Utility-weighted sampling in decisions from experience

Falk Lieder, Tom Griffiths, Ming Hsu
UC Berkeley
Extreme potential outcomes influence people as if they were far more likely than they really are.
Extreme potential outcomes influence people as if they were far more likely than they really are.
Extreme potential outcomes influence people as if they were far more likely than they really are.
Extreme potential outcomes influence people as if they were far more likely than they really are.

38% of Americans say they are less likely to travel overseas because of 9/11.
Extreme potential outcomes influence people as if they were far more likely than they really are.

38% of Americans say they are less likely to travel overseas because of 9/11.
Expected Utility Theory
Expected Utility Theory

Take action $\arg\max_a E_{p(O|a)}[u(O)]$

utility of outcome $O$

expected value
Take action \( \arg \max_a E_{p(O|a)} \left[ u(O) \right] \)

\( \prod p(o|a) \cdot u(o) \)
Expected Utility Theory

Take action $\operatorname{argmax}_a E_{\rho(o|a)} \left[u(O)\right]$

$\prod \rho(o|a) \cdot u(o)$ do

Intractable!
Expected Utility Theory

Take action \( \text{argmax}_a \mathbb{E}_{p(o|a)} [u(O)] \)

\[ \prod p(o|a) \cdot u(o) \]  

Violated!

Intractable!
EU can be Approximated by Sampling

\[ EU = \sum p(o|a) \cdot u(o) \, do \]
EU can be Approximated by Sampling

\[ EU = \sum p(o|a) \cdot u(o) \]

\[ o_1, \ldots, o_s \sim p(o|a) \]

simulated outcomes
EU can be Approximated by Sampling

\[ EU = \sum_{o \mid a} p(o \mid a) \cdot u(o) \, \text{do} \]

\[ o_1, \cdots, o_s \sim p(o \mid a) \]

simulated outcomes \rightarrow EU estimates

\[ \hat{U}(a) = \frac{1}{s} \sum_{i=1}^{s} u(o_s) \]
EU can be Approximated by Sampling

$$EU = \prod p(o \mid a) \cdot u(o) \, do$$

$$o_1, \ldots, o_s \sim p(o \mid a)$$

$$\hat{a}^* = \arg\max_a \hat{U}(a)$$

simulated outcomes \quad \rightarrow \quad EU estimates \quad \rightarrow \quad decision$$

$$\hat{U}(a) = \frac{1}{s} \sum_{i=1}^{s} u(o_s)$$
EU can be Approximated by Sampling

$$\text{EU} = \prod_{o}(p(o|a) \cdot u(o))$$

$$o_1, \ldots, o_s \sim p(o|a)$$

simulated outcomes $\Rightarrow$ EU estimates $\Rightarrow$ decision

$$\hat{a}^* = \arg \max_a \hat{U}(a)$$

$$\hat{U}(a) = \frac{1}{s} \sum_{i=1}^{s} u(o_s)$$

finite time $\Rightarrow$ finitely many simulated outcomes
Representative sampling is dangerous

5 out of 6 people enjoy Russian Roulette
Representative sampling is dangerous.
Representative sampling is dangerous variance

5 out of 6 people enjoy Russian Roulette
Representative sampling is dangerous variance.
Representative sampling is dangerous

bias

5 out of 6 people enjoy Russian Roulette
Representative sampling is dangerous vs. bias

5 out of 6 people enjoy Russian Roulette
Utility estimation by importance sampling
Utility estimation by importance sampling
Utility estimation by importance sampling

\[ o_1, o_2, o_3 \sim q \]

simulated outcomes

\[ \text{fun} \quad \text{death} \]
Utility estimation by importance sampling

\[ \hat{U}_{q,s}(a) = \frac{1}{\sum_{j=1}^{s} w_j} \sum_{j=1}^{s} w_j \cdot u(o_j) \]
Utility estimation by importance sampling

\[ o_1, o_2, o_3 \sim q \]

simulated outcomes

\[ \hat{U}_{q,s}^{IS}(a) = \frac{1}{\sum_{j=1}^{s} w_j} \sum_{j=1}^{s} w_j \cdot u(o_j) \]

\[ w_i = \frac{p(o_i)}{q(o_i)} \]
Utility estimation by importance sampling

\( o_1, o_2, o_3 \sim q \)

simulated outcomes

EU estimates

\[
\hat{U}_{q,s}^{IS}(a) = \frac{1}{\sum_{j=1}^{s} w_j} \sum_{j=1}^{s} w_j \cdot u(o_j)
\]

\[ w_1 = \frac{p(o_1)}{q(o_1)} \]

...  

