Multi-Armed Bandit

Schwartz

September 24, 2017

The Multi-Armed Bandit

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

"Discussion of Dr Gittins' paper", Journal of the Royal Stat. Society, Series B, issue 41, 1979.

The Multi-Armed Bandit

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

"Discussion of Dr Gittins' paper", Journal of the Royal Stat. Society, Series B, issue 41, 1979.

The Multi-Armed Bandit

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

"Discussion of Dr Gittins' paper", Journal of the Royal Stat. Society, Series B, issue 41, 1979.

Objectives

1. Bayesian versus Frequentist A/B testing
2. Bayesian Priors
3. Multi-armed bandit deployment
   - Regret: all about it and having none of it
   - ϵ-greedy, softmax, UCB1, and the Bayesian Bandit

A/B Testing reminder

		
Trial 1		
Trial 2		
Trial 3		
Trial 4		
$\vdots$		
Trial $n_A + n_B$		
Total		

A/B Testing reminder

		
Trial 1	$\emptyset$	
Trial 2		
Trial 3		
Trial 4		
$\vdots$		
Trial $n_A + n_B$		
Total		

A/B Testing reminder

		
Trial 1	$\emptyset$	
Trial 2		$\emptyset$
Trial 3		
Trial 4		
$\vdots$		
Trial $n_A + n_B$		
Total		

A/B Testing reminder

		
Trial 1	$\emptyset$	
Trial 2		$\emptyset$
Trial 3		✓
Trial 4		
$\vdots$		
Trial $n_A + n_B$		
Total		

A/B Testing reminder

		
Trial 1	$\emptyset$	
Trial 2		$\emptyset$
Trial 3		$\checkmark$
Trial 4	$\emptyset$	
$\vdots$		
Trial $n_A + n_B$		
Total		

A/B Testing reminder

		
Trial 1	$\emptyset$	
Trial 2		$\emptyset$
Trial 3		$\checkmark$
Trial 4	$\emptyset$	
$\vdots$		
Trial $n_A + n_B$		
Total	$\hat{p}_A = \frac{\sum X_A^{(i)}}{n_A}$	$\hat{p}_B = \frac{\sum X_B^{(j)}}{n_B}$

A/B Testing reminder

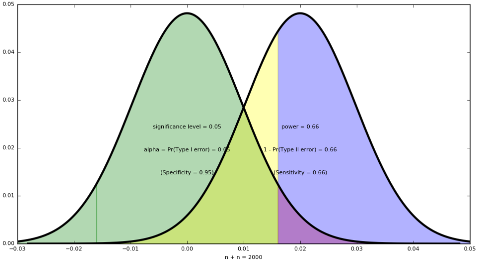
$X_A^{(i)} \sim \text{Bern}(\theta_A)$ $X_B^{(j)} \sim \text{Bern}(\theta_B)$		
Trial 1	$\emptyset$	
Trial 2		$\emptyset$
Trial 3		$\checkmark$
Trial 4	$\emptyset$	
$\vdots$		
Trial $n_A + n_B$		
Total	$\hat{p}_A = \frac{\sum X_A^{(i)}}{n_A}$	$\hat{p}_B = \frac{\sum X_B^{(j)}}{n_B}$

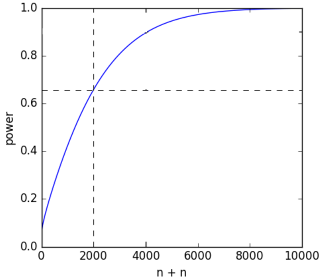
$$\hat{p}_A - \hat{p}_B \sim N\left(\theta_A - \theta_B, \frac{\theta_A(1-\theta_A)}{n_A} + \frac{\theta_B(1-\theta_B)}{n_B}\right)$$

and if $\theta_A = \theta_B = \theta$ then $\hat{p}_A - \hat{p}_B \sim N\left(0, \frac{2\hat{p}(1-\hat{p})}{n_A+n_B}\right)$

