MAX-PLANCK-INSTITUT
FÜR BIOLOGISCHE KYBERNETIK

The Art and Science of Modeling Human Decision-Making

Marcel Binz

September 8, 2022

Max Planck Institute for Biological Cybernetics
Computational Principles of Intelligence Lab
https://marcelbinz.github.io/

Main goal: Teach you a general set of modeling skills that you can apply to your own research projects afterward.

Part 1: Modeling decision-making on multi-armed bandit problems.

Part 2: How models can be used to draw inferences about human decision-making.

Computational Modeling

Computational models are immensely useful for improving our understanding of the human mind.

> *"Verbally expressed statements are sometimes flawed by internal inconsistencies, logical contradictions, theoretical weaknesses and gaps. A running computational model, on the other hand, can be considered as a sufficiency proof of the internal coherence and completeness of the ideas it is based upon."*
>
> Fum, D., Del Missier, F., & Stocco, A. (2007). The cognitive modeling of human behavior: Why a model is (sometimes) better than 10,000 words. Cognitive Systems Research, 8(3), 135-142.

Computational Modeling

~~How to model?~~

Why to model?

> "[Models are] *tools for exploring the implications of ideas.*"
>
> McClelland, J. L. (2009). The place of modeling in cognitive science. Topics in Cognitive Science, 1(1), 11-38.

Computational Modeling

The Art of Modeling

Building models is sometimes more of an art than a science.

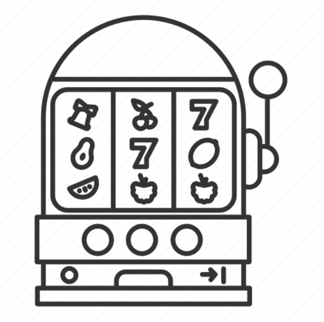
There is not a single right way to do it.

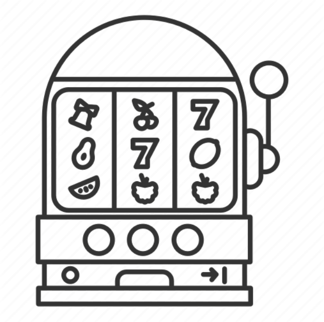
> "*How to be a great painter cannot be taught in words; one learns by trying many different approaches that seem to surround the subject.*"
>
> Hamming, R. R. (1997). Art of doing science and engineering: Learning to learn. CRC Press.

I think this is also true for modeling!

