

# Dilemme exploration- exploitation en recherche

Nicolas Guiard

Aix-Marseille Université - DirNum

# À propos de moi

Nicolas Guiard

- **Poste** : Ingénieur logiciel à la DirNum d'AMU
- **Intérêts** : statistiques, machine learning, psychologie, cybersécurité, méthodologie de la recherche, quiz adaptatifs (AMUQuiz), algorithmes...

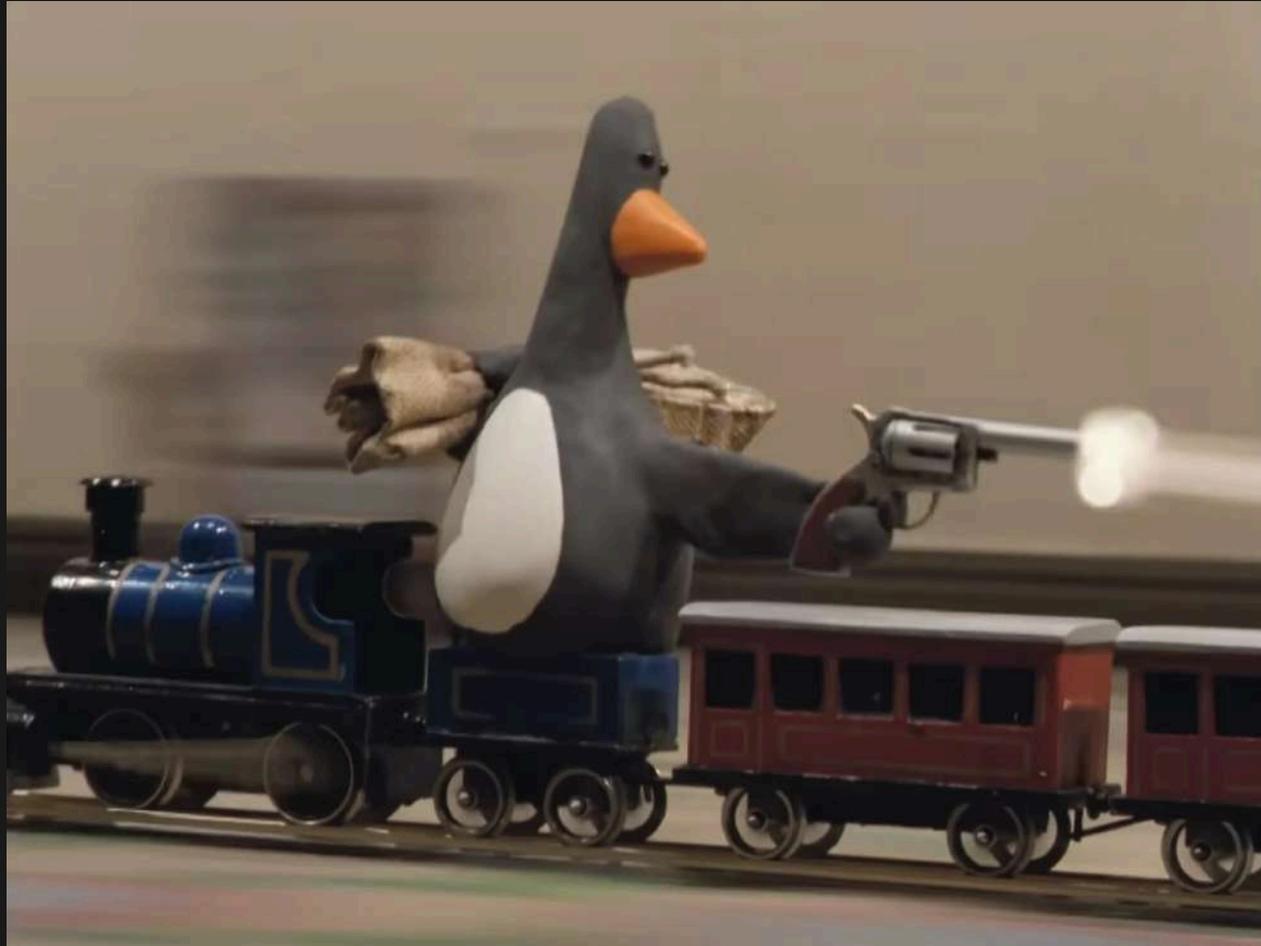
Cette présentation :

- 3 parties
- focus sur les méthodes
- pauses questions

# Pourquoi ça peut vous intéresser ?

- Désir parfois implicite d'optimiser quelque chose :
  - réduire l'anxiété
  - augmenter le bonheur
  - améliorer l'apprentissage
  - ...
- Juste comprendre le fonctionnement d'un système ?
- Quelque chose de meta
- Points communs, vastes applications

# Bandit manchot



# Votre mission

- Une nouvelle maladie inconnue touche le peuple des Kerbal
- Plusieurs traitements candidats
- Sauver le maximum
- Multi-armed bandit... Traduit en français par “bandit manchot”
- Problème d’optimisation

**KERBAL**  
SPACE PROGRAM



# Stratégie : Oracle

Si vous savez déjà quel est le meilleur traitement... C'est facile.