A/B **Hypothesis Testing** *power calculations*

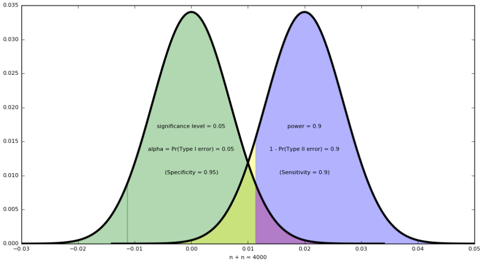
- Effect size is $0.06 - 0.04$
- With $2n$ alternating trials
  $\hat{p} = 0.05$
- $se = \sqrt{2\frac{\hat{p}(1-\hat{p})}{n}}$

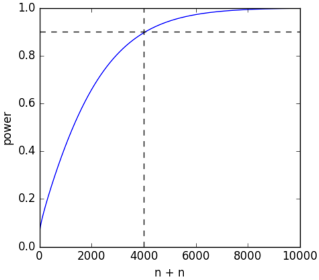




A/B **Hypothesis Testing** *power calculations*

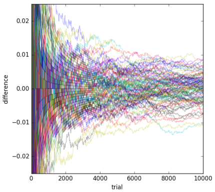
- Effect size is $0.06 - 0.04$
- With $2n$ alternating trials
  $\hat{p} = 0.05$
- $se = \sqrt{2\frac{\hat{p}(1-\hat{p})}{n}}$

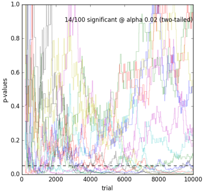


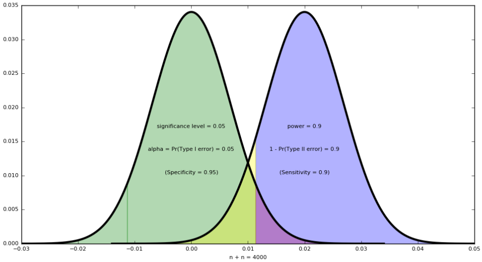


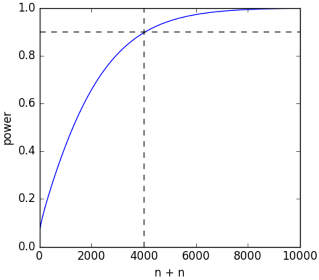
A/B **Hypothesis Testing** *power calculations*

- Effect size is $0.06 - 0.04$
- With $2n$ alternating trials $\hat{p} = 0.05$
- $se = \sqrt{2\frac{\hat{p}(1-\hat{p})}{n}}$







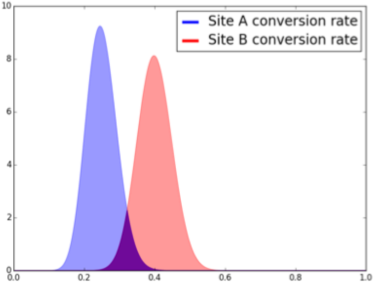


Bayesian *Philosophy*

- Rather than θ_A & θ_B being fixed constants to be estimated...

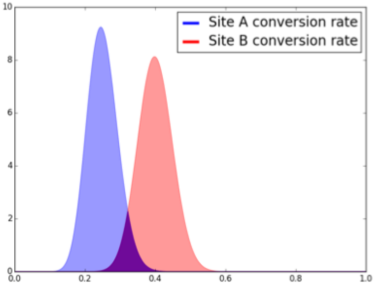
Bayesian *Philosophy*

- Rather than θ_A & θ_B being fixed constants to be estimated...
- they are random variables with distributions $p(\theta_A)$ & $p(\theta_B)$



Bayesian *Philosophy*

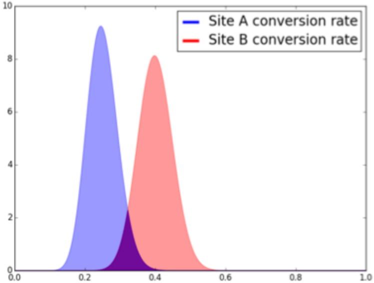
- Rather than θ_A & θ_B being fixed constants to be estimated...
- they are random variables with distributions $p(\theta_A)$ & $p(\theta_B)$



- The distributions model the parameter uncertainty given data
- And give probability statements of beliefs about parameters

Bayesian *Philosophy*

- Rather than θ_A & θ_B being fixed constants to be estimated...
- they are random variables with distributions $p(\theta_A)$ & $p(\theta_B)$

