

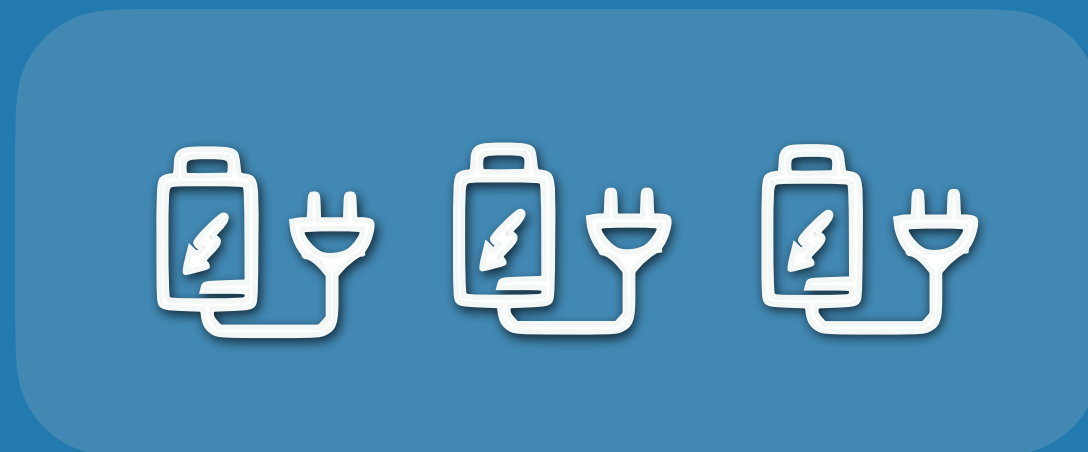
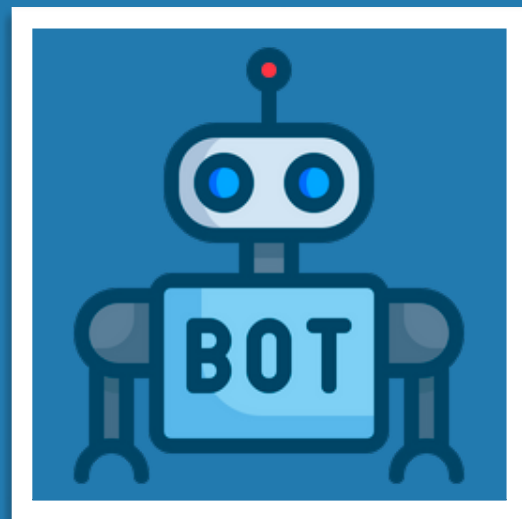
# Thompson Sampling

