

Rational Self-Doubt

Sherri Roush
Department of Philosophy
Logic Group
U.C., Berkeley
roush@berkeley.edu

This work was funded by NSF grant SES -
0823418:

“Fallibility and Revision
in Science and Society”



Judgmental Self-Doubt

Second-guessing your opinion on the basis of evidence about your cognitive skills or circumstances. Example occasions:

- You are confident in your memory of certain events, and then find out you are on mind-altering drugs.
- You are confident that the murderer is number 3 in the lineup, and then find out that human beings are generally unreliable in eyewitness testimony.
- You are watching for a tiger and your visual field suddenly goes blank.
- You are a woman not fully confident of your argument, and then discover that women are generally underconfident.

More problematically ...

The evolutionist admits to the Creationist that our theories *might be wrong*. The latter concludes his view is just as good. Every view is just a hypothesis! 60% of Americans appear to agree.

You admit that most past scientific theories have been false, and are expected to withdraw confidence in yours. (Pessimistic induction over the history of science.)

You are contemplating marriage and read that the divorce rate is 60%.

Evidence about q vs. evidence about yourself wrt q

You are confident that the murderer is #3 in the line-up, and then remember that the murderer was missing his left pinky finger. #3 is not.

vs.

You are confident that the murderer is #3 in the line-up, and then find out that human beings are generally unreliable in eyewitness testimony.

Evidence about the murderer vs. evidence about yourself

Evidence about q vs. evidence about yourself

q ... There is no tiger around

You are confident of q , and then see an orange, furry rustling in the trees.

vs.

You are confident of q , and then your visual field goes entirely blank.

In second case, you gain no evidence about tigers, yet intuitively you should change your confidence in q .

If so, then this revision cannot be modeled at the first order.

Terms – 1st- and 2nd- order evidence

There are different orders of evidence. So:

1st-order evidence about q, It will rain, is:

The barometer reads low pressure.

2nd-order evidence is relevant to your beliefs and their properties:

The owner of the house says the barometer has been rusted stuck for five years.

Conditionalization: assimilating new evidence by probabilistically “updating” all of your other beliefs.

Terms – 1st – and 2nd- order belief

Anything that can be evidence or hypothesis for you is one of your beliefs:

Let q be any proposition not containing a belief predicate, e.g.:
The murderer used a knife.

You have a ***first-order belief*** in q when you believe q . You express this:
“The murderer used a knife.”

You have a ***second-order belief*** (belief about a belief) when you believe that you believe q . To express this belief you say:
“I believe that the murderer used a knife”

If I describe you, I use one belief predicate vs. two belief predicates
I say: “He *believes* q ” vs. “He *believes* that he *believes* q ”

What is the rational thing to do?

1) Assimilate 2nd-order evidence:

-- Surely you have to *do something*.

-- Otherwise you violate the Principle of Total Evidence.

2) Ignore 2nd-order evidence:

-- If you start second-guessing, how do you stop? Arbitrarily?

-- 2nd-order probabilities don't exist, are trivial, are incoherent.

-- Miller's Principle (a.k.a. Self-Respect)

-- 2nd order revision could be distorting

-- Is there any added value?

Degrees of belief as probabilities

$q, r, s, \dots, q', r', s', \dots, P_S(q) = x, P_S(r) = y, \dots$ propositions

$P_S(q) = x$... Subject S has degree of belief x in q

$P_S(q/r) = z$... S's degree of belief in q *given* r is z

$P_T(P_S(q) = x) = x'$... Subject T has degree of belief x' *that* subject S has degree of belief x in q .

$P_S(P_S(q) = x) = x''$... Subject S has degree of belief x'' *that* subject S has degree of belief x in q .

Second-Order Probabilities

A second-order probability is a degree of belief about a degree of belief.

Don't exist: A belief is an expression. People can't have opinions about them.

A degree of belief is a disposition to act that has a certain strength. (Frank Ramsey) It can be measured by betting behavior. Same for beliefs about *what your beliefs are*: we ask you to bet about how you would bet.