\[ w_3 = \frac{p(o_3)}{q(o_3)} \]
Utility estimation by importance sampling

$o_1, o_2, o_3 \sim \mathbf{q}$

simulated outcomes

$\hat{a}^* = \arg \max_a \hat{U}^{\text{IS}}_{q,s}(a)$

EU estimates → decision

$\hat{U}^{\text{IS}}_{q,s}(a) = \frac{1}{\sum_{j=1}^{s} w_j} \sum_{j=1}^{s} w_j \cdot u(o_j)$

$w_1 = \frac{p(o_1)}{q(o_1)}$

...$

w_3 = \frac{p(o_3)}{q(o_3)}$
Utility estimation by importance sampling

\[ \hat{a}^* = \arg\max_a \hat{U}_{q,s}(a) \]

Which distribution should the brain sample from?

\[ o_1, o_2, o_3 \sim q \]

simulated outcomes

\[ p(o_1)/q(o_1) \]

\[ w_1 \]

... 

\[ w_3 = p(o_3)/q(o_3) \]

EU estimates 

decision

\[ \hat{U}_{q,s}(a) = \frac{1}{\sum_{j=1}^{s} w_j} \sum_{j=1}^{s} w_j \cdot u(o_j) \]
Answer: Utility-Weighted Sampling (UWS)

$$\tilde{q}(o) \propto p(o) \cdot |u(o)|$$

- simulation frequency
- extremity

probability
Answer: Utility-Weighted Sampling (UWS)

\[ \tilde{q}(o) \propto p(o) \cdot |u(o)| \]

Lieder, Hsu, Griffiths (2014)
Decisions from Experience (Ludvig, et al., 2014)
Decisions from Experience (Ludvig, et al., 2014)
Decisions from Experience (Ludvig, et al., 2014)
Decisions from Experience (Ludvig, et al., 2008)
Decisions from Experience (Ludvig, et al., 2008)
Decisions from Experience (Ludvig, et al., 2008)
Inconsistent Risk Preferences Emerge from Learning

Ludvig et al. (2014), Experiments 1-2

![Graph showing risk choice percentage over block number for different scenarios.](image-url)
UWS Can Emerge from Reward-Modulated Associative Learning

Actions:
UWS Can Emerge from Reward-Modulated Associative Learning

Actions:

\[ a_t \]
UWS Can Emerge from Reward-Modulated Associative Learning

Outcomes:  -40  -20  +20  +40

Actions:  

\( o_t \)

\( a_t \)
UWS Can Emerge from Reward-Modulated Associative Learning

Outcomes: -40 -20 +20 +40

Actions:

\( w_t \)

\( a_t \)

\( o_t \)
UWS Can Emerge from Reward-Modulated Associative Learning

Outcomes:  -40  -20  +20  +40

Actions:

$$w_t$$

$$o_t$$

$$a_t$$

Prediction error:

$$\left| \text{PE}(o_t) \right| = |r(o_t) - \bar{r}|$$
UWS Can Emerge from Reward-Modulated Associative Learning

Outcomes: -40 -20 +20 +40

Actions: 

learning rate

\[
(w_t(a_t, o_t) + \alpha \cdot |PE(o_t)|)
\]

prediction error

\[
|PE(o_t)| = |r(o_t) - \bar{r}_t|
\]
UWS Can Emerge from Reward-Modulated Associative Learning

Outcomes:  
-40  -20  +20  +40

Actions:

\[ (1 - \gamma) \cdot (w_t(a_t, o_t) + \alpha \cdot |PE(o_t)|) \]

learning rate

prediction error

\[ |PE(o_t)| = |r(o_t) - \bar{r}_t| \]
UWS Can Emerge from Reward-Modulated Associative Learning

For all $o \neq o_t$:

$$(1 - \gamma) \cdot w_t(a_t, o)$$

Outcomes:

-40  -20  +20  +40

Prediction error:

$$|PE(o_t)| = |r(o_t) - \bar{r}_t|$$

Actions:

Learning rate:

$$(1 - \gamma) \cdot (w_t(a_t, o_t) + \alpha \cdot |PE(o_t)|)$$

Forgetting rate:

$$(1 - \gamma) \cdot w_t(a_t, o)$$
Learning Rule Convergences to Utility-Weighted Sampling

Utility-weighted learning converges to

\[ w_{a,o} \propto p(o\mid a) \cdot |u(o)| \text{ with } u(o) = \text{PE}(o) \]
Learning Rule Convergences to Utility-Weighted Sampling