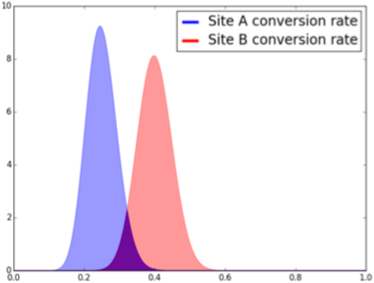


- The distributions model the parameter uncertainty given data
- And give probability statements of beliefs about parameters

  E.g., $Pr(\theta_A > \theta_B)$, $Pr(\theta_A - \theta_B > 0.05)$, or $Pr(\frac{\theta_A}{\theta_B} > 1.1)$

Bayesian *Philosophy*

- Rather than θ_A & θ_B being fixed constants to be estimated...
- they are random variables with distributions $p(\theta_A)$ & $p(\theta_B)$



- The distributions model the parameter uncertainty given data
- And give probability statements of beliefs about parameters

  E.g., $Pr(\theta_A > \theta_B)$, $Pr(\theta_A - \theta_B > 0.05)$, or $Pr(\frac{\theta_A}{\theta_B} > 1.1)$

- Previously, confidence intervals/p-values characterized the (sampling) variability of statistical estimation procedures

Questions snarky Bayesians ask Frequentists

- Can you say “there’s 95% probability that A beats B”?

Questions snarky Bayesians ask Frequentists

- Can you say "there's 95% probability that A beats B"?
  Nope. But I can.

Questions snarky Bayesians ask Frequentists

- Can you say "there's 95% probability that A beats B"?
  Nope. But I can.
- Can you stop the test early based on surprising results?

Questions snarky Bayesians ask Frequentists

- Can you say "there's 95% probability that A beats B"?
  Nope. But I can.

- Can you stop the test early based on surprising results?
  Nope. But I can.

Questions snarky Bayesians ask Frequentists

- Can you say “there’s 95% probability that A beats B”?
  Nope. But I can.
- Can you stop the test early based on surprising results?
  Nope. But I can.
- Can you update the test parameters while it’s running?

Questions snarky Bayesians ask Frequentists

- Can you say “there’s 95% probability that A beats B”?
  Nope. But I can.
- Can you stop the test early based on surprising results?
  Nope. But I can.
- Can you update the test parameters while it’s running?
  Nope. But I can.

Bayesian *Theory*

- Data X that informs our knowledge of θ_A and θ_B allows us to "update" our beliefs about θ_A and θ_B using *Bayes' Theorem*

$$f(\theta|X) = \frac{f(X|\theta)f(\theta)}{f(X)}$$

Bayesian *Theory*

- Data X that informs our knowledge of θ_A and θ_B allows us to "update" our beliefs about θ_A and θ_B using *Bayes' Theorem*

$$f(\theta|X) = \frac{f(X|\theta)f(\theta)}{f(X)}$$

- Bayes' Theorem derives the posterior (theta given data) as a function of the likelihood, the prior and the marginal likelihood

Bayesian *Theory*

- Data X that informs our knowledge of θ_A and θ_B allows us to "update" our beliefs about θ_A and θ_B using *Bayes' Theorem*

$$f(\theta|X) = \frac{f(X|\theta)f(\theta)}{f(X)}$$

- Bayes' Theorem derives the posterior (theta given data) as a function of the likelihood, the prior and the marginal likelihood
- The posterior is (marginal likelihood) proportional to:

$$f(\theta|X) \propto f(X|\theta)f(\theta)$$

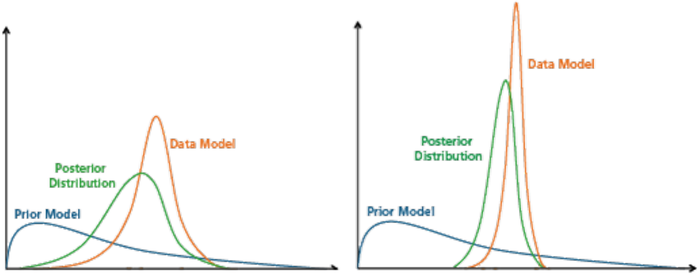
Bayesian *Theory*

- Data X that informs our knowledge of θ_A and θ_B allows us to "update" our beliefs about θ_A and θ_B using *Bayes' Theorem*

