Pragmatics, probabilities & psychologists: A Bayesian perspective on some reasoning problems

Dani Navarro
School of Psychology
University of New South Wales
compcogscisydney.com
Contributors, coauthors, collaborators

Keith Ransom
Wouter Voorspoels
Steph Banner
Ally Tingey
Jessica Lee

Amy Perfors
Brett Hayes
Drew Hendrickson
Peter Lovibond
Michelle Keshwa
Matthew Welsh
Matt Dry
Michael Lee
Chris Donkin

Kristy Martire
Ben Newell
Wai Keen Vong
Lauren Kennedy
Steve Langsford
Candy Liu
Anastasia Ejova
Rachel Stephens
Gert Storms
Pat Shafto
Baxter Eaves
Charles Kemp
Nancy Briggs
Titia Benders
Sean Tauber

Funding
What are the connections between human reasoning and statistical inference?
What should we do with this *sample* of evidence?

These birds have plaxium blood
The problem of inductive generalisation
What factors shape our inductive inferences?

Similarity and typicality of the sample
What factors shape our inductive inferences?

Size and diversity of the sample
Reasoners consider hypotheses

- small birds
- large birds
- aquatic birds
- all birds
- etc.
The sample rules out some and not others…

small birds

all birds
small birds

all birds

Inductive generalisation is based on hypotheses consistent with the sample
Traditional view of reasoning

Sample data → Belief about the world

Properties of the sample shape learning
Reasoning as intuitive statistics

\[ P(h|d) = \frac{P(d|h)P(h)}{\sum_{h' \in \mathcal{H}} P(d|h')P(h')} \]

Sample data \rightarrow \text{Belief about the world} \rightarrow \text{State of the world} \rightarrow \text{Sample data}
Critical prediction: Learning depends on sampling

\[ P(h|d) = \frac{P(d|h)P(h)}{\sum_{h' \in H} P(d|h')P(h')} \]

State of the world

Sample data

The evidentiary value of the sample depends on how the learner thinks it was generated, or how it came to their attention.

Sample shape

Properties of the sample shape learning

Sample data

Belief about the world

Sample data

"learning"

Sample data

"sampling"
Epistemic vigilance: Statistical reasoning about untrustworthy data
These birds have plaxium blood

Does this bird have plaxium blood?
This is silly, but “it’s all made up” is absolutely a legitimate sampling assumption.

What if I told you the sample is a lie?

Does this bird have plaxium blood?
The price of inductive freedom is epistemic vigilance

informant knowledge

informant beliefs

evidence

Three year olds are easily deceived...

Mascaro & Sperber (2009)

The price of inductive freedom is epistemic vigilance

informant knowledge

informant beliefs

informant trustworthiness

evidence

... but four year olds are savvy statisticians


Mascaro & Sperber (2009)
Why epistemic vigilance?

People will try to “mislead with a half truth” if the listener is **naive**…

Ransom, Voorspoels, Perfors & Navarro (2017)
They rarely try this when the listener is suspicious!
Everyday reasoning about the world is intertwined with social reasoning about other people.

Why are you telling me this?
Where did you hear this?
Do you even know what you’re talking about?
What do you want me to do with this information?
What does all this buy us?
Taking a hint from a helpful teacher

Ransom, Perfors & Navarro (2016). *Cognitive Science*
Inductive reasoning when a **helpful teacher** provides the data
Ah, I get it - you’re calling my attention to sparrows

Inductive reasoning when a helpful teacher provides the data
Inductive reasoning when an indifferent world provides the data
Inductive reasoning when an *indifferent world* provides the data.

bloody trap is too small to fit anything except sparrows
Random:

Helpful:

Sampling mechanism:

“select items at random”

“select items to efficiently communicate an idea”
Adding positive instances has minimal effect if they’re too similar to things I already know about.

Adding positive instances from the same category conveys intent, and drives attention to that category.
Previous experience? (filler trials)

Cover story?