Thompson Sampling in Multi-Armed Bandits

Hannah Do | Feb 27th, 2022

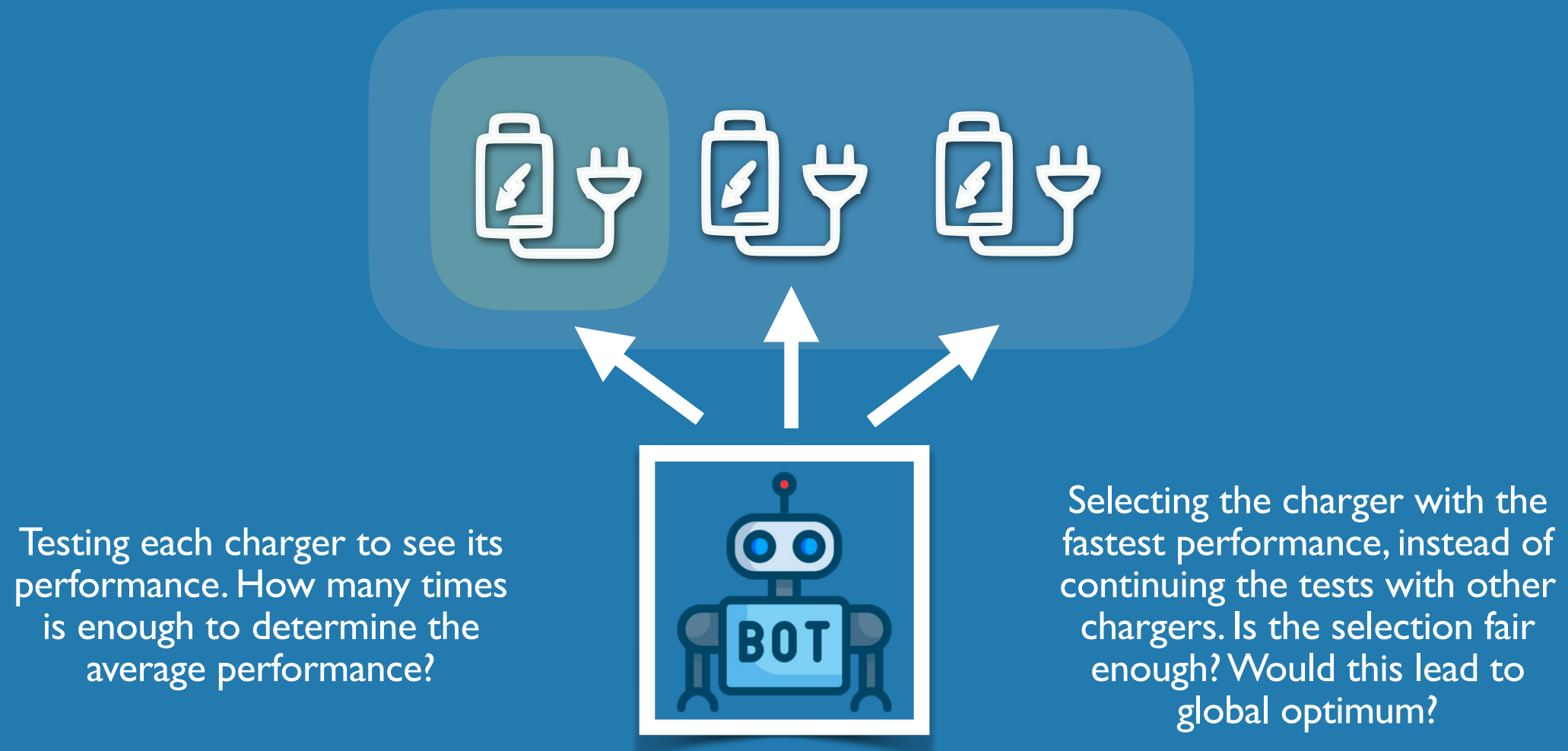
---

**Multi-Armed Bandit (MAB)** is a Machine Learning framework in which an agent has to select actions (arms) in order to maximize its cumulative reward in the long term.



In each round, the agent receives some information about the current state (context), then it chooses an action based on this information and the experience gathered in previous rounds. At the end of each round, the agent receives the reward associated with the chosen action.

