

Is it rational to have faith?

Looking for new evidence, Good's Theorem,
and Risk Aversion.

Lara Buchak

UC Berkeley

buchak@berkeley.edu

*Special thanks to Branden Fitelson, who unfortunately
couldn't be here today.

What is Faith?

Paradigm examples of faith:

- “I have faith that God exists.”
- “I have faith that my spouse isn't cheating on me.”
- “I have faith that my car will start when I leave for work this morning.”

Characteristics of having faith in some proposition X:

- X is something whose truth one has a stake in.
 - We wouldn't say: “I have faith that in the 1860s there was a civil war in America”
- The person with faith in X has *some* evidence in favor of X.
 - We wouldn't say: “Even though I am entirely convinced by the argument from evil, I have faith that God exists.” Or “Even though my car hasn't worked for the last two weeks, I have faith that it will work this morning.” (At least, I claim that we *shouldn't* say these things.)
- The person with faith in X goes beyond the evidence in some sense.
 - We wouldn't say “I have faith that $2+2 = 4$.”

The first part of this talk will be devoted to spelling out the sense in which faith requires one to go beyond the evidence.

Faith and Evidence

- First thought: Faith in X requires being determined not to abandon one's belief in X under any circumstances.
 - Problem: Seems to require assigning $p(X) = 1$.
 - This clearly conflicts with Bayesian rationality.
 - And it seems to be too strong of a requirement anyway. Surely I could have faith in my spouse's constancy while recognizing that I would abandon this belief if I caught him cheating.
- Second thought: Faith in X requires treating the circumstances in which one would abandon one's belief in X as epistemically impossible.
 - Better, because it seems to explain how a person could lose faith: he could find himself in circumstances he previously took to be epistemically impossible.
 - Problem: Again, conflicts with Bayesian rationality: seems to require having degree of belief 0 to encountering evidence against X. E.g. requires that if one holds that $p(X | C) < p(X)$, then one must hold that $p(C) = 0$.
 - Again, seems to be too strong of a requirement. Again, I could have faith in my spouse's constancy while recognizing that catching him cheating is not incompatible with my current evidence.

Faith and Evidence II

- Better thought: Faith in X requires *not actively looking for evidence for or against X*. That is, not actively engaging in an inquiry whose only purpose is to figure out whether X is true or false.
 - Example: Hiring a private detective to investigate one's spouse.
 - Example: Being presented with an envelope which one knows contains evidence for or against one's spouse's constancy.
 - Example: Doubting Thomas.
- Doesn't entail assigning $p(X) = 1$ or having degree of belief 0 in encountering evidence against X.
 - Not actively looking for evidence for or against X is compatible with thinking that evidence against X might be out there. It is also compatible with accidentally coming across evidence against X, and then revising one's beliefs (losing faith).
- That a person could lose his faith in response to evidence seems entirely compatible with the claim that he really has faith; but that he actively seeks out evidence that he knows might make him revise his beliefs does not.
- This also explains why we think of faith as involving a commitment: it involves a commitment to not look for more evidence. Or we might say it involves a kind of trust that evidence against X won't be forthcoming, even though we recognize that encountering this evidence is epistemically possible.

Faith and Rationality

- This requirement of faith – that faith in X requires one to refrain from further inquiry into whether or not X is true – doesn't obviously conflict with Bayesian rationality, because it doesn't require a person to have any particular degrees of belief.
- But it turns out to conflict with rationality in a more subtle way.
- Decision theory claims that rational agents maximize expected utility.
- **Theorem** (I.J. Good): It always maximizes expected utility to make a new observation and use it, provided the cost of making the observation is negligible.
- So if one can only have faith in X by refraining from (purposely) making new cost-free observations, then having faith is always irrational.

Good's Theorem

Good's Theorem

If an agent is choosing between three acts, A_1 , A_2 , and A_3 , we can calculate the expected utility of each act:

$$\sum_i p(S_i) u(A_j | S_i)$$

The expected utility for a rational agent will be the maximum expected utility among the acts:

$$\max_j \sum_i p(S_i) u(A_j | S_i).$$

For each possible experimental result E_k , the expected utility of getting that result and using it will be:

$$\max_j \sum_i p(S_i | E_k) u(A_j | S_i).$$

Therefore, the expected utility of performing the experiment will be:

$$\sum_k p(E_k) \max_j \sum_i p(S_i | E_k) u(A_j | S_i).$$

Good proves that unless the act recommended is the same no matter what the result of the experiment is – that is, if the same act maximizes expected utility for each E_k – then this value is always higher than the agent's expected utility before making the observation.