- **Helpful** cover story, filler trials imply helpful
- **Neutral** cover story, filler trials imply helpful
- **Neutral** cover story, filler trials imply random
- **Random** cover story, filler trials imply random
Change in argument strength

Target 1  Target 2  Control

Both Relevant  Relevant Fillers  Random Fillers  Both Random
Humans

<table>
<thead>
<tr>
<th>Condition</th>
<th>Change in argument strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both Relevant</td>
<td>-0.2</td>
</tr>
<tr>
<td>Relevant</td>
<td>-0.1</td>
</tr>
<tr>
<td>Fillers</td>
<td>0.0</td>
</tr>
<tr>
<td>Random</td>
<td>0.1</td>
</tr>
<tr>
<td>Both Random</td>
<td>0.2</td>
</tr>
<tr>
<td>Random Fillers</td>
<td></td>
</tr>
</tbody>
</table>

Diagram showing bar charts for different conditions.
Knowledge about animal categories *(theory of the world)* creates structural differences between the different arguments.

The sampling model *(theory of the context)* describes how “adding more data” can have different effects across conditions and arguments.
Using negative evidence to take hints from helpful teachers

Voorspoels, Navarro, Perfors, Ransom & Storms (2015). *Cognitive Psychology*
Mozart produces alpha waves in the brain

Positive evidence

This seems helpful!
Negative evidence

This... not so much

The sound of a falling rock does not

see Hempel (1945), Good (1960), etc
Okay, we start by telling people that Mozart does produce alpha waves…
… and they reason sensibly

Bach
Nirvana
waterfall

+Mozart
Adding Metallica as a negative example has a modest, sensible effect on inferences about Nirvana.
Bach  Nirvana  waterfall

judged likelihood

Bach

Nirvana

waterfall

Um.

+Mozart

-Falling rock
Negative evidence is interpreted as marking the category boundary.
Bayesian reasoners with a random sampling assumption do not produce the effect
Bayesian reasoners with a helpful sampling assumption do produce the effect
What does it mean to be “helpful” anyway?

\[ P(x|h) \propto P(h|x)^\alpha \]

The data \( x \) sampled by the communicator… … is designed to maximise the learner’s degree of belief in hypothesis \( h \)

Mozart but **not** rocks. Wink wink

Gotcha!
Prediction:

If the negative evidence is perceived as a **helpful hint** we should continue to get the effect.

If it is construed as an **arbitrary fact**, the effect should vanish.
Here’s the experimental results:

**Hint**

**Arbitrary**
Superficially useless information can have a huge effect when it is deemed to be helpful.

WTF is this “falling rocks” thing? It must be relevant somehow, so…
Extension: Negative evidence, fear conditioning & inductive reasoning

(work in progress!)
Fear conditioning

CS → US
Similarity based generalisation
Negative evidence along the same dimension ("near" CS-)

Lee, Lovibond, Hayes & Navarro (in prep)
CS- decreases generalisation on this side
CS- increases generalisation on this side
What happens when the “far” CS- has no value on the blue-green dimension?
CS- increases generalisation across the whole dimension
These are essentially the same design

Near negative

Distant negative
We needed a fancy sampling assumption for this. What about this?
Bayesian reasoning with random sampling produces the wrong pattern

(aside: compare to animal results, Switalski et al 1966)
Bayesian reasoning that assumes an intentional* sampling process works*
Taking the *wrong* hint because your teacher is a jerk

(another work in progress!)
Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Which is more probable?
(a) Linda is a bank teller
(b) Linda is a feminist bank teller

Kahneman & Tversky (1983)
The social/pragmatic account

Linda is blah blah blah…

... she’s feminist, obviously, why else would you tell me all that stuff

see Hertwig & Gigerenzer (1999)
The social/pragmatic account

What is Linda?

... feminist?