# Stratégie : Epsilon-first

Le fameux RCT (Randomized Controlled Trial) ou essai contrôlé randomisé.

# Stratégie : Thompson Sampling

- Thompson (1933)
- Redécouvert indépendamment plusieurs fois

# Distribution Beta

Intuition...

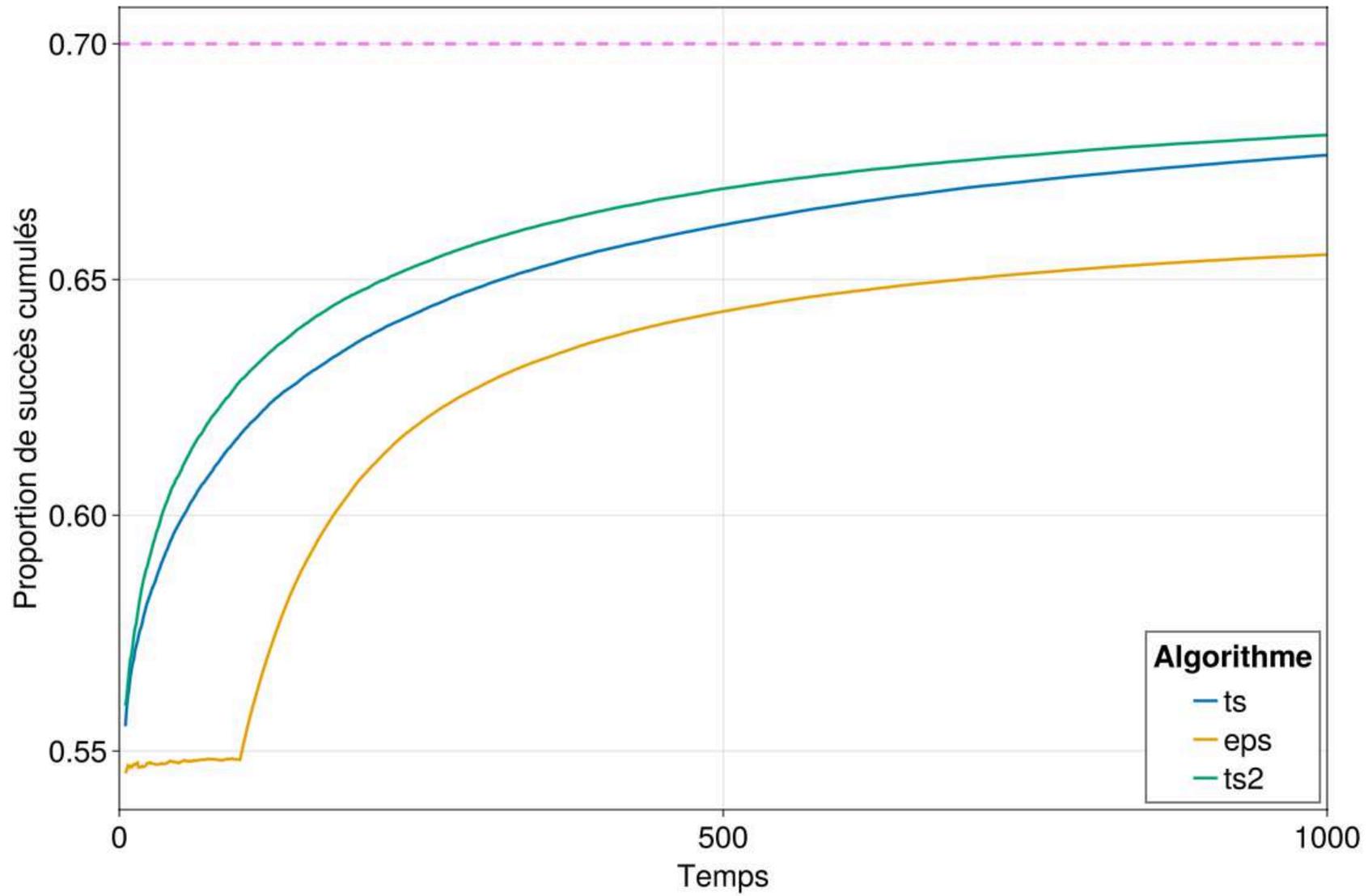
# Thompson Sampling : Demo

Démo !

# Autres métriques

- Typiquement : minimiser le *regret* (ou maximiser la proportion de succès)