Are trivial: All of them will be 0 or 1.

Being rationally fully confident and accurate requires an infallibility we don't have.

Better: Bayesian rationality (probabilistic coherence) shouldn't have requirements on substantive knowledge. This area is no exception.

Skyrms, Brian. 1980. "Higher-Order Degrees of Belief." In D. H. Mellor (ed.), *Prospects for Pragmatism: Essays in Memory of F. P. Ramsey*.

Agenda – Generalizing Bayesian Rationality

Rule for 2nd-order revision

Defense

Applications

More defense

The Principle that Gets in the Way

$$P(q/P(q) = x) = x$$

Self-Respect (SR)

What you think your degree of belief in q is is
what your degree of belief should be.

Don't disapprove of your own degrees of belief.

(Traditionally called "Miller's Principle.")

Restricted Self-Respect (RSR)

$$P(q/P(q) = x) = x$$

provided there is no statement of probability which when combined with “ $P(q) = x$ ” is relevant to q .

This says: The mere fact that you have a degree of belief is not a reason to have a different degree of belief.

Restricted Self-Respect (RSR)

$$P(q/P(q) = x) = x$$

*provided there is no statement of probability
which when combined with “ $P(q) = x$ ” is
relevant to q .*

This says: The mere fact that you have a degree of belief
is not a reason to have a different degree of belief.



Unrestricted Self-Respect (USR)

$$P(q/P(q) = x \cdot r) = x$$

*for any statement of probability r
for which P has a value*

I.e., no matter what else the subject believes.



Unrestricted Self-Respect (USR)

$$P(q/P(q) = x \cdot r) = x$$

for any statement of probability r

for which P has a value

I.e., no matter what else the subject believes.

We need to generalize away from this, preserving RSR.

Expressing Reliability

S is reliable to degree y when believing q to degree x iff

$$**PR**(q/P(q) = x) = y$$

(same as *fallible to degree 1-y*)

“**PR**” means objective probability, whatever kind you like.


Expressing the Quandary

Explicitly:

$$P(q/P(q) = x \cdot \mathbf{PR}(q/P(q) = x) = y) = ?$$

Expressing the Quandary

Explicitly:

$$P(q/ \quad) = ?$$


Incoming evidence: $P(q) = x \cdot PR(q/P(q) = x) = y$

Expressing the Quandary

Explicitly:

$$P(q/P(q) = x \cdot \mathbf{PR}(q/P(q) = x) = y) = ?$$

USR says r doesn't matter:

$$\mathbf{P(q/P(q) = x \cdot PR(q/P(q) = x) = y) = x}$$

Expressing the Quandary

Literally:

$$P(q/P(q) = x . PR(q/P(q) = x) = y) = ?$$

$$\text{USR: } P(q/P(q) = x . PR(q/P(q) = x) = y) = x$$

But consider:

$$P(q/P(q) = x . PR(q/P(q) = x) = y) = ?$$

$$=* P(q/PR(q) = y) = ?$$



$$P(\mathbf{PR}(P(q) = x) = 1/P(q) = x) = 1,$$

which is an instance of

$$P(\mathbf{PR}(A) = 1/A) = 1.$$

I.e., you are certain given A that the objective probability of A is 1.

Principal Principle

$$P(q/PR(q) = y \cdot s) = y$$

I.e., Your degrees of belief should conform to what you think the objective probability is.

(provided s is admissible, etc., but no need for s at all here.)

Forced Choice

$$P(q/P(q) = x \cdot PR(q/P(q) = x) = y) = ?$$

USR: $P(q/P(q) = x \cdot PR(q/P(q) = x) = y) = x$

Principal Principle:

$$P(q/P(q) = x \cdot PR(q/P(q) = x) = y) =^* y$$

What's good for the goose is good for the gander.