Because why else would you tell me all that stuff???
Social / pragmatic context
(a) Emily F. has diabetes
(b) Andrew J. is anaemic + Charlotte L. is hypertensive

(a) Sophie P. is short sighted + Jack N. has anxiety
(b) Ethan K. is overweight + Jack N. has anxiety

(a) Chloe M. has diabetes
(b) Chloe M. has diabetes + Chloe M. is overweight

Random / disconnected fact condition
Social / pragmatic

Random

Navarro, Tingey, Perfors & Keshwa (in prep)
Generate a description that implies but does not openly state that “Linda is a feminist”:

Linda is 31 and has had a rough upbringing, growing up with an abusive father which restricted her mother and her freedom. This upbringing was what made her decide to major in sociology and psychology within university. She has strong views on politics and other similar matters that affect men and women. She regularly attends rallies and protests on the weekend.
The “taboo” task

Generate a description that implies but does not openly state that “Paula is a bank teller”:

*Paula is 30, and loves buying clothes even at her age of 30. She is in contact with money so much that she has been able to calculate the exact change given before the cashier has given it to her. Her skills in counting are ingrained within her brain that she cannot turn it off, due to years dealing with cash.*
The “taboo” task

Generate a description that implies but does not openly state that “Brenda is a feminist & bank teller”:

*Brenda is 32 years old, methodical, logical, and passionate about her beliefs. She is very good with both people and numbers and is often able to spot errors. She is trusted by her friends to handle the money when planning an overseas trip. She is also a very individual woman and looks up to celebrities such as Emma Watson*
Several different versions

"Feminist / Bank Teller",
"Engineer / Jazz Musician",
"Introvert / Chef",
"Journalist / Anxious Person",
"Painter / Accountant",
"Extrovert / Statistician",
"Pacifist / Boxer",
"Butcher / Empath",
"Writer / Mechanic"

Navarro, Tingey, Perfors & Keshwa (in prep)
“Mind reading” task:

Isabelle is 41 years old and is very bright and good with numbers. Her creative flair has always been a passion although until recently she didn’t act on it. As a woman with two professions she works extremely hard and ensures that her conflicting logical and free spirited natures are harmonious in all aspects of life.

Which of the following do you believe the writer was trying to communicate when they wrote this description:

- Isabelle is a painter
- Isabelle is an accountant
- Isabelle is both a painter and an accountant
- None of the above
Navarro, Tingey, Perfors & Keshwa (in prep)
Standard conjunction task:

Ryan is 26. He spends his spare time unwinding and sitting on a couch at the end of the day reading or watching a movie. He has a small but tight knit group of friends. He likes talking to them individually and dislikes group outings.

How likely is it that this person belongs to each of the following categories? Please give an estimate of the probability from 0 to 100% for each category (0 being impossible and 100 being certain).

Teacher:   %
Introvert:  %
Activist:  %
Chef and Introvert: %
Neurosurgeon and pessimist: %
In progress: social vs random vignettes
In progress: social vs random vignettes

- Damien is [Fact X]
- He is also [Fact Q]
- ...

Fact A
Fact B
Fact C
...
Fact Z
...

Dice roll:

- Roll A
- Roll B
- Roll C
More tensions between social and random sampling: variations on the Monty Hall Dilemma

(yet another work in progress!)
The Monty Hall dilemma
A suitably constrained host:

<table>
<thead>
<tr>
<th></th>
<th>If A is correct</th>
<th>If B is correct</th>
<th>If C is correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host opens A</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host opens B</td>
<td></td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Host opens C</td>
<td></td>
<td></td>
<td>0%</td>
</tr>
</tbody>
</table>

Host won’t open the prize door
A suitably constrained host:

<table>
<thead>
<tr>
<th></th>
<th>If A is correct</th>
<th>If B is correct</th>
<th>If C is correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host opens A</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Host opens B</td>
<td></td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Host opens C</td>
<td></td>
<td></td>
<td>0%</td>
</tr>
</tbody>
</table>

Nor will they open the door you chose (A)
A suitably constrained host:

<table>
<thead>
<tr>
<th></th>
<th>If A is correct</th>
<th>If B is correct</th>
<th>If C is correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host opens A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host opens B</td>
<td>50%</td>
<td></td>
<td>100%</td>
</tr>
<tr>
<td>Host opens C</td>
<td>50%</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