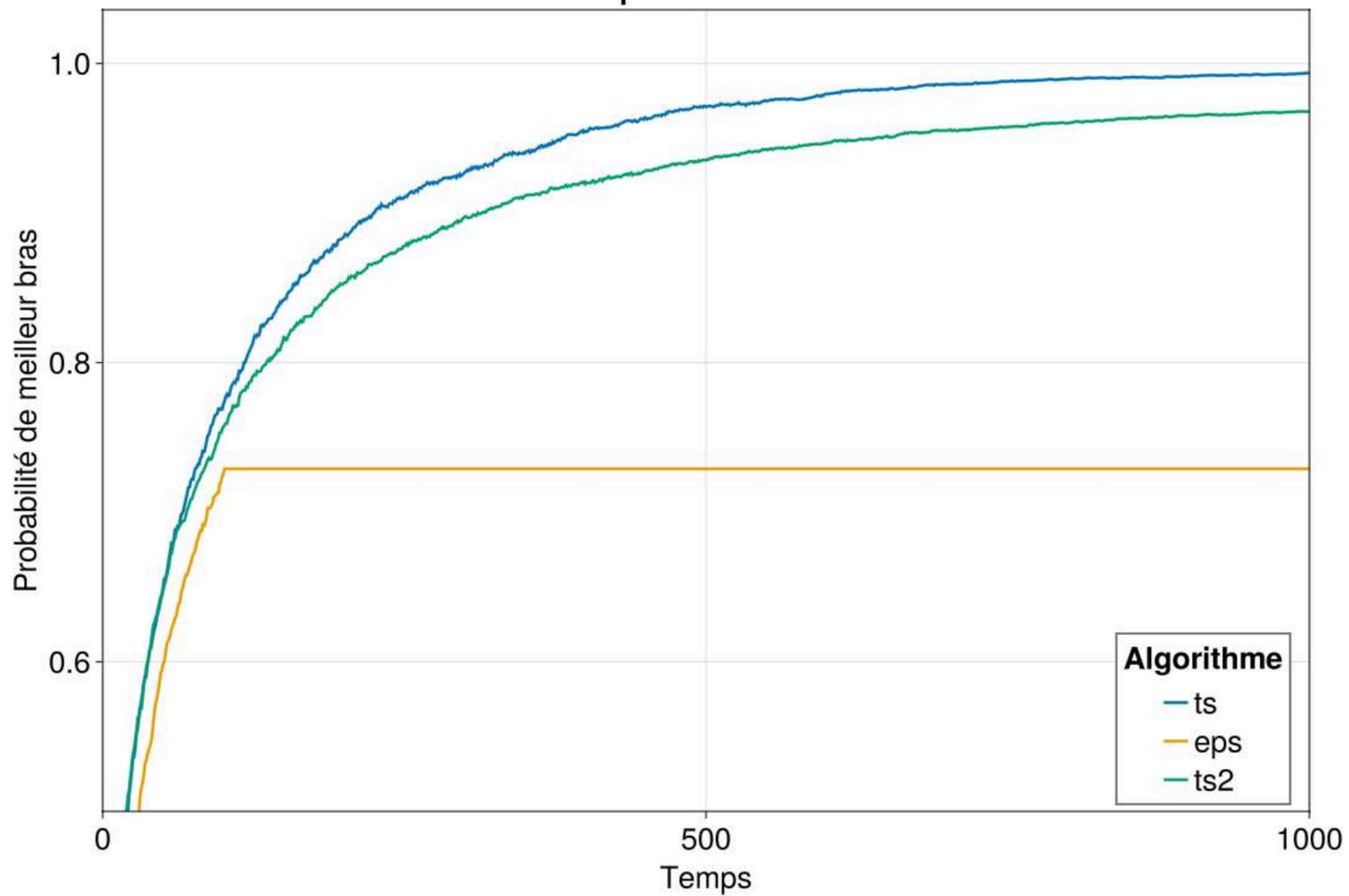
Évolution des succès



# Autres métriques

- Typiquement : minimiser le *regret* (ou maximiser la proportion de succès)
- Mais on peut vouloir maximiser la probabilité d'avoir trouvé le *meilleur bras*

Évolution de la probabilité du meilleur bras



# Infos diverses

## Autres algos

- Upper Confidence Bound (UCB) - Auer, Cesa-Bianchi, and Fischer ([2002](#))
- Gittins ([1979](#))
- ...

## Variantes

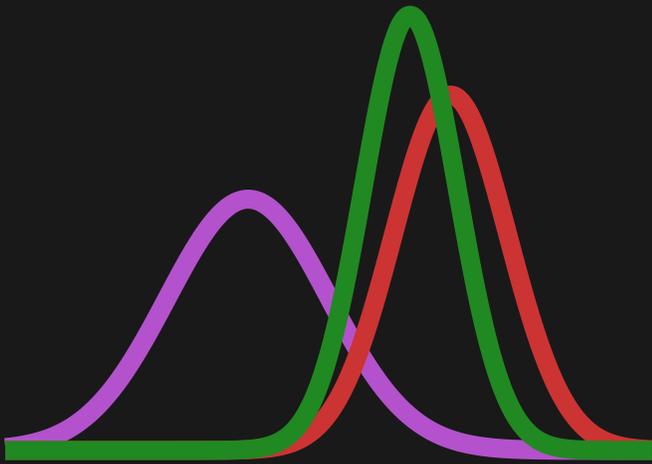
- Bandit contextuel
- Bandit non-stationnaire
- Bandit adversarial