Our respect for the judgment of others is not unconditional:

$$P_T(q/P_S(q) = x \cdot \mathbf{PR}(q/P_S(q) = x) = y) = y$$

$$P_T(q/P_T(q) = x \cdot \mathbf{PR}(q/P_T(q) = x) = \mathbf{y}) = \mathbf{y}$$

Calibration and Re-calibration

Cal (synchronic constraint)

$$P(q/P(q) = x \cdot PR(q/P(q) = x) = y) = y$$

Re-Cal (diachronic constraint)

$$P_f(q) = P_i(q/P_i(q) = x \cdot PR(q/P_i(q) = x) = y) = y$$

To be *calibrated* is for one's confidence to match one's believed reliability. ($x = y$)

To *re-calibrate* is to update one's confidence in light of information about one's reliability. ($x \rightarrow y$)

Calibration and Re-calibration

Cal (synchronic constraint)

$$P(q/P(q) = x \cdot PR(q/P(q) = x) = y) = y$$

Re-Cal (diachronic constraint)

$$P_f(q) = P_i(q/P_i(q) = x \cdot PR(q/P_i(q) = x) = y) = y$$

Defended in:

Roush, "Second-Guessing: A Self-Help Manual," *Episteme* 6.3
(2009): 251-268.

Roush, "The Re-Calibrating Bayesian," manuscript

Applications

Eyewitness case

Tiger case

Creationist case

Pessimistic Induction

Roush, “Optimism about the Pessimistic Induction,”
New Waves in Philosophy of Science,
Palgrave-MacMillan, 2010.

Taking Stock

Supposed reasons to ignore 2nd-order evidence:

- If you start second-guessing, how do you stop?
Arbitrarily?
- 2nd-order probabilities ~~don't exist~~, are ~~trivial~~, are **incoherent**.
- ~~Miller's Principle (a.k.a. Self-Respect)~~
- 2nd order revision could be distorting
- Is there any added value?

Incoherence charges

- 1) Against Cal and Re-Cal
- 2) Against applying the function P to P -statements.

Incoherence of Cal?

1. $P(q) = x$,
2. $P(P(q) = x) = 1$, and
3. $P(PR(q/P(q) = x) = y) = 1$
4. $x \neq y$

Cal says:

$$P[q/(P(q) = x \cdot PR(q/P(q) = x) = y)] = y$$

From 2., 3., Cal, and *, we get $P(q) = y$.

But by assumption 1., $P(q) = x$. Contradiction.

Incoherence of Cal? Nah.

1. $P(q) = x$,
2. $P(P(q) = x) = 1$, and
3. $P(PR(q/P(q) = x) = y) = 1$
4. $x \neq y$

Cal says:

$$P[q/(P(q) = x \cdot PR(q/P(q) = x) = y)] = y$$

From 2., 3., Cal, and *, we get $P(q) = y$.

But by assumption 1., $P(q) = x$. Contradiction.

Incoherence of Second-order Probability?

Power Set problem:

Assume that the domain of the probability function P is all propositions of form q, r , etc., and $P(q) = x, P(r) = y$, and all Boolean combinations thereof.

Let S_1, S_2, S_3, \dots be the subsets of the set of propositions (in the domain of P). For each one, we can construct a distinct proposition (in the domain of P):

$$P(p) = z \text{ if and only if } p \text{ is a member of } S_1.$$

→ We've just mapped the set of subsets of the propositions into the set of propositions.

Solutions

Typed theory? Won't do the job

Re-Cal $P_f(q) = P_i(q/P_i(q) = x) \cdot PR(q/P_i(q) = x) = y = y$

→ The class of propositions is not a set.

But then probability must be definable on proper classes.

It is:

Rubin, Herman, A new approach to the foundations of probability.
1969 Foundations of Mathematics (Symposium Commemorating Kurt Gödel, Columbus, Ohio, 1966) pp. 46--50 Springer, New York

Taking Stock

Supposed reasons to ignore 2nd-order evidence:

-- If you start second-guessing, how do you stop?

Arbitrarily?

-- 2nd-order probabilities ~~don't exist~~, are ~~trivial~~, are ~~incoherent~~.