Good's Theorem: An Example

- Consider a person with the following degrees of belief in G ("God exists") and R ("When I pray for something trivial in such-and-such circumstance I receive it").

$p(G) = 0.98$	$p(G R) = 0.99$	$p(G \sim R) = 0.5$	$p(R) = 48/49$
$p(\sim G) = 0.02$	$p(\sim G R) = 0.01$	$p(\sim G \sim R) = 0.5$	$p(\sim R) = 1/49$
- The agent is deciding between two acts, C and $\sim C$, associated with the following utility payoffs:

$u(C \& G) = 10$	$u(C \& \sim G) = 4$
$u(\sim C \& G) = 0$	$u(\sim C \& \sim G) = 7$
- Act C is something that's very good to do if God exists but not good to do otherwise, whereas $\sim C$ is nearly the same either way, though slightly better if God does not exist. We want to know whether the agent should make the decision now, or perform the experiment before making the decision, assuming that performing the experiment has no cost (including the cost of postponing the decision).
- If he makes the decision now, his expected utility values are:

$u(C) = (0.98)(10) + (0.02)(0) = 9.8$
$u(\sim C) = (0.98)(4) + (0.02)(7) = 4.06$
- So he will do C, with expected utility 9.8. If he carries out the experiment, then his expected utility values would be:

If R comes out true: $u(C) = 9.9$; $u(\sim C) = 4.03$. He will do C, with EU 9.9.
If $\sim R$ comes out true: $u(C) = 5.0$; $u(\sim C) = 5.5$. He will do $\sim C$, with EU 5.5
- Since he will get result R with probability 48/49 and $\sim R$ with probability 1/49, the EU of performing the experiment is:

$u(E) = (48/49)(9.9) + (1/49)(5.5) = 9.81$.

- So he should perform the experiment, as the theorem predicts.

A Challenge to Expected Utility Maximization

- Ignoring issues about faith, one serious challenge to expected utility theory is an objection to how it handles *risk aversion*.
- Most people are risk averse in that they prefer \$x to a gamble that yields \$x on average. E.g. most people prefer \$50 to a coin flip between \$0 and \$100.
- Expected utility theory handles this by allowing that the utility function diminishes marginally. This locates the badness of risk in the outcomes: risk aversion is a property of how an agent evaluates particular outcomes.
- Intuitively, though, risk is a property of a gamble as a whole: a gamble is risky insofar as it has a high variance, a low minimum, and so forth. Risk is a "global" property of a gamble, rather than a "local" property. And intuitively, risk aversion has to do with how an agent takes global features of gambles into account: for example, if an agent is averse to risk, he cares more about the minimum than an agent who is risk neutral.
- Furthermore, as the following examples show, allowing the utility function to diminish marginally cannot account for many typical attitudes towards risky gambles.

Risk Aversion: Examples

- **First Example: Allais Paradox**
- Consider the choice between A and B, and the choice between C and D, where the gambles are as follows:

A: \$5,000,000 with probability 0.1, \$0 otherwise.

B: \$1,000,000 with probability 0.11, \$0 otherwise.

C: \$1,000,000 with probability 0.89, \$5,000,000 with probability 0.1, \$0 with probability 0.01.

D: \$1,000,000 with probability 1.

- People tend to choose A over B, and D over C, but there are no utility values we can assign to \$0, \$1M, and \$5M such that these choices maximize expected utility.

Risk Aversion: Examples

- **Second Example: Gambles with Independent Goods**
- Consider two goods that are independent in the sense that having one does not increase or decrease the value of the other. E.g., non-independent goods include a right-hand glove and a left-hand glove, or a dollar and a second dollar. Independent goods might include a pair of gloves and a nice dinner.
- In utility terms, two goods are independent when $u(A \& \sim B) + u(B \& \sim A) = u(A \& B)$, assuming $u(\text{status quo}) = 0$.
- Consider the choice between the following deals. People tend to prefer deal 2 to deal 1, but there are no utility values we can assign to the goods such that they are independent and deal 2 is preferred to deal 1.

	HH	HT	TH	TT
Deal 1	Dinner	Dinner and gloves	Nothing	Gloves
Deal 2	Dinner	Dinner	Gloves	Gloves

Generalizing the Theory to Include a Subjective Risk Function

- Again, what these examples show is that people tend to care about global properties of gambles, e.g. the minimum (what happens in the worst state), the maximum, the variance, etc.