# Limites

- Latence d'obtention de données
- Malédiction des généralités

# Pause questions 1

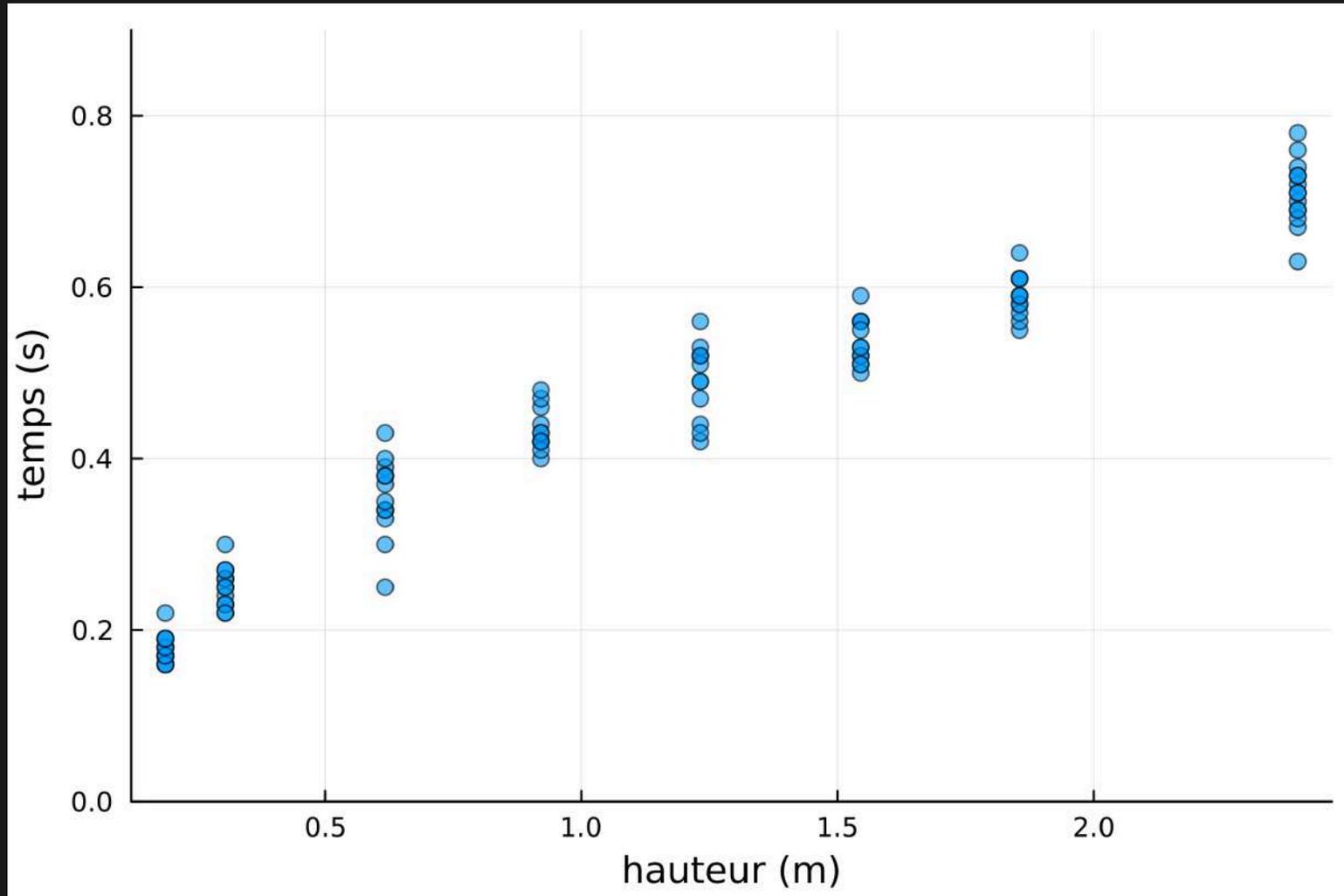
# Inférence bayésienne et programmation probabiliste



# La gravité de la situation

Une histoire de procrastination...