Part 1: Models of Decision-Making

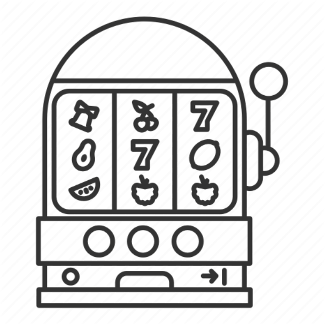
Multi-Armed Bandits



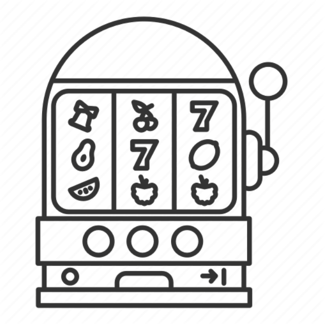


Multi-Armed Bandits

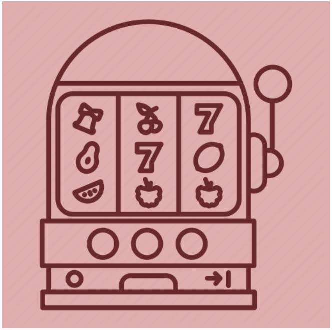
-41



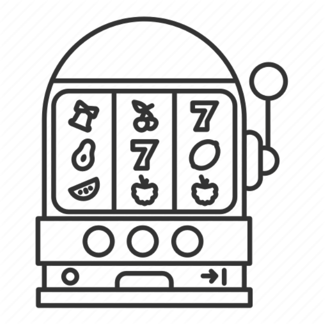
Multi-Armed Bandits



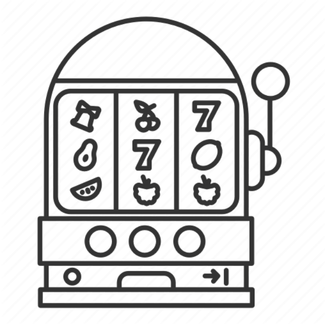
40

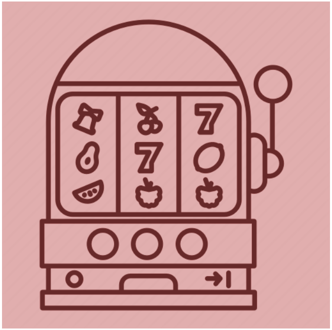


-48



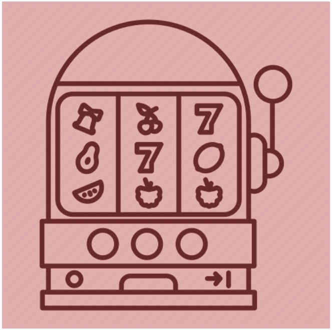
Multi-Armed Bandits



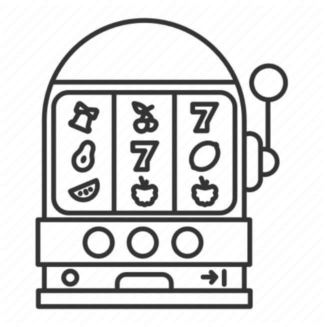


41

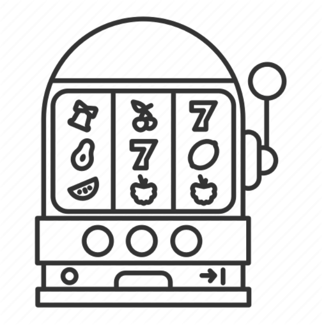
Multi-Armed Bandits



-33

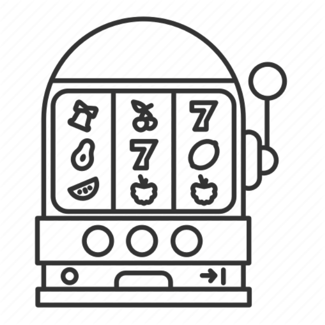


Multi-Armed Bandits



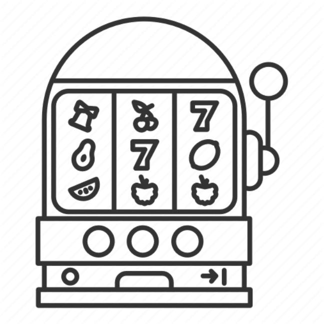
29

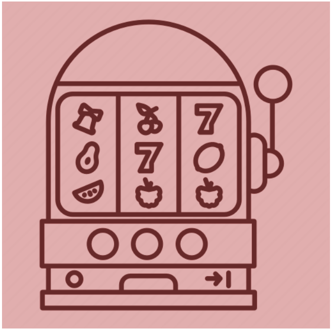
Multi-Armed Bandits



11

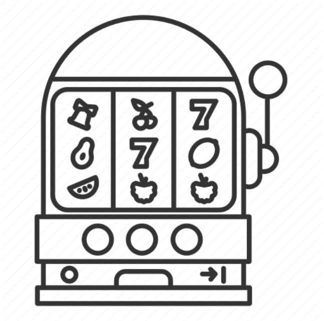
Multi-Armed Bandits



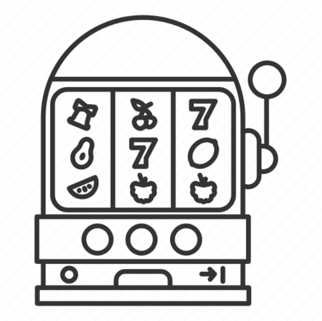


24

Multi-Armed Bandits



$p(R_t|A_t = 1, \mu)$



$p(R_t|A_t = 2, \mu)$

Multi-Armed Bandits

"Originally considered by Allied scientists in World War II, it proved so intractable that, according to Peter Whittle, the problem was proposed to be dropped over Germany so that German scientists could also waste their time on it."

Multi-armed bandit. (2022, August 2). In Wikipedia. ***https://en.wikipedia.org/wiki/Multi-armed_bandit***

Decision-making without a brain: how an amoeboid organism solves the two-armed bandit

Chris R. Reid[1,†], Hannelore MacDonald[1], Richard P. Mann[2], James A. R. Marshall[3,4], Tanya Latty[5] and Simon Garnier[1]

[1]Department of Biological Sciences, New Jersey Institute of Technology, Newark, NJ 07102, USA
[2]School of Mathematics, University of Leeds, Leeds LS2 9JT, UK
[3]Department of Computer Science, University of Sheffield, Sheffield S1 4DP, UK
[4]Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK
[5]School of Life and Environmental Sciences, University of Sydney, Sydney, New South Wales 2006, Australia

SG, 0000-0002-3886-3974

Several recent studies hint at shared patterns in decision-making between taxonomically distant organisms, yet few studies demonstrate and dissect mechanisms of decision-making in simpler organisms. We examine decision-making in the unicellular slime mould *Physarum polycephalum* using a classical decision problem adapted from human and animal decision-making studies: the two-armed bandit problem. This problem has previously only been used to study organisms with brains, yet here we demonstrate that a brainless unicellular organism compares the relative qualities of multiple options, integrates over repeated samplings to perform well in random environments, and combines information on reward frequency and magnitude in order to make correct and adaptive decisions. We extend our inquiry by using Bayesian model selection to determine the most likely algorithm used by the cell when making decisions. We deduce that this algorithm centres around a tendency to exploit environments in proportion to their reward experienced through past sampling. The algorithm is intermediate in computational complexity between simple, reactionary heuristics and calculation-intensive optimal performance algorithms, yet it has very good relative performance. Our study provides insight into ancestral mechanisms of decision-making and suggests that fundamental principles of decision-making, information processing and even cognition are shared among diverse biological systems.

Multi-Armed Bandits

If the decision-maker knows $\mu = [\mu_1, \mu_2]$ the task is trivial.

We thus consider the case where μ is unobserved.

The goal of the decision-maker is to infer it.

We will assume that the decision-maker learns in a rational way.

This is an assumption that we should in practice test against alternative theories.

Bayesian Inference

In an uncertain world, optimal solutions are specified by **Bayesian inference**.

The general idea:

1. The decision-maker maintains a belief over unobserved parameters $p(\mu)$.
2. This belief is updated via Bayes' rule once a new data-point (a_t, r_t) is observed:

$$p(\mu|R_t = r_t, A_t = a_t) = \frac{p(R_t = r_t|A_t = a_t, \mu)p(\mu)}{\int p(R_t = r_t|A_t = a_t, \mu)p(\mu)d\mu}$$

Bayesian Inference

Bayesian is optimal if prior and likelihood match how data is generated in the world.

Many justifications for this (Dutch book arguments, free energy minimization, performance-based measures, . . .).

We won't cover the details today, but if you want to learn more, see here.

Bayesian Inference

Finding analytical expressions for the posterior is often tricky.

It is, however, possible if we assume a normally distributed prior and likelihood:

- $p(\mu_a) = \mathcal{N}(m_a, s_a)$
- $p(R_t|A_t = a, \mu_a) = \mathcal{N}(\mu_a, \sigma)$

Noise variance σ^2 is known.