-- ~~Miller's Principle~~

-- 2nd order revision could be distorting

-- Is there any added value?

Taking Stock

Supposed reasons to ignore 2nd-order evidence:

-- If you start second-guessing, how do you stop?
Arbitrarily?

-- 2nd-order probabilities ~~don't exist~~, are ~~trivial~~, are ~~incoherent~~.

-- ~~Miller's Principle~~

-- 2nd order revision could be distorting

-- Is there any added value?

Regress and Pathology

Re-Cal

$$P_f(q) = P_i(q/P_i(q) = x \cdot PR(q/P_i(q) = x) = y) = y$$

This yields a new first-order degree of belief in q ,
so it looks like Re-Cal is applicable again.



Regress and Pathology - *No*

Re-Cal

$$P_f(q) = P_i(q/P_i(q) = x \cdot PR(q/P_i(q) = x) = y) = y$$

This yields a new first-order degree of belief in q , so Re-Cal is applicable again?

1. Yes, *provided* you have more evidence about reliability.
2. Yes, that's how conditionalization works.

Regress and Pathology - *No*

Re-Cal

$$P_f(q) = P_i(q/P_i(q) = x \cdot PR(q/P_i(q) = x) = y) = y$$

This yields a new first-order degree of belief in q , so Re-Cal is applicable again ...

1. Yes, provided you have more evidence.
2. Yes, that's how conditionalization works.

Be precise, be happy!

Is Recalibration distorting?

Seidenfeld, Teddy, 1985. "Calibration, Coherence, and Scoring Rules." *Philosophy of Science* 52(2): 274-94.

Problem: You can be calibrated in your confidence about rain by knowing that 20% of the days in the year it rains in your locale and announcing 20% chance of rain every day.

You have no discrimination. You could hedge your bets this way.

→ Calibration is an improper scoring rule.

Re-Cal is not a scoring rule, but a principle of conditionalization.

Distorting in the long run?

The (Principal Principle) direct inference embedded in Re-Cal goes like this, for day j :

$$P[q_j / P(q_j)=x_j \cdot \mathbf{Ch}(q_v/P(q_v)=x)=y(x) \cdot K] = y(x)$$

You will have the right degree of belief in q if you have the right chance (or objective probability) hypothesis (function), i.e., the correct $y(x)$.

So, the issue is whether in the long run we can converge on the correct chance hypothesis (function), e.g., the chance of rain on a given day (q_v) on which the subject believes to degree x that it will rain.

The outcomes that give us evidence for that are instances of degrees of belief in rain and whether it rained, on all the previous days.

Likelihood Ratio Convergence Theorem (LRCT)

Hawthorne, James, “Inductive Logic,” *Stanford Encyclopedia of Philosophy*.

Likelihood Ratio: $P(e/h)/P(e/-h)$

Suppose you will be fed *separating evidence*, that is h, -h predict at least some different outcomes in the stream of evidence you’re going to get. (All the work in application is here.)

Then if h is the true hypothesis it is probable that you will see outcomes that rule out all the –h hypotheses within a certain number of trials. (Law of Large Numbers)

Likelihood Ratio Convergence Theorem (LRCT)

Hawthorne, James, “Inductive Logic,” *Stanford Encyclopedia of Philosophy*.

Likelihood Ratio: $P(e/h)/P(e/-h)$

Suppose you will be fed *separating evidence*, that is h , $-h$ predict at least some different outcomes in the stream of evidence. Then it is probable that you will see outcomes that rule out all the $-h$ hypotheses within a certain number of trials.

An LRCT theorem can be proven for the chance hypothesis (function) Of Re-Cal, i.e., where $h = Ch(q_v, P(q_v)=x)=y(x)$

→ Adding 2nd-order conditioning to 1st-order conditioning is not distorting in the long run. Nor is 2nd order alone.

Pointless?

Seidenfeld: 1st-order conditionalization alone gets you to the truth about q (and to calibration) in the long run. Why bother with the re-calibration rule?