# Données



# Turing.jl

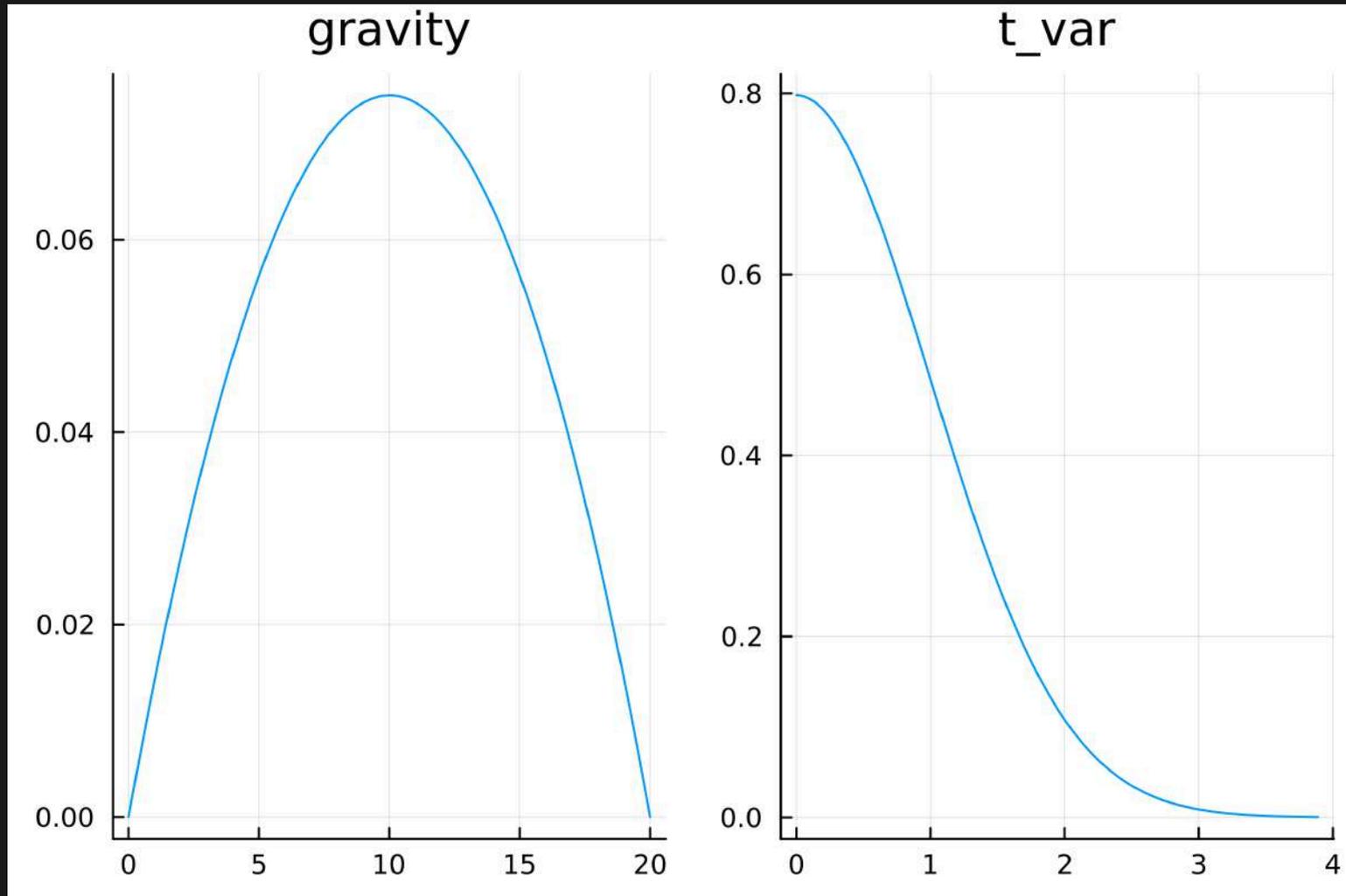
Fjelde et al. (2025)

- **inférence bayésienne** : partir d'une croyance *a priori* (prior) sur une hypothèse ou un paramètre qui nous intéresse, et la mettre à jour avec les données observées.
- **programmation probabiliste** : découpler la spécification du modèle et les calculs d'inférence. Pas besoin d'être très expert en programmation ou en statistiques computationnelles.
- implémenté dans le langage **Julia**
- extrêmement versatile

# Exemple de modèle

```
1 @model function acceleration_model(t, d)
2   gravity ~ 20 * Beta(2, 2) + 0
3   t_bias ~ Normal(0, 0.20)
4   t_var ~ truncated(Normal(0, 2), lower = 0)
5
6   time_from_distance(dist) = sqrt(2dist / gravity)
7
8   for i in eachindex(t)
9     t[i] ~ Normal(time_from_distance(d[i]) + t_bias, sqrt(t_var))
10  end
11 end
```

