

Le Paradoxe de Thib Dossier mathématique complet

Auteur : Pascal Thibodeau

Affiliation : Indépendant, Sorel-Tracy, Québec, Canada

Date : Septembre 2025

Site web : <https://kshiotsn.gensparkspace.com>

Objectif. Ce document expose, de manière pédagogique mais rigoureuse, l'intégralité des composantes mathématiques du *Paradoxe de Thib* : inférence bayésienne, utilité espérée, valeur d'information (EVSI), règles de seuil sous coûts/bénéfices asymétriques, score opérationnel, agrégation multi-signaux, et calibration. Les formules sont données avec notations claires, dérivations, sensibilités et exemples numériques.

Symbole	Définition (unité)
H_A / H_N	Hypothèse d'intérêt (artificiel / nouvelle physique) vs hypothèse naturelle
D	Données/observations disponibles
$K = p(D H_A)/p(D H_N)$	Facteur de Bayes (rapport de vraisemblance)
$O_A = p(H_A)/p(H_N)$	Odds a priori
$O_A = K \cdot O_N$	Odds a posteriori
$\pi = p(H_A D) = O_A/(1+O_A)$	Probabilité a posteriori de H_A
C	Coût marginal de l'investigation
B, b	Bénéfice attendu si H_A vrai (B) ; si H_N vrai ($b \geq 0$)
EVSI	Valeur d'information attendue (pré-investigation)
S	Score opérationnel $A \cdot I \cdot R \cdot U / (C \cdot P)$
A, I, R, U	Intensité d'anomalie, information, risque/réversibilité, urgence
P	Pénalité a priori $\sim 1/O_N$ (ou fonction croissante en $1/O_N$)

1. Noyau bayésien : facteur de Bayes, odds et postérieur

On compare deux hypothèses H_A et H_N . À partir de données D , le *facteur de Bayes* est $K = p(D|H_A) / p(D|H_N)$. Les *odds* a priori sont $O_A = p(H_A)/p(H_N)$. Le *théorème de Bayes* en écriture d'odds donne $O_A = K \cdot O_N$. La probabilité a posteriori de H_A est alors $\pi = O_A/(1+O_A)$. En pratique, on travaille souvent en log-espace : $\log K$, $\log O_A$ et $\log O_N = \log K + \log O_A$.

2. Règle de décision par utilité espérée

On note C le coût marginal d'une étape d'investigation (heures télescope, analyses), B le bénéfice si H_A est vrai (découverte à haute valeur), et b le bénéfice si H_N est vrai (clarification à valeur scientifique).

L'*accroissement d'utilité espérée* d'une investigation est $\Delta EU \approx \pi \cdot B + (1-\pi) \cdot b - C$. La **règle optimale** est : *investiguer si* $\Delta EU > 0$. En isolant π , on obtient un *seuil* $\pi > \pi^*$ avec $\pi^* = (C - b)/(B - b)$ (valide si $B > b$ et $C > b$).

En réécrivant en fonction du *facteur de Bayes*, la condition se formule $K > \tau$ avec $\tau = [(C - b)/(B - b)] \cdot (1/O_A)$. Cas conservateur (quand b est négligeable) : $\tau \approx (C/B) \cdot (1/O_A)$. En log : $\log K > \log \tau$.