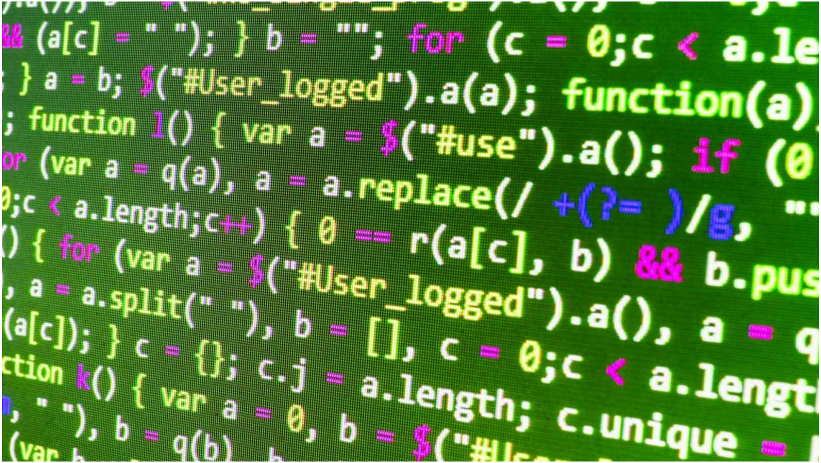
Bayesian Inference

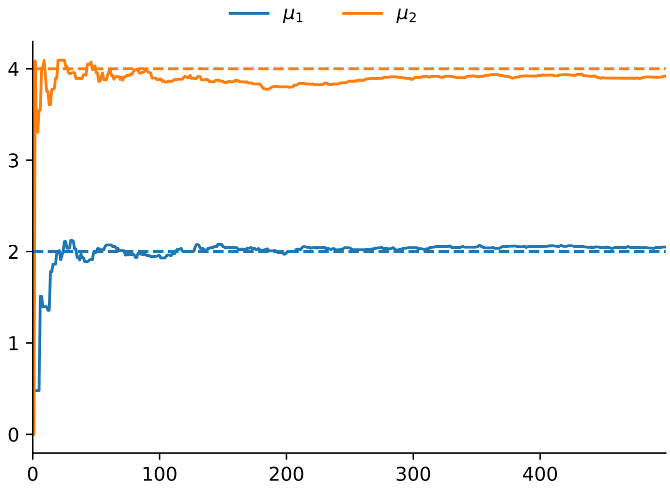
With these assumptions, Bayesian inference amounts to:

$$
\begin{aligned}
\alpha_t &= \frac{s_{a_t}^2}{s_{a_t}^2 + \sigma^2} \\
(m_a, s_a^2) &= \begin{cases} (m_a, s_a^2) & \text{if } a_t \neq a \\ (m_a + \alpha_t(r_t - m_a), s_a^2 - \alpha_t s_a^2) & \text{if } a_t = a \end{cases}
\end{aligned}
$$

We won't cover the derivation today, but if you want to learn more, see this video.

Coding Exercise




μ_1
μ_2
4
3
2
1
0
0
100
200
300
400

Decision-Making

How do people make decisions based on their current beliefs?

General recipe:

1. Come up with a hypothesis.
2. Turn it into a model.
3. Explore the implications of your hypothesis.

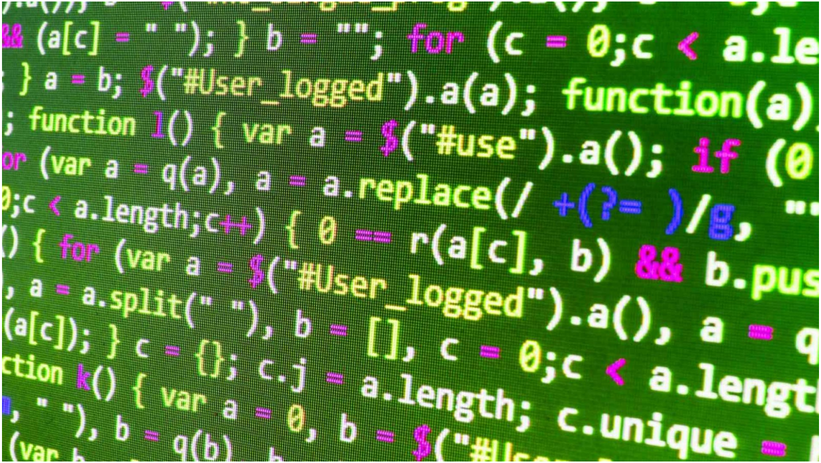
Decision-Making: Exploitation

Hypothesis 1: people always select the action they think is best.

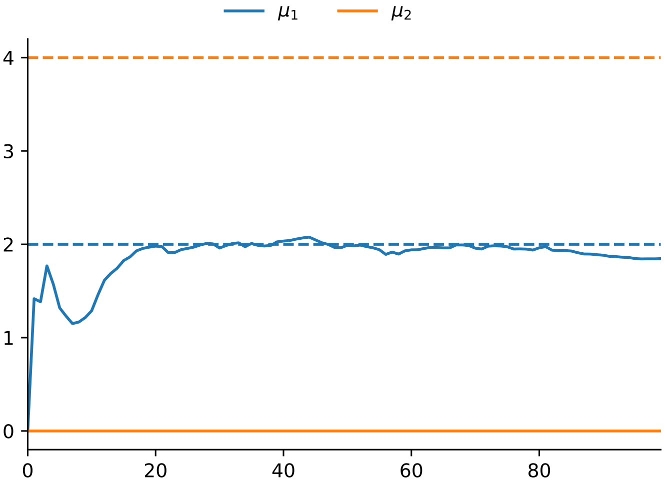
Formally, pick the action with highest expected reward:

$$a_t = \arg\max_{a \in \mathcal{A}} m_a$$

This kind of strategy is also called **exploitation**.