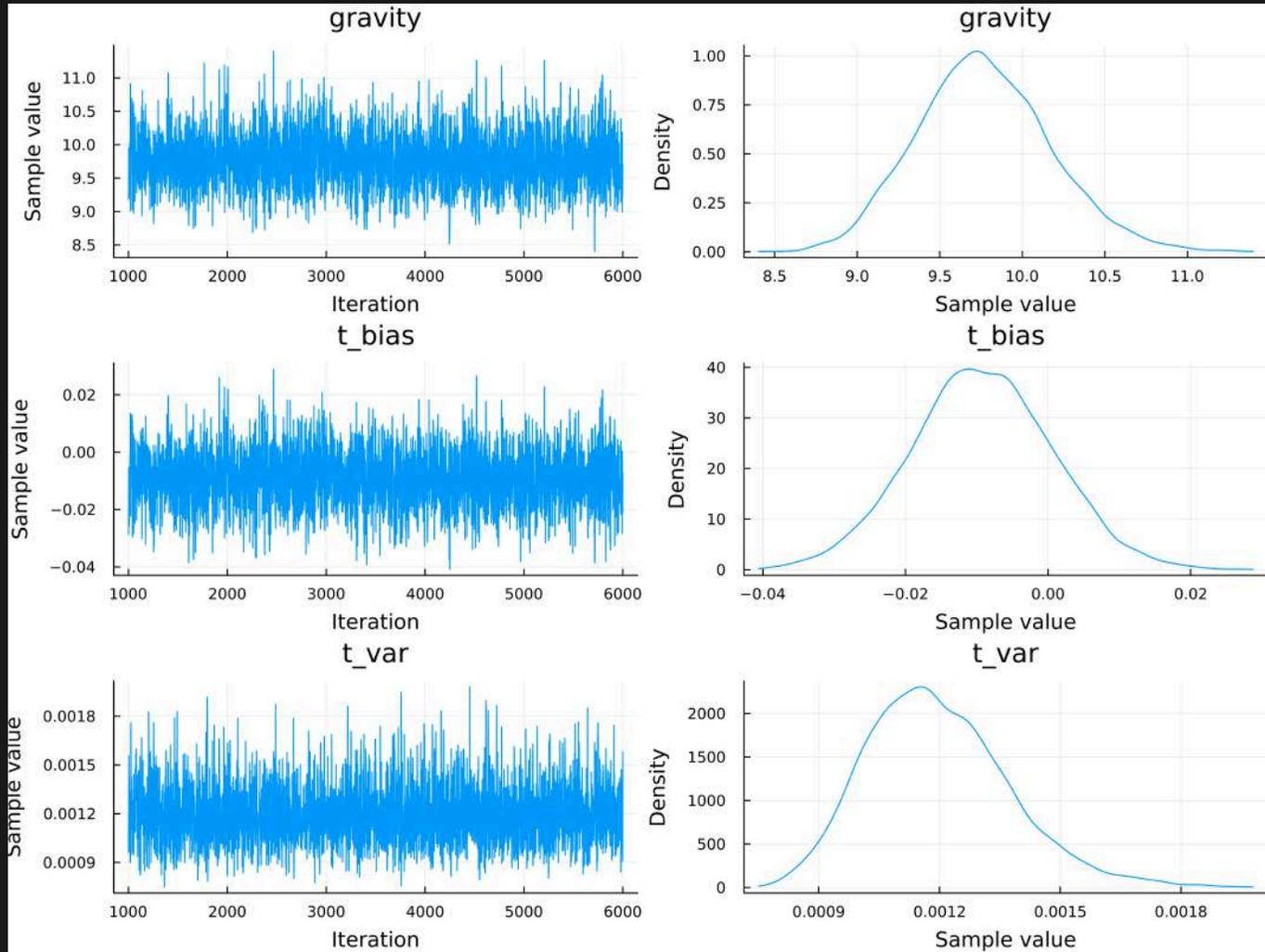
# Nos priors



# Exemple de modèle

```
1 @model function acceleration_model(t, d)
2   gravity ~ 20 * Beta(2, 2) + 0
3   t_bias ~ Normal(0, 0.20)
4   t_var ~ truncated(Normal(0, 2), lower = 0)
5
6   time_from_distance(dist) = sqrt(2dist / gravity)
7
8   for i in eachindex(t)
9     t[i] ~ Normal(time_from_distance(d[i]) + t_bias, sqrt(t_var))
10  end
11 end
```

# Résultats (visuels)



# Résultats (résumés stats)

## Summary Statistics

<b>parameters</b>	<b>mean</b>	<b>std</b>	<b>mcse</b>	<b>ess_bulk</b>	<b>ess_tail</b>	<b>rhat</b>	<b>ess_per_sec</b>
Symbol	Float64	Float64	Float64	Float64	Float64	Float64	Float64
gravity	9.7676	0.4039	0.0090	2016.7659	2239.0942	1.0010	554.0566
t_bias	-0.0093	0.0099	0.0002	1861.2921	2118.0926	1.0013	511.3440
t_var	0.0012	0.0002	0.0000	2602.2371	2430.2199	0.9999	714.9003

## Quantiles

<b>parameters</b>	<b>2.5%</b>	<b>25.0%</b>	<b>50.0%</b>	<b>75.0%</b>	<b>97.5%</b>
Symbol	Float64	Float64	Float64	Float64	Float64
gravity	9.0337	9.4926	9.7502	10.0299	10.6154
t_bias	-0.0286	-0.0158	-0.0094	-0.0027	0.0102
t_var	0.0009	0.0011	0.0012	0.0013	0.0016

# Pause questions 2

# Thompson Sampling + Turing

# Contexte

Kim a mal au dos.

# Protocole

- Participante experte motivée
- Protocole co-construit au fil de discussions, essais, simulations
- Questionnaire quotidien, co-construit
- Allocation dynamique et quotidienne des variables expérimentales
- But général : **minimiser la douleur**

# Variables expérimentales

```
1  const CONDS_MARCHE = ["0", "20", "20x2", "40"]
2  const CONDS_EAU = ["Libre", "1,5 L", "2 L", "2,5 L", "3 L"]
3  const CONDS_POSTURE = [
4      "Pas de travail",
5      "Assise toute la journée", # 1
6      "Assise + 10 minutes debout / h", # 2
7      "Assise + 10 minutes ballon / h", # 3
8      "1h assise, 1h ballon", # 4
9      "45 minutes assise/debout/assise/ballon", # 5
10 ]
11 const CONDS_SPORT = ["Pas de sport", "Doux", "Moyen", "Intense"]
12 const CONDS_MEDITATION = ["non", "15 min"]
13 const CONDS_ETIREMENTS = [0, 1]
14 const CONDS_ECHAUFFEMENTS = [0, 1]
```