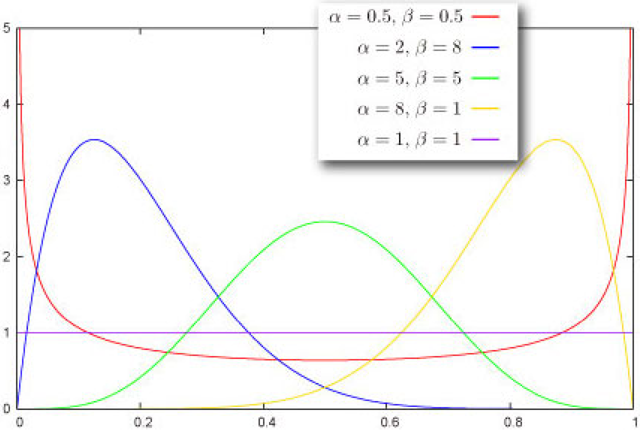
$$f(\theta|X) = \frac{f(X|\theta)f(\theta)}{f(X)}$$

- Bayes' Theorem derives the posterior (theta given data) as a function of the likelihood, the prior and the marginal likelihood
- The posterior is (marginal likelihood) proportional to:

$$f(\theta|X) \propto f(X|\theta)f(\theta)$$



Beta Distribution



$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

Bayesian A/B Testing

- Likelihood

$$
\begin{aligned}
f(X_1, X_2, \cdots, X_n|\theta) &= f(X_1|\theta) f(X_2|\theta) \cdots f(X_n|\theta) \\
&= \prod_{i=1}^{n} \theta^{X_i} (1-\theta)^{1-X_i} \qquad \text{[beta(s)]} \\
&= \theta^{\sum_{i=1}^{n} X_i} (1-\theta)^{n - \sum_{i=1}^{n} X_i}
\end{aligned}
$$

Bayesian A/B Testing

- Likelihood

$$\begin{aligned} f(X_1, X_2, \cdots, X_n|\theta) &= f(X_1|\theta)f(X_2|\theta)\cdots f(X_n|\theta) \\ &= \prod_{i=1}^{n} \theta^{X_i}(1-\theta)^{1-X_i} \qquad \text{[beta(s)]} \\ &= \theta^{\sum_{i=1}^{n} X_i}(1-\theta)^{n-\sum_{i=1}^{n} X_i} \end{aligned}$$

- Conjugate* Prior

$$f(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \qquad \text{[beta]}$$

Bayesian A/B Testing

- Likelihood

$$f(X_1, X_2, \cdots, X_n|\theta) = f(X_1|\theta)f(X_2|\theta)\cdots f(X_n|\theta)$$
$$= \prod_{i=1}^{n} \theta^{X_i}(1-\theta)^{1-X_i} \qquad \text{[beta(s)]}$$
$$= \theta^{\sum_{i=1}^{n} X_i}(1-\theta)^{n-\sum_{i=1}^{n} X_i}$$

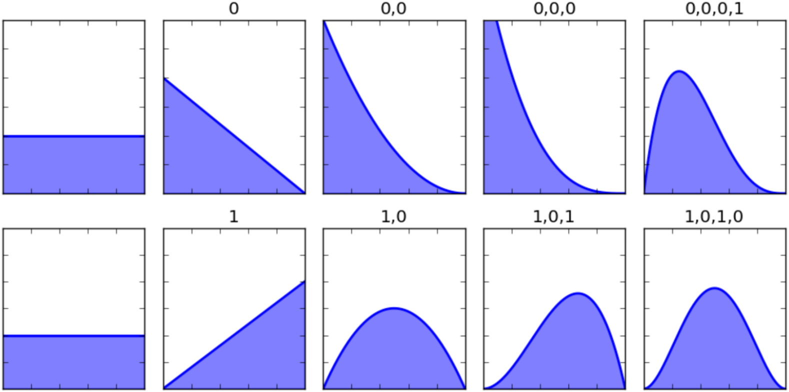
- Conjugate* Prior

$$f(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1} \qquad \text{[beta]}$$