&& (a[c] = " "); } b = ""; for (c = 0;c < a.le
} a = b; $("#User_logged").a(a); function(a)
; function l() { var a = $("#use").a(); if (0
or (var a = q(a), a = a.replace(/ +(?=)/g, "
0;c < a.length;c++) { 0 == r(a[c], b) && b.pus
() { for (var a = $("#User_logged").a(), a = q
, a = a.split(" "), b = [], c = 0;c < a.lengt
(a[c]); } c = {}; c.j = a.length; c.unique =
ction k() { var a = 0, b = $("#Us
, " "), b = q(b)

Decision-Making: Exploitation



Decision-Making: Exploitation

The best action was never selected (and likely will never be selected).

Pure exploitation ignores long term benefits of acquiring new knowledge.

Example of the **exploration-exploitation dilemma**.

Instead of always exploiting current knowledge, we need sometimes need to gather more information.

Decision-Making: Exploration

How to explore?

Previous work has identified two important ideas:

1. **Random exploration:**
   Inject some form of stochasticity into the decision-making process.
2. **Directed exploration:**
   Provide bonus rewards that encourage the agent to visit parts of the environment that ought to be explored.

Decision-Making: Exploration

These are two vague definitions.

No one-to-one correspondence with computational models.

Many different implementations for both random and directed exploration.

Let's take a look at some examples.

Decision-Making: Boltzmann Exploration

Hypothesis 2: people add noise into their decision-making process (random exploration).

Boltzmann exploration provides one way to implement this idea:

$$p(A_t = 1) = \frac{e^{wm_1}}{\sum\limits_{a \in \mathcal{A}} e^{wm_a}}$$

The higher the expected reward, the more likely an action is selected.

The degree of noise is controlled by an inverse temperature parameter w.

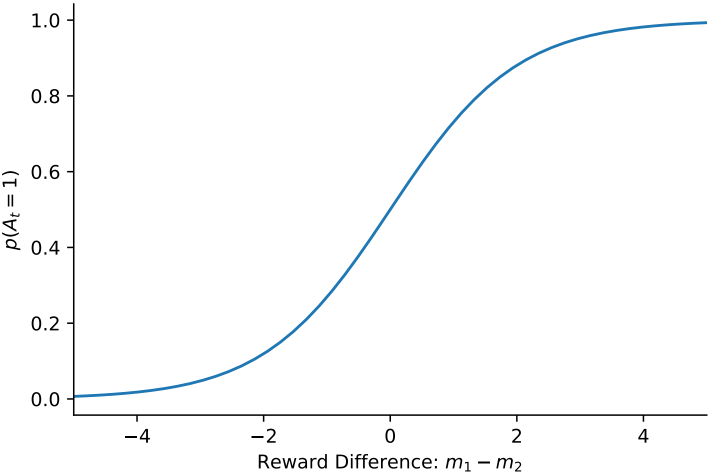
Decision-Making: Boltzmann Exploration

We can rewrite this a bit for the case of two actions:

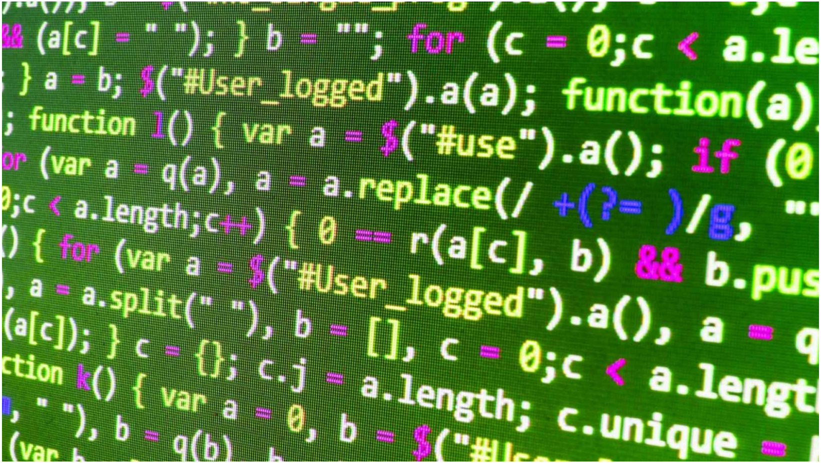
$$
\begin{aligned}
p(A_t = 1) = \frac{e^{wm_1}}{\sum\limits_{a \in \mathcal{A}} e^{wm_a}} &= \frac{e^{wm_1}}{e^{wm_1} + e^{wm_2}} \\
&= \frac{1}{1 + e^{w(m_2 - m_1)}} \\
&= \frac{1}{1 + e^{-w(m_1 - m_2)}} \\
&= \sigma(w\,(m_1 - m_2))
\end{aligned}
$$

Thus, the probability of selecting an action is given by the sigmoid function of the (scaled) reward difference.

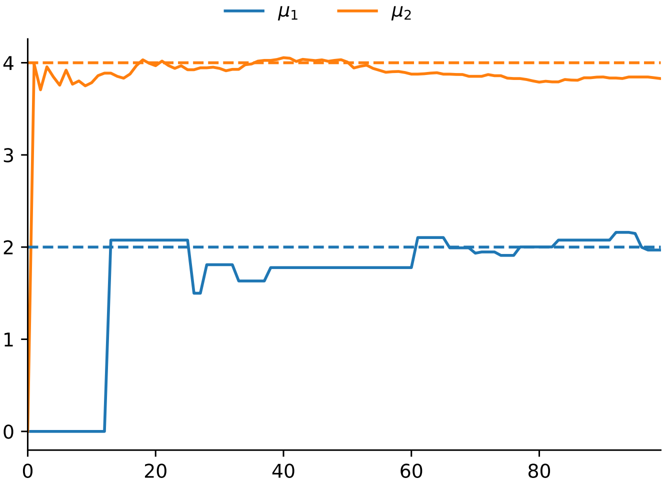
Decision-Making: Boltzmann Exploration



Coding Exercise



Decision-Making: Boltzmann Exploration



Decision-Making: Upper Confidence Bound

Hypothesis 3: people select arms for which they can gather more information (directed exploration).