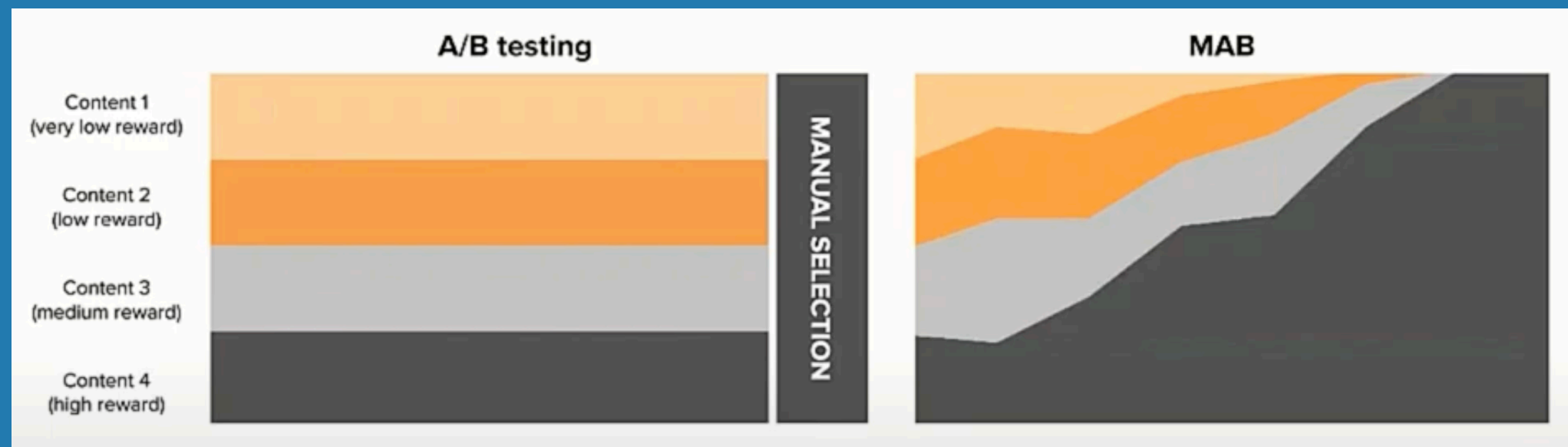
# Exploration VS Exploitation



## Methods :

Greedy Algorithm, Epsilon-first, Epsilon-greedy, UCB1, Thompson Sampling

# Exploration VS Exploitation

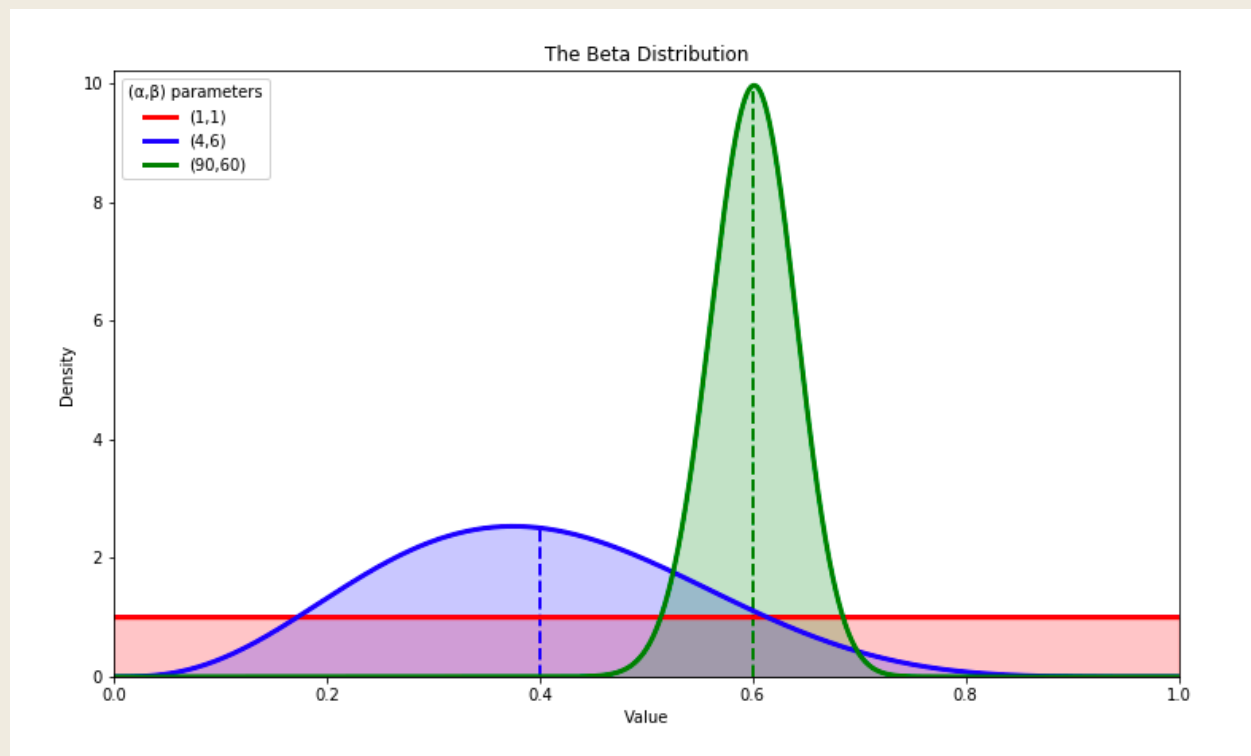


**MAB** continuously assigns traffic to the means that produces the best performance. Takes much less time compared to A/B testing, while trying out all the variations.

# Thompson Sampling

Rather than just refining an estimate of the mean reward, Thompson Sampling (or **Bayesian** Bandits algorithm) builds up a probability model (beta distribution –  $\beta(1,1)$ ) from the obtained rewards, and then samples from this to choose an action.