Standard theory:

{A, p; B, 1 - p} A gamble that yields A with probability p and B with probability $1 - p$
 Expected utility: $u(\{A, p; B, 1 - p\}) = p(u(A)) + (1 - p)u(B)$
 $= u(B) + p[u(A) - u(B)]$

- Decision makers have subjective utilities and probabilities.
- Decision makers are risk averse if the utility function is concave. (Or, more generally, if the goods involved are not independent.) This implies that risk aversion is a property of how an agent values individual outcomes.

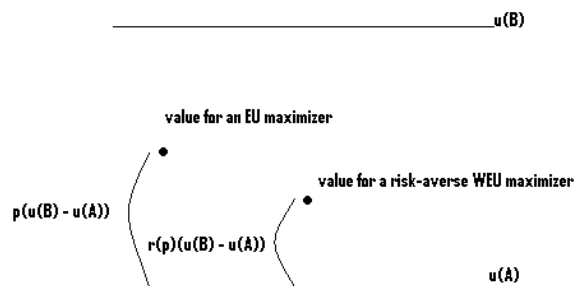
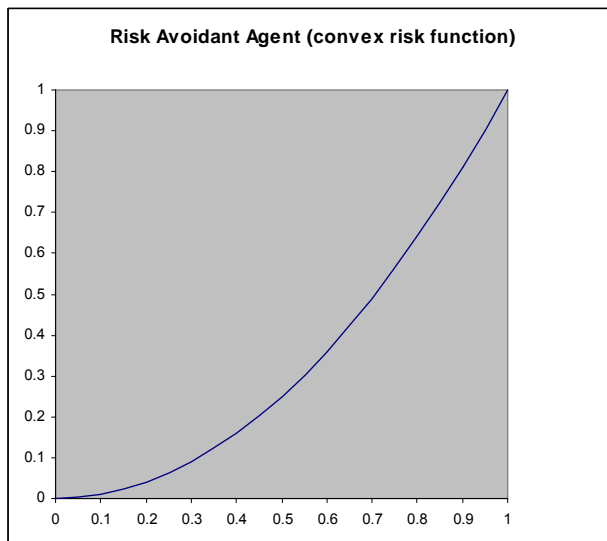
Generalization:

If A is at least as good as B:

Weighted-expected utility: $u(\{A, p; B, 1 - p\}) = u(B) + r(p)[u(A) - u(B)]$,

- $r: [0, 1] \rightarrow [0, 1]$ is a function that represents how the agent takes risk into account.
- Decision makers have subjective utilities, probabilities, and evaluation of risk.

Example of a Risk Function



Risk Function

- Decision makers are risk averse if the risk function is convex. This implies that risk aversion is a property of how an agent values *arrangements of outcomes across the possibility space*.
- Decision makers are risk seeking if the risk function is concave. Decision makers are risk neutral if $r(p) = p$ – they are ordinary expected utility maximizers.
- If $r(p)$ is convex, the minimum is weighted more heavily than for the risk-neutral agent. Limit case: $r(p) = 0$ if $p \neq 1$, the agent who uses maximin. If $r(p)$ is concave, the maximum is weighted more heavily than for the risk-neutral agent. Limit case: $r(p) = 1$ if $p \neq 0$, the agent who uses maximax.
- The introduction of a risk function allows for rational agents to have the preferences in the two original examples. It also allows for people to prefer \$50 to a coin flip between \$0 and \$100, while still valuing these amounts of money linearly. And it explains why the St. Petersburg gamble is worth so little.

Gambles over multiple options, e.g. $\{O_1, p_1; O_2, p_2; \dots; O_n, p_n\}$

Expected utility:

$$u(O_1) + \left(\sum_{i=2}^n p_i\right)(u(O_2) - u(O_1)) + \left(\sum_{i=3}^n p_i\right)(u(O_3) - u(O_2)) + \dots + (p_n)(u(O_n) - u(O_{n-1}))$$

Weighted-expected utility:

$$u(O_1) + r\left(\sum_{i=2}^n p_i\right)(u(O_2) - u(O_1)) + r\left(\sum_{i=3}^n p_i\right)(u(O_3) - u(O_2)) + \dots + r(p_n)(u(O_n) - u(O_{n-1}))$$