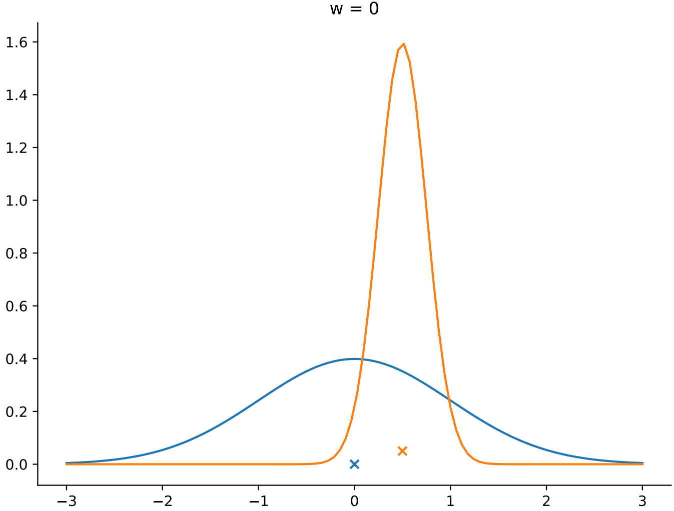
We can use the uncertainty in the posterior distribution to guide exploration:

$$a_t = \arg\max_{a \in \mathcal{A}} \left[m_a + w s_a \right]$$

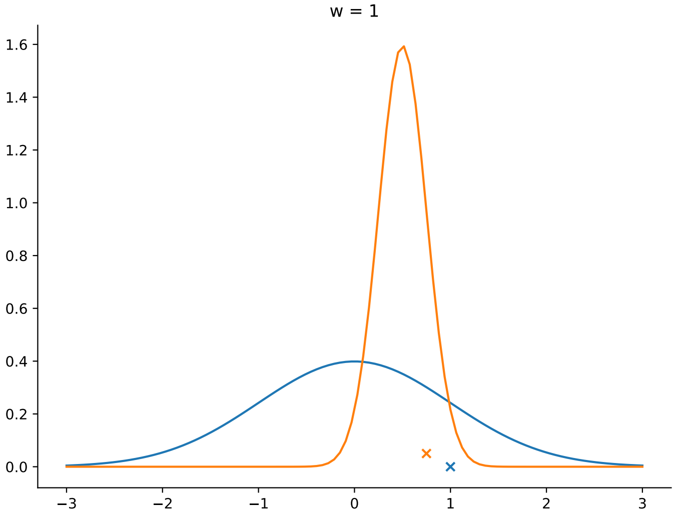
The trade-off between exploration and exploitation is controlled by the parameter w.

This is part of the broader family of **upper confidence bound** (UCB) algorithms.

Decision-Making: Upper Confidence Bound



Decision-Making: Upper Confidence Bound



Decision-Making: Thompson Sampling

Hypothesis 4: people sample action relative to their proportion of being correct.

This idea goes under many names:

- Thompson sampling
- Posterior sampling
- Probability matching
- . . .

It is another example of random exploration.

Decision-Making: Thompson Sampling

There are different ways of implementing this idea.

Implementation 1:

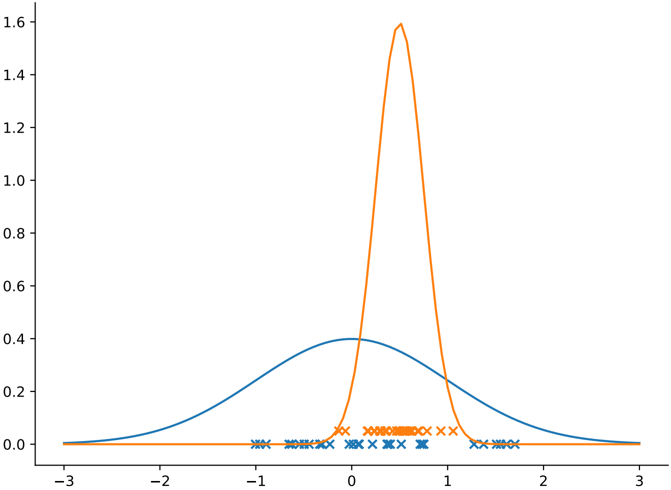
1. Draw a sample from the current posterior:

$$\hat{\mu}_a \sim p(\mu_a)$$

2. Pick the action with the higher sampled value:

$$a_t = \arg\max_{a \in \mathcal{A}} \hat{\mu}_a$$

Decision-Making: Thompson Sampling



Implementation 2:

1. Compute the probability that the mean of action 1 is larger than the mean of action 2, i.e. $p(\mu_1 > \mu_2)$.
2. Draw a sample from the resulting probability distribution.

This also ensure that actions are selected relative to their proportion of being correct.

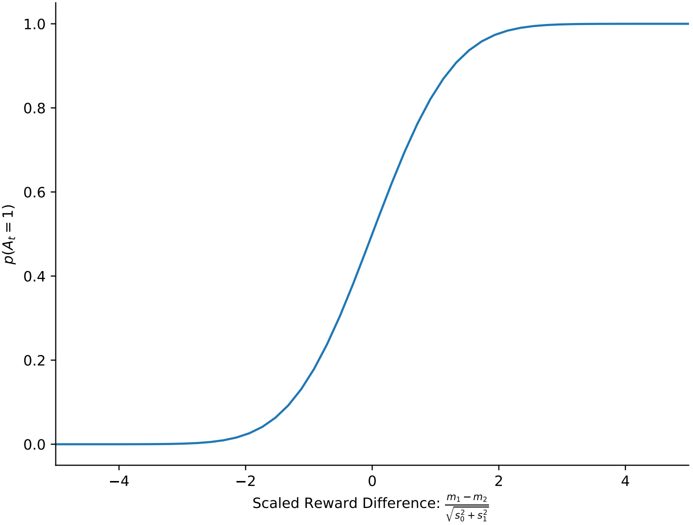
Decision-Making: Thompson Sampling

$p(\mu_1 > \mu_2)$ has a closed-form expression with our assumptions:

$$
\begin{aligned}
p(A_t = 1) &= p(\mu_1 > \mu_2) \\
&= p(\mu_1 - \mu_2 > 0) \\
&= \phi\left(\frac{m_1 - m_2}{\sqrt{s_1^2 + s_2^2}}\right)
\end{aligned}
$$

where ϕ is the cumultative distribution function of a standard normal distribution.