Otherwise random
A suitably constrained host: a Bayesian reason to switch

* this is the correct solution to the original problem as stated by vos Savant
An indifferent host chooses randomly

<table>
<thead>
<tr>
<th></th>
<th>If A is correct</th>
<th>If B is correct</th>
<th>If C is correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host opens A</td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>Host opens B</td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>Host opens C</td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
</tr>
</tbody>
</table>
An indifferent host: a Bayesian reason for indifference

<table>
<thead>
<tr>
<th></th>
<th>If A is correct</th>
<th>If B is correct</th>
<th>If C is correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host opens A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host opens B</td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>Host opens C</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A malicious host who never offers a bet when your choice was wrong!

<table>
<thead>
<tr>
<th></th>
<th>If A is correct</th>
<th>If B is correct</th>
<th>If C is correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host opens A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host opens B</td>
<td></td>
<td>p%</td>
<td></td>
</tr>
<tr>
<td>Host opens C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host does not open a door</td>
<td></td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>
A malicious host with discretion: a Bayesian reason to stay

<table>
<thead>
<tr>
<th></th>
<th>If A is correct</th>
<th>If B is correct</th>
<th>If C is correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host opens A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host opens B</td>
<td>p%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Host opens C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host does not open a door</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A helpful host with discretion:

<table>
<thead>
<tr>
<th></th>
<th>If A is correct</th>
<th>If B is correct</th>
<th>If C is correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host opens A</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Host opens B</td>
<td></td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Host opens C</td>
<td></td>
<td></td>
<td>0%</td>
</tr>
<tr>
<td>Host does not open a door</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A helpful host with discretion:

<table>
<thead>
<tr>
<th></th>
<th>If A is correct</th>
<th>If B is correct</th>
<th>If C is correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host opens A</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Host opens B</td>
<td></td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Host opens C</td>
<td></td>
<td></td>
<td>0%</td>
</tr>
<tr>
<td>Host does not open a door</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>
A helpful host with discretion: 
A Bayesian reason to switch

<table>
<thead>
<tr>
<th></th>
<th>If A is correct</th>
<th>If B is correct</th>
<th>If C is correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host opens A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host opens B</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Host opens C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host does not open a door</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Probability of switching

Choice Proportion

Switch

Perfors, Navarro, Benders & Donkin (in limbo)
People (incorrectly?) view the original MHD as most similar to the malicious version
Possibly people are treating MHD as a “social reasoning” problem, and thinking that the host is malicious?
Can people be sensitive to conditional sampling without requiring a social component?

(also in progress, but almost finished)
Most of these effects rely on sampling by people.

This problem can be solved using social cognition.

Maybe this is all social reasoning?
Sampling across spatial locations

Eurasian magpie

Not social cognition

Australian magpie
Sampling across time

Not social cognition
You are currently classifying predators according to whether they pose a threat to humans. *Your team*, working at *this location recently* collected 200 observations and found that 50 (25%) of them met this criterion. This week, you have made another 4 observations, of which 3 (75%) met the above criterion. What proportion of predators in the area do you estimate pose a threat to humans?
Let’s make this a little more sneaky…

20 small birds with *plaxium* blood (SP+)
**Category sampling**: select items based on category membership (i.e. small birds)
Property sampling: select items based on possession of the property (i.e. plaxium blood)
Lawson & Kalish (2009)

**Property**

**Category**

**Random**

- **Mean Projection Score**
- **Target**
  - Robin
  - Pigeon
  - Owl
  - Ostrich
  - Mouse
  - Lizard
Hypotheses a reasoner might consider

Hayes, Banner & Navarro (2017)
Hypotheses consistent with the data
Category sampling

Frame explains absence of LP+ and LP-

Hypothesis must account for absence of SP-

Hayes, Banner & Navarro (2017)
Category sampling

2 of 3 hypotheses allow LP+
… so generalisation to large birds is very plausible

Hayes, Banner & Navarro (2017)
**Property sampling**

*Frame* explains absence of SP- and LP-.

*Hypothesis* must account for absence of LP+.