Utility-weighted learning converges to

\[ w_{a,o} \propto p(o \mid a) \cdot |u(o)| \text{ with } u(o) = PE(o) \]

with activation function

\[ P(Y=1) \propto w^T \cdot x \]

the network learns to perform utility-weighted sampling.
Efficient coding (Summerfield & Tsetsos, 2015)

\[ |PE(o_t)| = |r(o_t) - \bar{r}| \]
Efficient coding (Summerfield & Tsetsos, 2015)

\[ |PE(o_t)| = |r(o_t) - \bar{r}| \]

\[ \bar{r}_t = \bar{r}_{t-1} + \eta \cdot (r_t - \bar{r}_{t-1}) \]
Efficient coding (Summerfield & Tsetsos, 2015)

\[ |\text{PE}(o_t)| = |r(o_t) - \bar{r}_t| \]

\[ r(o) = \frac{o}{o_t^{\text{max}} - o_t^{\text{min}}} + \mathcal{E} \]

\[ \mathcal{E} \sim \mathcal{N}(0, \sigma_{\mathcal{E}}^2) \]

\[ \bar{r}_t = \bar{r}_{t-1} + \eta \cdot (r_t - \bar{r}_{t-1}) \]
Model fitting

Maximum-Likelihood-Estimation of $s, \alpha, \gamma, \lambda$, and $\sigma^2_{\epsilon}$ from block-by-block choice frequencies in Experiments 1-4 by Ludvig et al. (2014).

A single set of parameters fits all experiments.
UWS captures that people learn to overweight extreme outcomes

Ludvig et al. (2014), Experiments 1-2

![Graph showing risky choice in percentage over block number for different outcomes and trials.](image_url)
UWS captures that people learn to overweight extreme outcomes

Ludvig et al. (2014), Experiments 1-2
Utility-Weighted Sampling Captures Memory Biases (Madan et al. 2014)

Which outcome comes to mind first?
Utility-Weighted Sampling Captures Memory Biases (Madan et al. 2014)

Which outcome comes to mind first?
Utility-Weighted Sampling Captures Memory Biases (Madan et al. 2014)

Which outcome comes to mind first?
Utility-Weighted Sampling Captures Frequency Estimation Bias (Madan et al. 2014)

How often did this door lead to each outcome?

+40: ___ %
+20: ___ %
0: ___ %
Utility-Weighted Sampling Captures Frequency Estimation Bias (Madan et al. 2014)

How often did this door lead to each outcome?

+40: ___ %
+20: ___ %
0: ___ %
Utility-Weighted Sampling Captures Frequency Estimation Bias (Madan et al. 2014)

How often did this door lead to each outcome?

+40: ___ %
+20: ___ %
0: ___ %
Biased Beliefs Predict Risk Seeking

$r_{people} = +0.16; p<0.05$
Biased Beliefs Predict Risk Seeking

$r_{\text{UWS}} = +0.23$

$r_{\text{people}} = +0.16; \ p<0.05$
Biased Beliefs Predict Risk Seeking

\[ r_{\text{UWS}} = +0.23 \]
\[ r_{\text{people}} = +0.16; \ p<0.05 \]
\[ r_{\text{people}} = -0.48; \ p<0.05 \]
Biased Beliefs Predict Risk Seeking

\[ r_{UWS} = +0.23 \]
\[ r_{people} = +0.16; \ p < 0.05 \]

\[ r_{UWS} = -0.44 \]
\[ r_{people} = -0.48; \ p < 0.05 \]
Conclusions
Conclusions

1. Utility-weighted sampling provides a unifying explanation for biases in memory, judgment, and decision making.
Conclusions

1. Utility-weighted sampling provides a unifying explanation for biases in memory, judgment, and decision making.

2. Utility-weighted sampling can emerge from reward-modulated associative learning.
Conclusions

1. Utility-weighted sampling provides a unifying explanation for biases in memory, judgment, and decision making.
2. Utility-weighted sampling can emerge from reward-modulated associative learning.
3. People overweight extreme events, because it is rational to focus on the most important eventualities.
Conclusions

1. Utility-weighted sampling provides a unifying explanation for biases in memory, judgment, and decision making.
2. Utility-weighted sampling can emerge from reward-modulated associative learning.
3. People overweight extreme events, because it is rational to focus on the most important eventualities.
4. Some cognitive biases may serve or reflect the rational allocation of finite cognitive resources.
Thank you!

This work was supported by grant number N00014-13-1-0341 from the Office of Naval Research to Thomas L. Griffiths, grant number RO1 MH098023 from the National Institutes of Health to Ming Hsu, and a Berkeley Graduate Fellowship to Falk Lieder.

We thank Elliot Ludvig, Quentin Huys, and Mike Pacer for their suggestions and feedback.