- Posterior

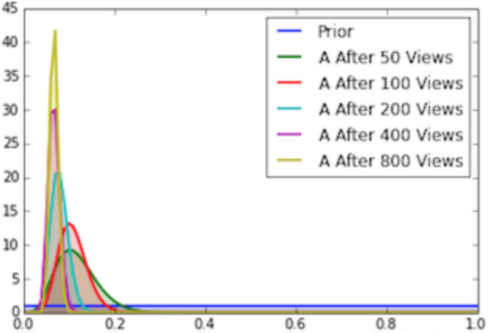
$$f(\theta|X) \propto \theta^{\sum X_i}(1-\theta)^{n-\sum X_i}\theta^{\alpha-1}(1-\theta)^{\beta-1}$$
$$= \theta^{\alpha+\sum X_i-1}(1-\theta)^{\beta+n-\sum X_i-1} \qquad \text{[beta]}$$

Beta Distribution Updates



$$\text{Beta}\left(\alpha + \sum X_i, \beta + n - \sum X_i\right)$$

Beta Distribution Updates



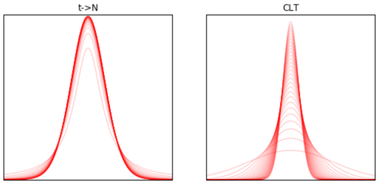
Probability of Success

Rev. Thomas Bayes (1701 – 1761)

$$f(\theta|X) = \frac{f(X|\theta)f(\theta)}{f(X)} = \text{Beta}\left(\alpha + \sum X_i, \beta + n - \sum X_i\right)$$

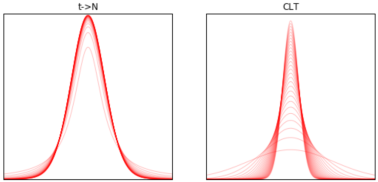
A brief synopsis of the world of the world to date

A brief synopsis of the world of the world to date

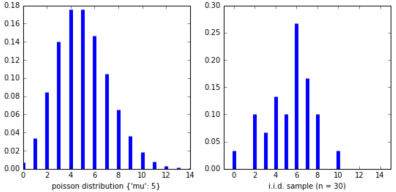


$$n \longrightarrow \infty$$

A brief synopsis of the world of the world to date

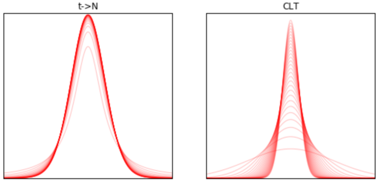


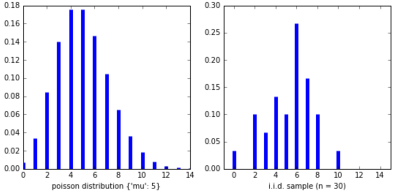
$n \longrightarrow \infty$



nonparametrics

A brief synopsis of the world of the world to date



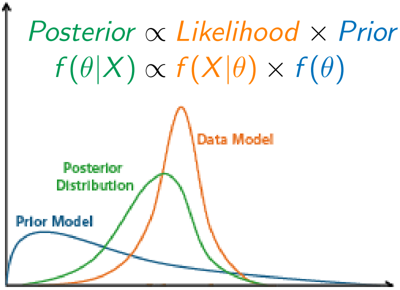


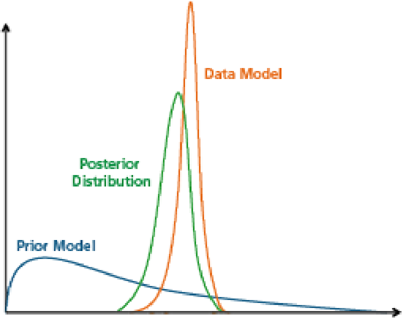
$n \longrightarrow \infty$

nonparametrics

$$\textit{Posterior} \propto \textit{Likelihood} \times \textit{Prior}$$

$$f(\theta|X) \propto f(X|\theta) \times f(\theta)$$





Parametric Versus Nonparametric

We might characterize an analysis framework as parametric if

1. Results are buttressed or bolstered by modeling assumptions

Parametric Versus Nonparametric

We might characterize an analysis framework as parametric if

1. Results are buttressed or bolstered by modeling assumptions
   E.g. leveraging the structure of normality via a z-test increases power but comes at a cost of loss of robustness compared to nonparametric tests free of distributional assumptions.