# Variables observationnelles

```
1 const C_SOMMEIL = 33
2 const C_STRESS = 34
3 const C_FATIGUE = 35
4 const C_PERTU_FAM_NEG = 36
5 const C_INTER_SOCIALES = 37
6 const C_TIPPING_POINT = 38
7 const C_PAIN_EVENT = 39
```

# Modèle

- Beaucoup de variables mais relativement simple
- Pas d'interactions
- Variables avec effets différés jusqu'à 4 jours

# Modèle

```
1 @model function kimi_backpain_complex(data::Array{Float64, 2})
2   baseline ~ Normal(55, 20)
3   unexplained_var ~ InverseGamma(0.1, 0.4)
4
5   # ...
6
7   es_posture_1 ~ Normal(2.5, 20) # 12
8   es_posture_2 ~ Normal(-5, 15) # 13
9
10  # ...
11
12  es_sommeil ~ truncated(Normal(-0.5, 3); upper=0.)
13
14  # ...
15
```

# L'importance des simulations

Pour un modèle relativement complexe, faire des simulations en amont du déploiement réel est essentiel.

L'algorithme va prendre des décisions automatisées, avec de **vrais effets** dans le monde.

- Décisions bien prises → effets très positifs
- Bugs ou erreurs dans l'algo ou le modèle → potentielle catastrophe
- Avec une bonne simulation, on peut vérifier que le système est bien capable de trouver les bonnes conditions

# Résultats

## Sur l'algo

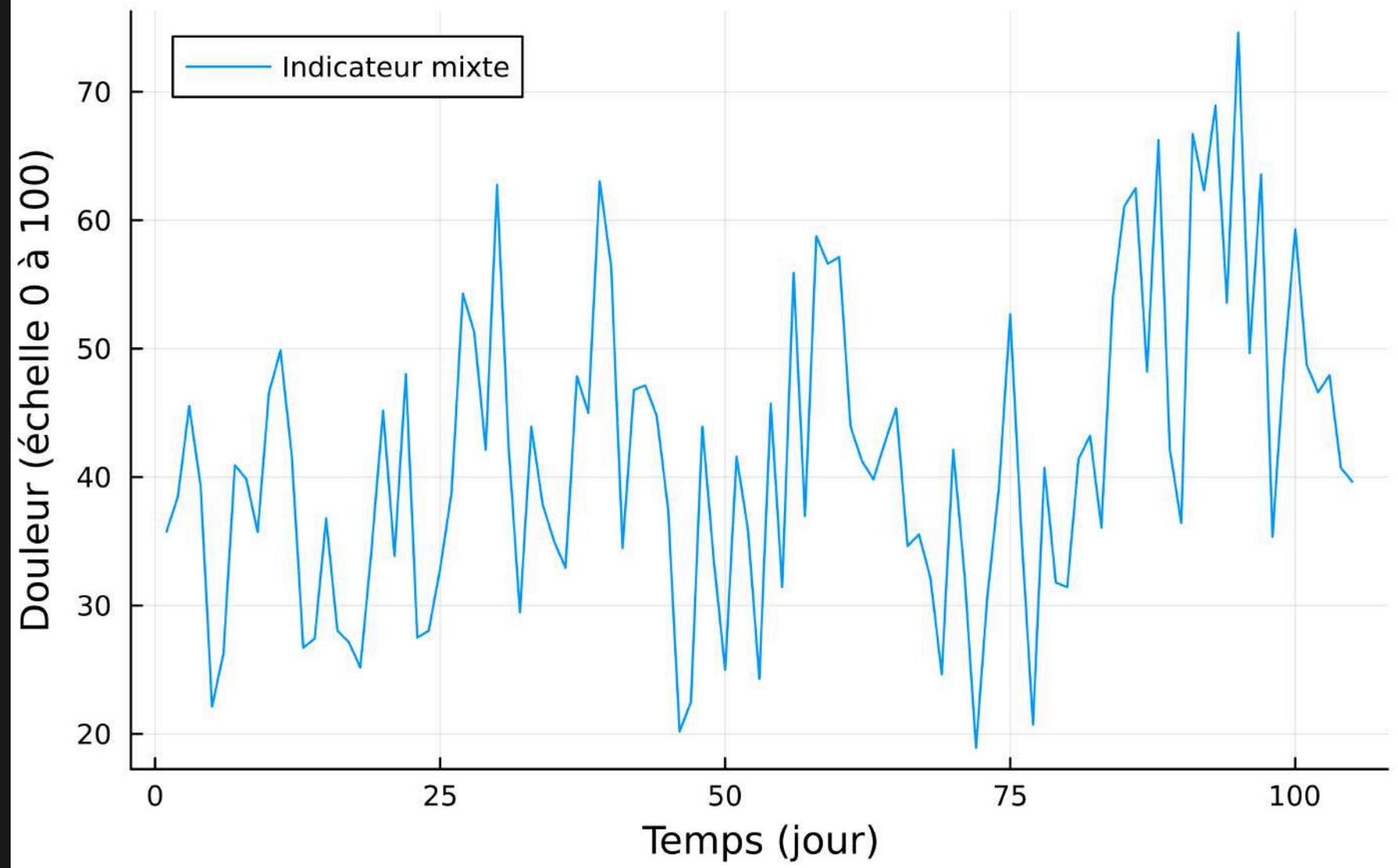
- A très vite détecté et évité les mauvaises conditions
  - sport intense
  - ballon
  - 40 min de marche
- Et favorisé les conditions aidantes
  - échauffements le matin
  - posture assise 50 min + debout 10 min
- Globalement très positif