Point

Seidenfeld: Why bother with the second order re-calibration rule?

- If you think PP is a short-run rationality constraint, and you are a human being, then you have to follow Re-Cal.
- Re-Cal can change extreme degrees of belief.
- Conditionalization on 2nd-order *alone* will converge to truth about q .
- Assumptions the subject treats as unfalsifiable at the first order can be escaped by using Re-Cal.

Revising extreme probabilities

$$P_f(q) = P_i(q/P_i(q) = x \cdot PR(q/P_i(q) = x) = y) = y$$

If q is an empirical proposition, and x is 1, y may not be 1.

(Just because you are certain of q does not mean the objective probability of q is 1, or that your evidence tells you to believe that it is.)

Re-Cal Re-cap

News about reliability level could require:

- merely a moment of silence (if already calibrated)

- loss of confidence

- uptick in confidence

The effect depends on quantity, and is not runaway.

A rational subject retains unity and coherence not by perfect self-knowledge or unconditional self-respect but by coherently **1)** handling fallibility, and **2)** conforming to a single probability function at two orders.

Re-Cal **1)** is required by the Principal Principle, **2)** is coherent, **3)** is not distorting, **4)** converges in the long run, **5)** has added value over first-order conditionalization.

Re-Cal Regress

$$P_f(q) = P_i[q/(P_i(q) = n \cdot PR(q/P_i(q) = n) = z)] = \mathbf{z}$$

$$P_f(q) = P_i[q/\mathbf{P_i(q) = z \cdot PR(q/\mathbf{P_i(q) = z) = ?}] = ?$$

Sequential Forecasting with Re-calibration : $q = \text{It will rain}$

Day: 1 2 3 4 5 6 7 8 9 10

n_1 n_2 n_2 n_1 n_3 n_1 n_2 n_2 n_3 n_2 etc.
(initial degrees of belief in rain, i)

- - 0 1 - 1 .5 1/3 .5 2/3 etc.
(re-calibrated degrees of belief)

- - - - - 1 - - - - etc.
(re-re-calibrated degrees of belief)

→ **1** **0** **1** **1** **1** **1** **0** **1** **0** **0**

(1 = rain, 0 = rain)

If the moon is made of green cheese, then
the moon is made of cheese.

You may believe:

q is true if and only if I don't believe it.

I.e. you believe $PR(q/P(q) = 1) = 0$ and $PR(q/P(q) = 0) = 1$

Recalibration tells you to cycle between believing q and not believing q, until you get more information.

But if that's what you believe about yourself then that's obviously the rational thing to do!

Bummer of a belief to have, but not paradoxical. (Hope you get some new evidence, or third-order beliefs.)

Confidence, Accuracy do not imply USR

Confidence: For any claim q that the agent has some level of credence in, either $P(P(q)=n)=1$ or $P(P(q)=n)=0$.

Accuracy: If $P(P(q)=n)=1$, then $P(q)=n$; and if $P(P(q)=n)=0$, then $P(q)=n$.

$$(1) \quad P(q | P(q)=n)=x, \text{ for some } x$$

$$(2) \quad P(P(q)=n)=1$$

$$(3) \quad P(q | P(q)=n)=P(q) \quad \leftarrow \leftarrow \leftarrow$$

By (2) and Accuracy, we have:

$$(4) \quad P(q)=n$$

Putting (3) and (4) together, we get:

$$(5) \quad P(q | P(q)=n)=n$$

$$P_f(q) = P_i(q/P_i(q) = x \cdot PR(q/P_i(q) = x) = y) = y$$

Assume:

$$P(q) = x$$

$$P(P(q) = x) = 1$$

By definition of conditional probability:

$$\begin{aligned} P(q/P(q) = x) &= P(q \cdot P(q) = x) / P(P(q) = x) \\ &= P(q) \end{aligned}$$

$$\text{(because } P(P(q) = x) = 1)$$

$$\rightarrow P(q/P(q) = x) = x$$