Parametric Versus Nonparametric

We might characterize an analysis framework as parametric if

1. Results are buttressed or bolstered by modeling assumptions
   E.g. leveraging the structure of normality via a z-test increases power but comes at a cost of loss of robustness compared to nonparametric tests free of distributional assumptions.

2. Predicted values are based on "parameters."

Parametric Versus Nonparametric

We might characterize an analysis framework as parametric if

1. Results are buttressed or bolstered by modeling assumptions
   E.g. leveraging the structure of normality via a z-test increases power but comes at a cost of loss of robustness compared to nonparametric tests free of distributional assumptions.

2. Predicted values are based on "parameters."
   E.g., the β coefficients in linear regression.

Parametric Versus Nonparametric

We might characterize an analysis framework as parametric if

1. Results are buttressed or bolstered by modeling assumptions
   E.g. leveraging the structure of normality via a z-test increases power but comes at a cost of loss of robustness compared to nonparametric tests free of distributional assumptions.

2. Predicted values are based on "parameters."
   E.g., the β coefficients in linear regression.

3. Parameter estimation determines the specific instance of a model within a "model class" defined by those parameters.

Parametric Versus Nonparametric

We might characterize an analysis framework as parametric if

1. Results are buttressed or bolstered by modeling assumptions
   E.g. leveraging the structure of normality via a z-test increases power but comes at a cost of loss of robustness compared to nonparametric tests free of distributional assumptions.

2. Predicted values are based on "parameters."
   E.g., the β coefficients in linear regression.

3. Parameter estimation determines the specific instance of a model within a "model class" defined by those parameters.
   E.g., the CLT guarantees a normal distribution which is determined by estimating μ and $\sigma^2/\sqrt{n}$ with n large

Parametric Versus Nonparametric

We might characterize an analysis framework as parametric if

1. Results are buttressed or bolstered by modeling assumptions
   E.g. leveraging the structure of normality via a z-test increases power but comes at a cost of loss of robustness compared to nonparametric tests free of distributional assumptions.

2. Predicted values are based on "parameters."
   E.g., the β coefficients in linear regression.

3. Parameter estimation determines the specific instance of a model within a "model class" defined by those parameters.
   E.g., the CLT guarantees a normal distribution which is determined by estimating μ and $\sigma^2/\sqrt{n}$ with n large

4. The complexity of the model does grows as data size n grows.

Parametric Versus Nonparametric

We might characterize an analysis framework as parametric if

1. Results are buttressed or bolstered by modeling assumptions
   E.g. leveraging the structure of normality via a z-test increases power but comes at a cost of loss of robustness compared to nonparametric tests free of distributional assumptions.

2. Predicted values are based on "parameters."
   E.g., the β coefficients in linear regression.

3. Parameter estimation determines the specific instance of a model within a "model class" defined by those parameters.
   E.g., the CLT guarantees a normal distribution which is determined by estimating μ and $\sigma^2/\sqrt{n}$ with n large

4. The complexity of the model does grows as data size n grows.
   E.g. trees grow in complexity as data becomes richer while a normal distribution is defined by μ and σ^2 regardless of n.

The Multi-Armed Bandit

The Multi-Armed Bandit

Truth / Trial	0.01	0.02	0.04	0.03	0.05	0.04	0.02
1							
2							
⋮							

The Multi-Armed Bandit

Truth / Trial	0.01	0.02	0.04	0.03	0.05	0.04	0.02
1							
2							
⋮							

Exploration & Exploitation

The Multi-Armed Bandit and *Regret*

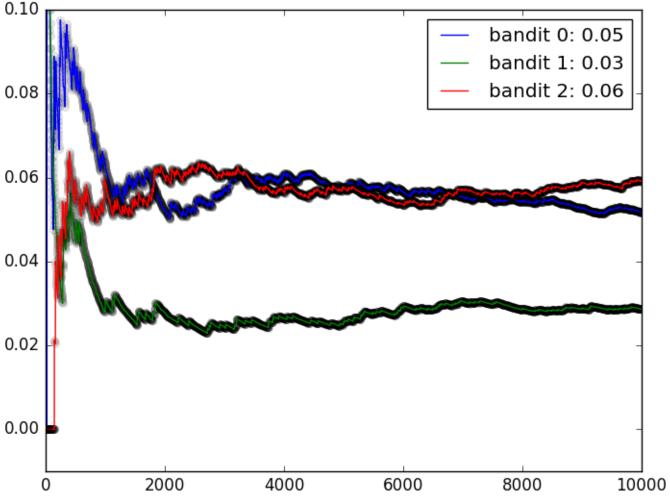
$$\text{regret} = \sum_t^T \max_k \theta_k - \theta^{(t)}$$