# Résultats

## Sur l'objectif

- Malheureusement, il semble qu'aucune des combinaisons de conditions n'ait réduit massivement la douleur par rapport à une gestion intuitive
- On a appris beaucoup
- Grande variabilité
- Abandon après 105 jours d'assiduité et de relevés de données
- Solution trouvée plus tard grâce à un kiné

# Évolution de la douleur



# Est-ce limité à un seul participant ?

Pas du tout.

# Simulation multi-participants longitudinale

Demo live !

# Exemples d'applications

Vous :

*faites de la recherche sur le vécu étudiant*

Et vous voulez :

*améliorer le bien-être des étudiants*

# Exemples d'applications

Vous :

*êtes Google*

Et vous voulez :

*maximiser la probabilité que les gens cliquent sur vos pubs*

# Exemples d'applications

Vous :

*faites de l'Experience Sampling*

Et vous voulez :

*maximiser le bonheur*

# Exemples d'applications

Vous :

*développez des algorithmes pour des quiz adaptatifs (AMUQuiz)*

Et vous voulez :

*trouver les meilleures paramètres pour favoriser l'apprentissage*

# Exemples d'applications

Vous :

*faites de la recherche sur la co-parentalité*

Et vous voulez :

*trouver des interventions qui améliorent la distribution de la charge mentale*

# Exemples d'applications

Vous :

*faites de la recherche sur les dys*

Et vous voulez :

*trouver des interventions utiles malgré la grande variabilité*

# Exemples d'applications

Vous :

*faites de la recherche sur la dépression*

Et vous voulez :

*la réduire*

# Utilisation pour les essais cliniques en médecine

Villar, Bowden, and Wason (2015)

- ça vient petit à petit
- quelques défis
- communauté habituée aux approches classiques avec une certaine aversion au risque

# Ok je suis convaincu, mais par où commencer ?

Discutons-en

- ça change un peu l'approche des questions de recherche :  
PPDAC en boucle "courte"
  - PPDAC : *Problem Protocol Data Analysis Conclusions*
- simulations

# Mais où sont les p-values ?

DID THE SUN JUST EXPLODE?  
(IT'S NIGHT, SO WE'RE NOT SURE.)

THIS NEUTRINO DETECTOR MEASURES  
WHETHER THE SUN HAS GONE NOVA.

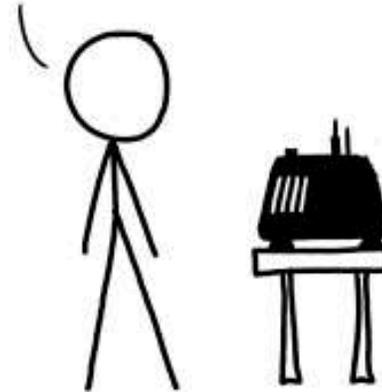
THEN, IT ROLLS TWO DICE. IF THEY  
BOTH COME UP SIX, IT LIES TO US.  
OTHERWISE, IT TELLS THE TRUTH.

LET'S TRY.  
DETECTOR! HAS THE  
SUN GONE NOVA?



FREQUENTIST STATISTICIAN:

THE PROBABILITY OF THIS RESULT  
HAPPENING BY CHANCE IS  $\frac{1}{36} = 0.027$ .  
SINCE  $p < 0.05$ , I CONCLUDE  
THAT THE SUN HAS EXPLODED.



BAYESIAN STATISTICIAN:

BET YOU \$50  
IT HASN'T.



# Pause questions 3

# Références

- Auer, Peter, Nicolo Cesa-Bianchi, and Paul Fischer. 2002. “Finite-Time Analysis of the Multiarmed Bandit Problem.” *Machine Learning* 47: 235–56.
- Fjelde, Tor Erlend, Kai Xu, David Widmann, Mohamed Tarek, Cameron Pfiffer, Martin Trapp, Seth D Axen, et al. 2025. “Turing. JI: A General-Purpose Probabilistic Programming Language.” *ACM Transactions on Probabilistic Machine Learning*.
- Gittins, John C. 1979. “Bandit Processes and Dynamic Allocation Indices.” *Journal of the Royal Statistical Society Series B: Statistical Methodology* 41 (2): 148–64.
- Thompson, William R. 1933. “On the Likelihood That One Unknown Probability Exceeds Another in View of the Evidence of Two Samples.” *Biometrika* 25 (3/4): 285–94.  
<http://www.jstor.org/stable/2332286>.
- Villar, Sofía S, Jack Bowden, and James Wason. 2015. “Multi-Armed Bandit Models for the Optimal Design of Clinical Trials: Benefits and Challenges.” *Statistical Science: A Review Journal of the Institute of Mathematical Statistics* 30 (2): 199.