Hayes, Banner & Navarro (2017)
No remaining hypotheses allow LP+... so generalisation to large birds is very implausible.

Hayes, Banner & Navarro (2017)
Replication of L&K 2009

Hayes, Banner & Navarro (2017)
Explicit negative evidence (actual LP-) attenuates value of implicit negative evidence (no LP+).

Hayes, Banner & Navarro (2017)
A toy model

Property Generalization Score

Hayes, Banner & Navarro (2017)
If we tell people large birds are common, then the absence of LP+ remains suspicious in the property sampling condition, and the effect replicates…

Hayes, Banner & Navarro (2017)
But if we tell people large birds are rare, then the absence of LP+ and LP- is attributed to the base rate and the effect vanishes.
People pay attention to *mechanistic* constraints on sampling processes (not just social cues), and this shapes our reasoning in a sensible way.
More extensions?
Choice: What drives people’s active sampling?

instrumental learning task

transfer task

curiosity-driven?

reward-focused?
Law: Evidence sampling and expertise in the courtroom

Martire, Growns & Navarro (under review)
Society: Trust-based sampling via self-organising social networks (fake news...)

Cultural Evolution With Self-Selected Sources

Pairwise Trust

with Amy Perfors
Development: Exploratory versus goal-directed sampling by preschoolers

Chapter 1. Probability Models

Situations that are mutually independent and identically distributed (i.i.d.), or $X$ might be some general quantity. The set of possible values for $X$ is the sample space and is often denoted as $\mathcal{X}$. The members $P_\theta$ of the parametric family will be distributions over this space $\mathcal{X}$. If $X$ is continuous or discrete, then densities or probability mass functions exist. We will denote the density or mass function for $P_\theta$ by $f_{X,\theta}(x)$. For example, if $X$ is a single random variable with continuous distribution, then

$$P_\theta(x < X \leq b) = \int_a^b f_{X,\theta}(x)dx.$$

If $X = (X_1, \ldots, X_n)$, where the function $f_{X,\theta}(x)$ when $\theta = \theta_0$,

$$f_{X,\theta}(x; \theta)$$

where $x = (x_1, \ldots, x_n)$. After all of the function in (1.1) as a function defined by $L(\theta)$. Section above structure based on the corresponding representation theorem 1.4B. Exercise 1.2, and Definition 1.4B.

1.1.2 Classical Statistics

Classical inferential techniques in matrices, maximum likelihood estimators. Those will be covered in a few important aspects. How to solve a few of these here. Suppose the parameter lies in one pair, then set up a hypothesis $H_0 : \Theta = \Theta_0$. The simplest sort of a subset $\Theta \subseteq \mathcal{X}$, and then reject a null hypothesis and reject the rejection region for $H_0$. Tests are composed power function of a test with rejection at is $P_{\Theta_0 \cap \Theta_1}(H_0)$. Omit.

Example 1.2. Suppose that $X$ is a distribution under $P_\theta$. The usual

\[ P_\theta(x < X \leq b) = \int_a^b f_{X,\theta}(x)dx. \]

Target

A: 70% red

Lure

B: 50% red

Lure

C: 30% red

Lure urns are low novelty

Lure urns are high novelty

1 Using the theory of measures (see Appendix A) we will be able to dispense with the distinction between densities and probability mass functions. They will both be special cases of a more general type of “density.”
Wrap-up:

On the origins of data and the rationality* of human reasoning
People are smart. Limited, but smart.

“Common sense” reasoning is infuriatingly cunning, and requires people to learn from complex data sources (e.g., other people)
We need to disentangle facts from agendas

with Amy Perfors and Pat Shafto
We need to detect trickery
We need to know when to reject the rules we’re given.

with Charles Kemp
We need to read the intention of potentially malicious agents too many collaborators to list.
Common sense reasoning requires uncommonly rich statistical models

Who? Why? Where?
How? When? Really?

\[
P(h|d) = \frac{P(d|h)P(h)}{\sum_{h' \in \mathcal{H}} P(d|h')P(h')}
\]
Thanks!