$$= T \cdot \max_k \theta_k - \sum_t^T \theta^{(t)}$$

The Multi-Armed *random* Bandit

Random: *A/B[/C/D/E/F/G]-test*

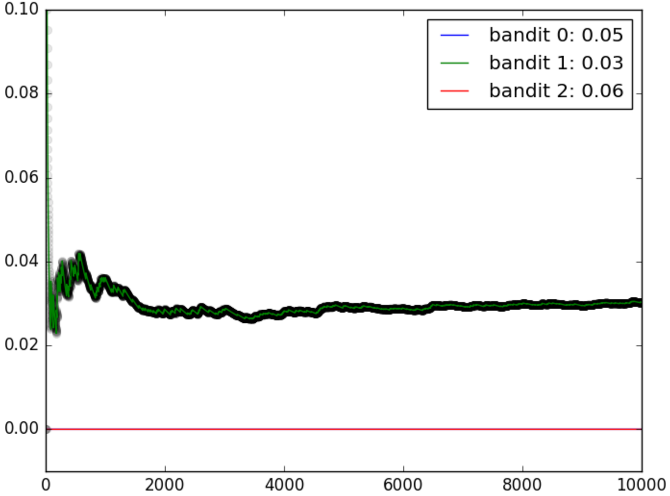
[when should you test?]



The Multi-Armed *max* Bandit

Max: *use the (currently) best performing option*

[perhaps after a burn-in?]

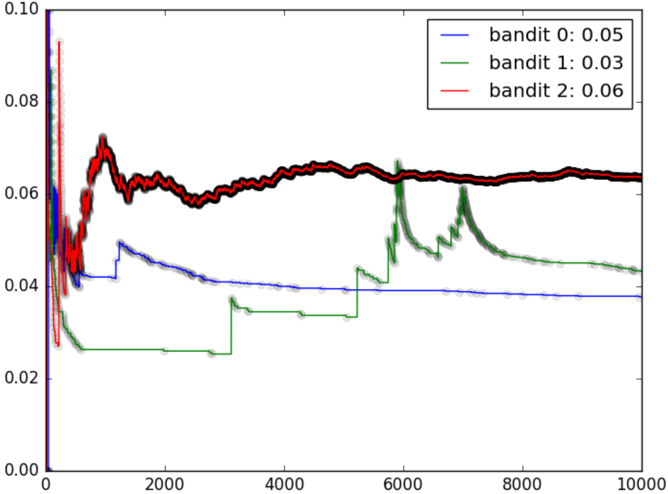


The Multi-Armed *softmax* Bandit

Softmax: *Select proportionally to the currently estimated success rates*

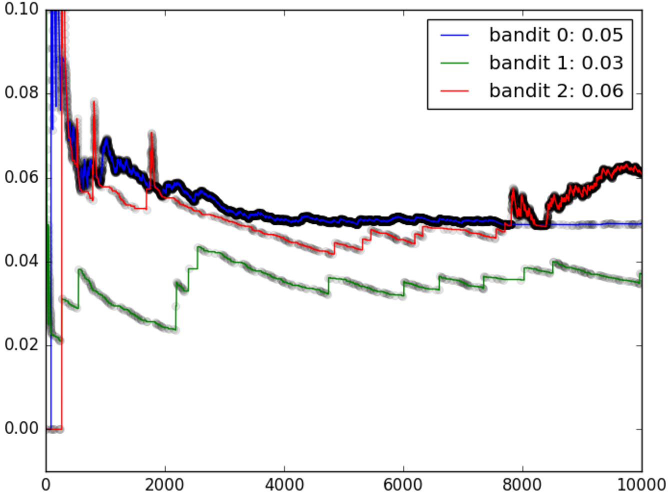
$$\Pr(\text{select } k) \propto \exp(\hat{p}_k/\tau), \text{ i.e.,}$$

$$\Pr(\text{select } k) = \frac{\exp(\hat{p}_k/\tau)}{\sum \exp(\hat{p}_k/\tau)}$$