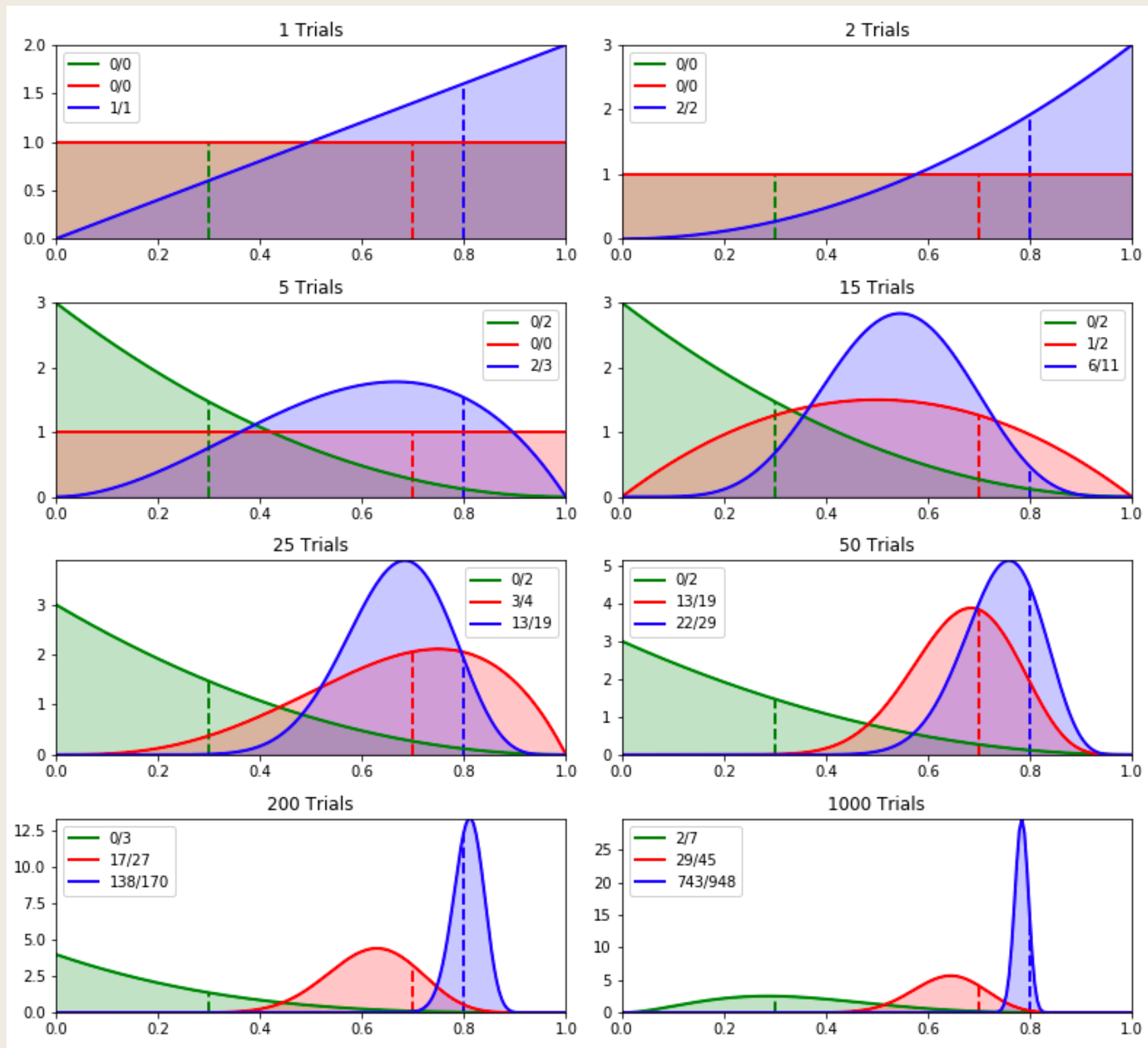
$$\frac{\alpha}{\alpha + \beta} = \frac{\text{number of successes}}{\text{total number of trials}}$$



If the charger returns some energy, the reward will be 1 and ‘ $\alpha$ ’, the count of the number of successes, will increment by 1.

If no reward was obtained, then ‘ $\alpha$ ’ will stay the same and ‘ $\beta$ ’ will increment by 1.

## THOMPSON SAMPLING



# Summary

Multi-Armed Bandit (MAB) can be used for **Website Optimization**. It can help you boost website traffic, convert visitors to customers, and increase many other measures of success.

- Different User Interfaces
- Recommendation Models

# References

1. **Multi-armed Bandit explained with Practical Examples** Frosmo Ltd. <https://www.youtube.com/watch?v=VrFZCGCwzVc&t=225s>
  2. **Intro to Bandit** Tensorflow - [https://www.tensorflow.org/agents/tutorials/intro\\_bandit](https://www.tensorflow.org/agents/tutorials/intro_bandit)
  3. **Thompson Sampling** Steve Roberts (Nov 2, 2020) - <https://towardsdatascience.com/thompson-sampling-fc28817eacb8>
-