Decision-Making: Thompson Sampling



Decision-Making: Thompson Sampling

This is similar to Boltzmann exploration, with two differences:

- The sigmoid function is replaced by ϕ.
- Reward differences are scaled by the total uncertainty.

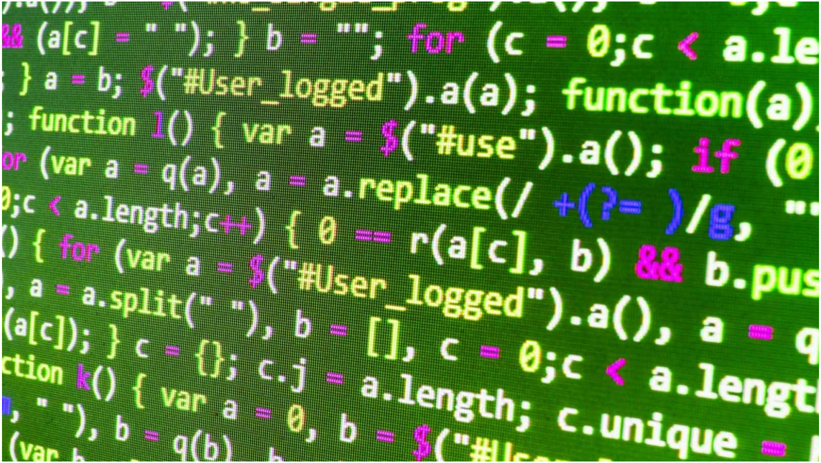
Decision-Making: Thompson Sampling

Both implementations have their own merit.

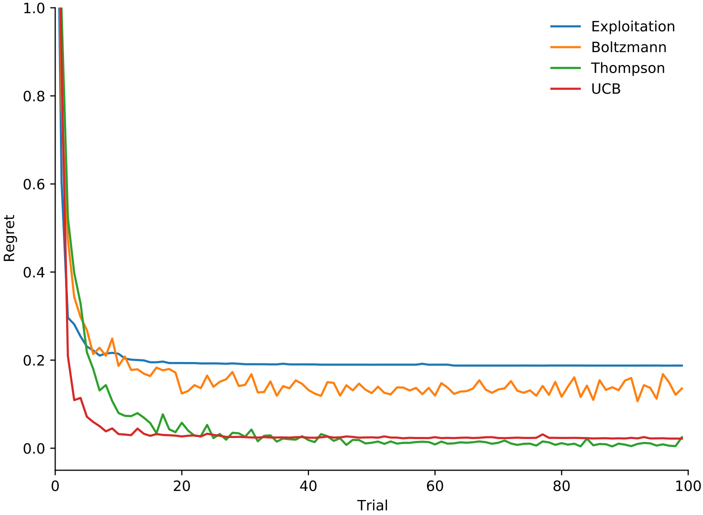
Implementation 1 provides a general recipe that can be applied even if $p(\mu_1 > \mu_2)$ has no closed-form expression.

Implementation 2 provides a closed-form expression of $p(\mu_1 > \mu_2)$, which is necessary for many follow-up analyses.

Coding Exercise



Decision-Making



Decision-Making: Interim Summary

We have discussed four strategies that people may apply to deal with the exploration-exploitation dilemma:

Exploitation: $$a_t = \underset{a \in \mathcal{A}}{\arg\max}\, m_a$$

Boltzmann exploration: $$p(A_t = 1) = \sigma(w\,(m_1 - m_2))$$

Upper confidence bounds: $$a_t = \underset{a \in \mathcal{A}}{\arg\max}\, [m_a + w s_a]$$

Thompson sampling: $$p(A_t = 1) = \phi\left(\frac{m_1 - m_2}{\sqrt{s_1^2 + s_2^2}}\right)$$

Decision-Making: Interim Summary

Pure exploitation | Boltzmann exploration | Upper Confidence Bounds, Thompson sampling | Gittins index, Bayes-Adaptive MDP

Next: which of these strategies are people actually using?

Part 2: Drawing Inferences about Human Decision-Making

Human Decision-Making

Up to now we have:

- Hypothesized how people explore.
- Implemented our hypotheses in computational models.
- Measured their performance.

We really what to gain actual insights in how people explore.

How do we do this?

Human Decision-Making

Here is one general recipe:

1. Design a unified model that contains all of your hypotheses.
2. Include parameters that allow to turn of particular components.
3. Fit parameters to behavioral data and see which of them are significant.

Here is one general recipe:

1. **Design a unified model that contains all of your hypotheses.**
2. **Include parameters that allow to turn of particular components.**
3. Fit parameters to behavioral data and see which of them are significant.