The Multi-Armed ϵ-*greedy* Bandit

ϵ-greedy: *Select randomly with probability* ϵ
otherwise use the current best option

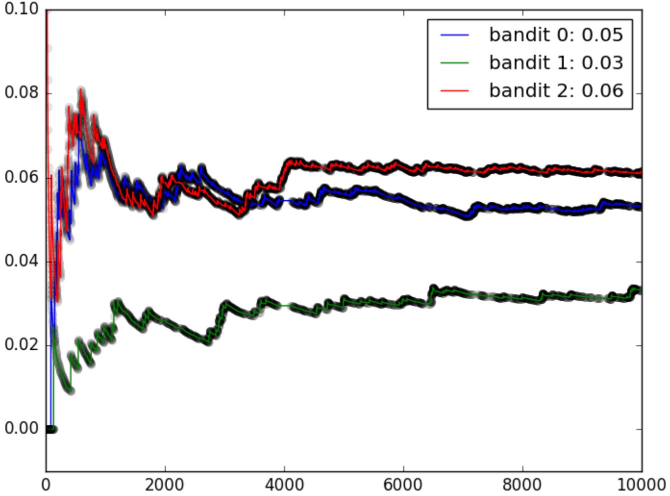


The Multi-Armed *UCB1* Bandit

UCB1: *Select the option which has the highest* **possible** *conversion potential*

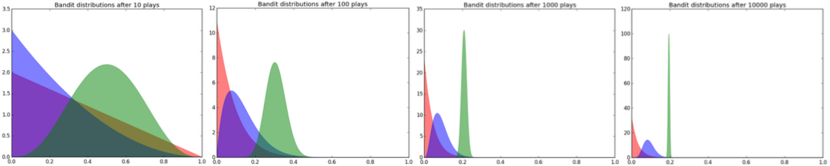
$$\max_k \hat{p}_k + \sqrt{\frac{2 log \sum n_k}{n_k}}$$

with n_k trials for bandit k



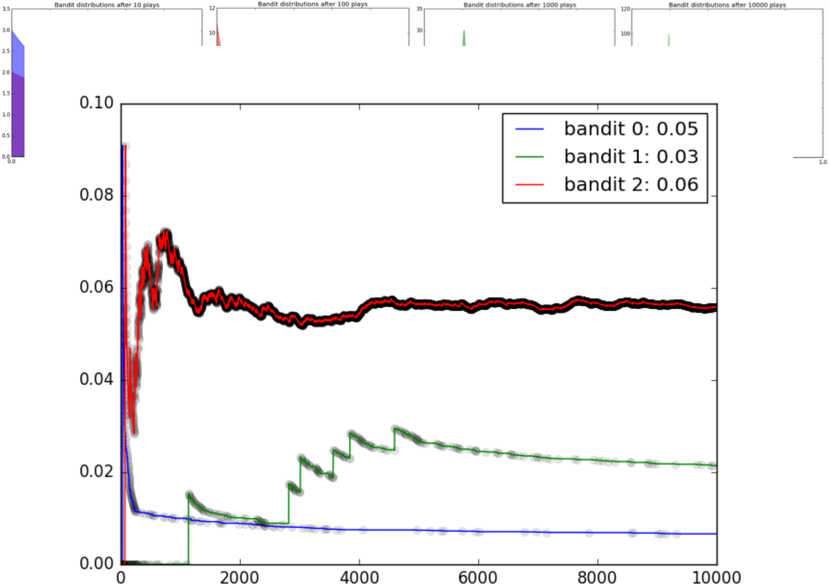
The Multi-Armed *Bayesian* Bandit

- *Sample current posteriors $\rightarrow$ Use the best $\rightarrow$ Update that posterior*



The Multi-Armed *Bayesian* Bandit

- *Sample current posteriors* $\rightarrow$ *Use the best* $\rightarrow$ *Update that posterior*



The Multi-Armed Bandit

Exploration & Exploitation