Example Redux

- As above:

$$\begin{array}{llll} p(G) = 0.98 & p(G | R) = 0.99 & p(G | \sim R) = 0.5 & p(R) = 48/49 \\ p(\sim G) = 0.02 & p(\sim G | R) = 0.01 & p(\sim G | \sim R) = 0.5 & p(\sim R) = 1/49 \\ u(C \& G) = 10 & u(C \& \sim G) = 0 & & \\ u(\sim C \& G) = 4 & u(\sim C \& \sim G) = 7 & & \end{array}$$

- Now let us assume our agent is risk averse, with $r(p) = p^2$. Before performing the experiment, his weighted-expected utility values are:

$$u(C) = 0 + r(0.98)(10 - 0) = 9.604$$

$$u(\sim C) = 4 + r(0.02)(7 - 4) = 4.0012.$$

- So he will do C, with WEU 9.604. If he carries out the experiment, then his WEU values are:

If R comes out true:

$$u(C) = 0 + r(0.99)(10 - 0) = 9.801$$

$$u(\sim C) = 4 + r(0.01)(7 - 4) = 4.0003$$

He will do C, with WEU 9.801

- If $\sim R$ comes out true:

$$u(C) = 0 + r(0.5)(10 - 0) = 2.5$$

$$u(\sim C) = 4 + r(0.5)(7 - 4) = 4.75$$

He will do $\sim C$, with WEU 4.75.

- Since he will get result R with probability 48/49 and $\sim R$ with probability 1/49, the WEU of performing the experiment is:

$$u(E) = 4.75 + r(48/49)(9.801 - 4.75) = 9.579.$$

- This is lower than the WEU of not performing the experiment. So he shouldn't perform the experiment!

Example Explained

- If someone believes in a proposition X with near certainty, then evidence in favor of that proposition won't drastically alter her degree of belief in that proposition, so it won't drastically alter her WEU.
- On the other hand, evidence for $\sim X$ will cause a drastic change.
- In the situation above, C is a risky option: it has a high variance, and does very well if X is true and very poorly if $\sim X$ is true. So the risk-averse agent needs a fairly high credence in X in order to perform it.
- Therefore, evidence for $\sim X$ will put the risk-averse agent in a situation in which it is rational to perform $\sim C$. But from her (new) point of view, there's still a fairly high chance (0.5) that this is misleading evidence: *evidence that leads her to do the action that, though rational, in the end gives her a lower payoff.*
- This is very different from a case in which an agent does not want to discover new evidence lest she find out something depressing. Instead, it is a case in which looking for more evidence comes with a significant chance that one will perform an action that is rational but, given the way the world turns out, wrong – it comes with a significant chance that one will make a mistake, so to speak.
- For example, if a private investigator turned up evidence that your spouse was cheating, it would indeed be rational to leave her, but, given your previously high credence in her faithfulness, *there is still a significant chance that by doing the rational thing you would be missing out on some great good*, namely a relationship with a spouse who is *in fact faithful*. And if God didn't respond to a believer's heartfelt prayers in a "testing" situation, it might indeed be rational for the believer to stop going to church, but even then, his credence in God's existence will still be fairly high, so there will be a significant chance that he is missing out on a great good.

Example Explained II

- If one's credence in X is antecedently very high, then carrying out an experiment that turns out to confirm X is not particularly helpful: it is very unlikely to change one's course of action, and it increases the weighted-expected utility of this action only slightly.
- On the other hand, carrying out an experiment that turns out to disconfirm X might lead one to change one's action (rationally, given the experimental result). But it might lead one to change one's action in a way that, a lot of the time, yields a worse result.
- In other words, carrying out an experiment is a risky endeavor. And if one's credence in X is antecedently very high, and one is risk-averse, the potential bad that comes from the possible negative result is not outweighed by the very minimal good that comes from the possible (and likely) positive result.
- Furthermore, for a risk-averse agent, the cases in which carrying out an experiment is going to be irrational are cases that have the characteristics mentioned at the beginning as typical of faith: the agent has an antecedently high degree of belief in X , and the agent has a stake in X in the sense that he is performing an act that is much worse if X turns out to be false.

Further Thoughts

- Conclusions if you think non-EU maximizers are irrational?
 - Even if being risk averse in my sense is irrational, experimental results suggest that many people are risk averse. So we might conclude from this talk that if you know that you are (irrationally) risk-averse, you shouldn't look for more evidence, as a way to monitor your future (irrational) behavior.
 - There is a surprising connection between whether gathering more evidence is helpful and an agent's attitude towards risk – or, more generally, towards global properties of gambles.
- Connection with the claim that faith requires “behaving as if X is true”?
- Exhaustive characterization of the conditions under which looking for more evidence is irrational?