Hybrid Model

How to apply to our setting?

$$p(A_t = 1|\mathbf{w}) = \sigma\left(w_1\left(m_1 - m_2\right) \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \right)$$

Hybrid Model

How to apply to our setting?

$$p(A_t = 1|\mathbf{w}) = \phi\left(w_1\left(m_1 - m_2\right) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\right)$$

Hybrid Model

How to apply to our setting?

$$p(A_t = 1|\mathbf{w}) = \phi\left(w_1\left(m_1 - m_2\right) + w_2\left(\frac{m_1 - m_2}{\sqrt{s_1^2 + s_2^2}}\right) \qquad\qquad\qquad \right)$$

Hybrid Model

How to apply to our setting?

$$p(A_t = 1|\mathbf{w}) = \phi\left(w_1\left(m_1 - m_2\right) + w_2\left(\frac{m_1 - m_2}{\sqrt{s_1^2 + s_2^2}}\right) + w_3\left(s_1 - s_2\right)\right)$$

Hybrid Model

How to apply to our setting?

$$p(A_t = 1|\mathbf{w}) = \phi\left(w_1\,(m_1 - m_2)\, + w_2\left(\frac{m_1 - m_2}{\sqrt{s_1^2 + s_2^2}}\right) + w_3\,(s_1 - s_2)\right)$$

For $\mathbf{w} = [w_1, w_2, w_3] = [w_1, 0, 0]$, we recover a Boltzmann-like strategy.

For $\mathbf{w} = [w_1, w_2, w_3] = [0, 1, 0]$, we recover Thompson sampling.

For $\mathbf{w} = [w_1, w_2, w_3] = [w_1, 0, w_3]$, we recover a noisy version of UCB.

Human Decision-Making

Here is one general recipe:

1. Design a unified model that contains all of your hypotheses.
2. Include parameters that allow to turn of particular components.
3. **Fit parameters to behavioral data and see which of them are significant.**

Parameter Fitting

We have a data-set of choices and corresponding rewards:

$$\mathcal{D} = \{(a_{t,k}, r_{t,k}) \mid 1 \leq t \leq T, 1 \leq k \leq K\}$$

where K is the total number of tasks and T the total number of trials per task.

Now all that is left is to fit the parameters $[w_1, w_2, w_3]$ to the human data.

This is called **parameter fitting**.

Maximum Likelihood Estimation

How should we fit parameters to the data?

Maximum likelihood estimation: Find the parameters $\mathbf{w}$ that assign the highest likelihood $p(A_{t,k} = a_{t,k}|\mathbf{w})$ to human choices.

Turns parameter fitting into an optimization problem.

Higher likelihoods indicate a better fit.

Maximum Likelihood Estimation

When dealing with multiple data-points, we maximize the joint probability:

$$\arg\max_{\mathbf{w}} \prod_{k,t} p(A_{t,k} = a_{t,k} | \mathbf{w})$$

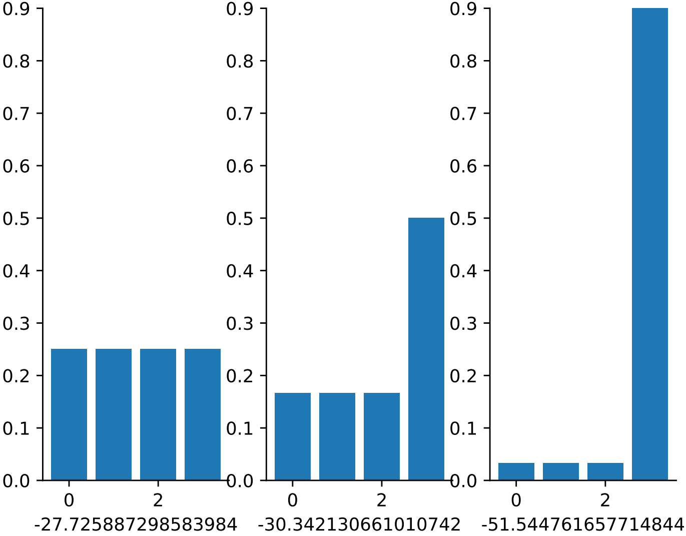
Equivalently, which can maximize the logarithm of this quantity:

$$\arg\max_{\mathbf{w}} \log \prod_{k,t} p(A_{t,k} = a_{t,k} | \mathbf{w})$$
$$= \arg\max_{\mathbf{w}} \sum_{k,t} \log p(A_{t,k} = a_{t,k} | \mathbf{w})$$

This is preferred as it improves numerical stability.

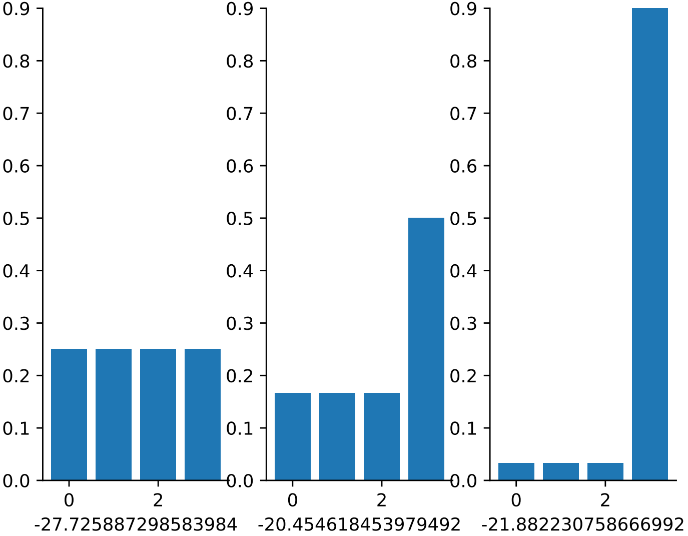
Maximum Likelihood Estimation

[1, 2, 3, 3, 2, 2, 2, 1, 3, 2, 0, 2, 3, 0, 1, 3, 1, 2, 0, 2]

