than previous research, since people's judgments were close to the optimal predictions produced by their Bayesian model.

(But what is 'close'?)

Furthermore, this shows that, at least for a range of everyday prediction tasks, people effectively adopt prior distributions that are accurately calibrated to the statistics of relevant events in the world.

Questions / Discussion

Several questions come to mind:

- Are these views really at odds with each other?
- If they are, how can they both have empirical backup?
- If they are not, to what cases does which view apply ?
  Could it be so simple as the distinction between laboratory studies and 'everyday' cognition?

Do you have suggestions? Questions?