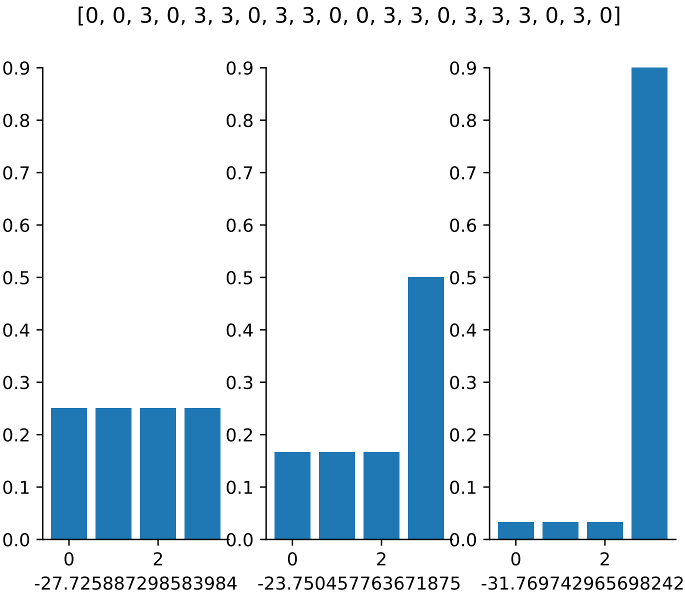


Maximum Likelihood Estimation

[3, 3, 1, 3, 3, 1, 3, 0, 3, 0, 1, 3, 3, 3, 2, 3, 3, 3, 3, 3]



Maximum Likelihood Estimation



Maximum Likelihood Estimation

$$\mathbf{w}^* = \arg\max_{\mathbf{w}} \sum_{k,t} \log p(A_{t,k} = a_{t,k} | \mathbf{w})$$

How do we solve this optimization problem?

- Grid search
- Gradient descent
- Bayesian optimization
- ...

Plotting

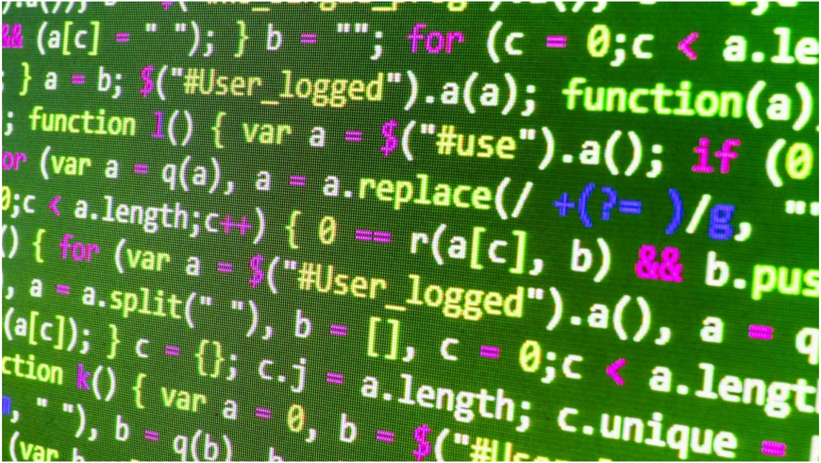
We typically perform parameter fitting for each individual participant.

This is us a set of three regression weights per participant.

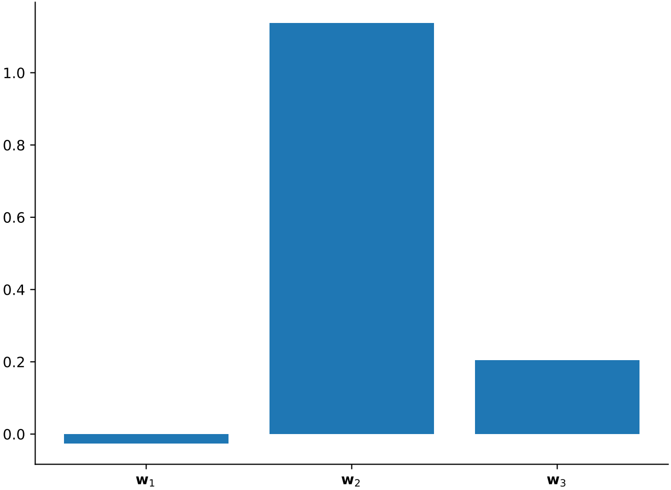
What should we report?

- Report summary statistics (e.g. mean).
- Ideally visualize estimated parameters for all participants.

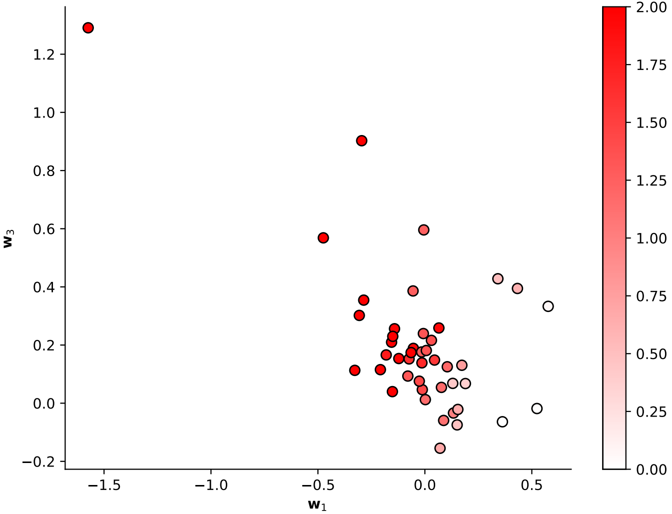
Coding Exercise



Results



Results



$\mathbf{w}_3$
$\mathbf{w}_1$
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−1.5
−1.0
−0.5
0.0
0.5
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00

Results

People employ a combination of random and directed exploration.

People use uncertainty estimates for exploration.

Summary

We have covered:

1. How to build models for a two-armed bandit task.
2. How to use these models to draw inferences about human decision-making.

Where do models come from?

- Introspection
- Previous work
- Rational analysis