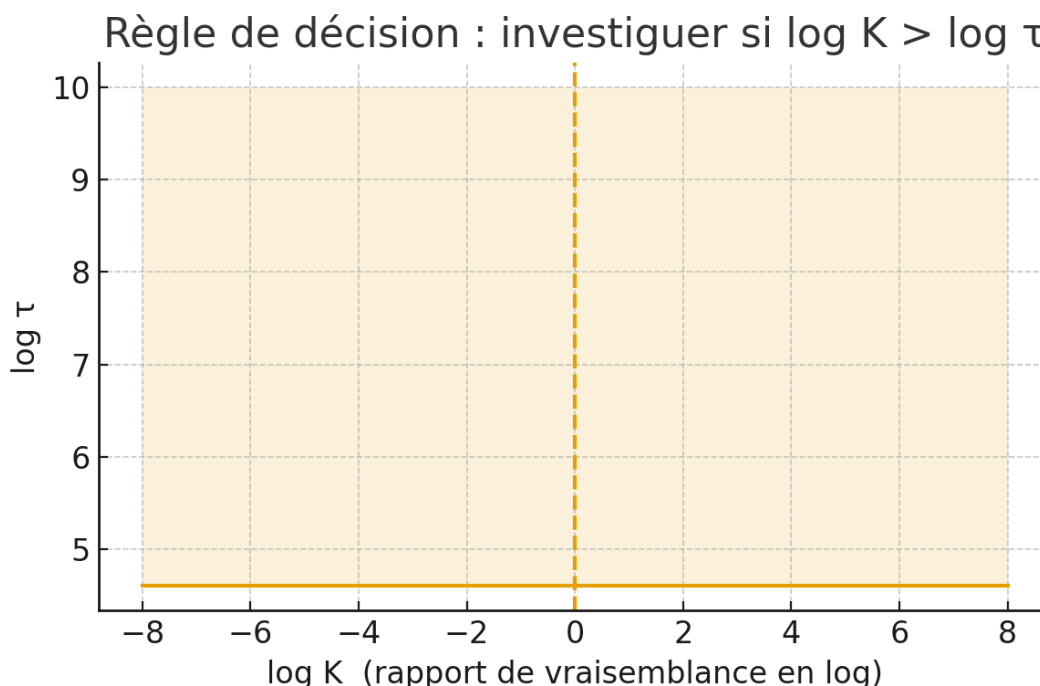


Figure 1. Région d'action : investiguer lorsque $\log K > \log \tau$.

3. Valeur d'information attendue (EVSI)

Avant d'engager une mesure M (p.ex. un créneau spectroscopique), on estime l'*EVSI* : le gain d'utilité espérée dû à l'actualisation de π après M , moins le coût de M . La règle de décision *séquentielle* est : *acquérir M si* $EVSI(M) > \text{Coût}(M)$. Dans les problèmes Σ ■étapes, on applique un raisonnement itératif « one■step look■ahead » : on choisit la prochaine action à *EVSI net* maximal, et on ré■évalue après observation. Cela évite la myopie et formalise l'urgence lorsque les fenêtres d'observation se ferment.

4. Score opérationnel S et équivalence bayésienne

Pour l'usage quotidien, on pose $S = (A \cdot I \cdot R \cdot U) / (C \cdot P)$ et l'on enquête si $S > 1$. Ici, A (anomalie) est un *proxy* croissant de K (par ex. $A = \exp(\alpha \cdot \text{LLR})$ avec $\text{LLR} = \log K$), I approxime l'*EVSI* de la prochaine mesure clé, R encode la réversibilité/risque (≥ 1 si réversible), U capture l'urgence (fenêtre courte $\Rightarrow U > 1$), C est le coût marginal, et P est une pénalité issue des *a priori* (p.ex. $P \propto 1/0$ ■). Sous un calibrage monotone, la règle $S > 1$ est *équivalente* à $K > \tau$.

5. Agrégation multi■signaux

Si des signaux (indépendants conditionnellement) D ■, ..., D_m sont disponibles, les facteurs de Bayes s'agrègent multiplicativement : $K_{\text{total}} = \prod K_i$ (ou $\log K_{\text{total}} = \sum \log K_i$). En cas de dépendances, on pondère ou on modélise explicitement les corrélations (graphes bayésiens). Une *bonne pratique* consiste à limiter la redondance et à valider la robustesse par analyses de sensibilité.

6. Calibration et échelles

- *Calibration a priori* : fixer 0 ■ par expertise (plausibilité physique, fréquence attendue d'artefacts).
- *Calibration anomalie*■ $\rightarrow K$: définir une carte A ■ K (p.ex. $\log K = \alpha \cdot A - \beta$) via rétro■analyse de cas historiques.
- *Coûts/bénéfices* : exprimer B , b , C dans une même unité utilité■équivalente (p.ex. « heures■téléscope normalisées »).
- *Seuils explicites* : publier τ et sa décomposition pour transparence et audit scientifique.

7. Sensibilité et élasticités

En $\log$ ■espace, la condition est $\text{LLR} \equiv \log K > \log \tau = \log(C - b) - \log(B - b) - \log 0$ ■. Les élasticités montrent comment bouge la frontière décisionnelle : augmenter B ou 0 ■ $\downarrow$ abaisse le seuil,

augmenter $C \uparrow$ le relève. Un tableau de bord peut tracer $\log \tau$ versus ces paramètres pour des audits rapides.

8. Exemple chiffré (vérifié)

Hypothèses de travail : $O = 1.0e-04$, $B = 1.0e+06$, $b = 1.0e+03$, $C = 1.0e+04$ (unités arbitraires mais homogènes). Seuil général : $\tau = ((C-b)/(B-b)) \cdot (1/O) = 9.009e+01$, donc $\log \tau = 4.501$. Règle : investiguer si $\log K > 4.501$.

9. Implémentation pas à pas

- 1) Fixer O (atelier d'experts, bornes plausibles).
- 2) Définir l'unité d'utilité commune et chiffrer B , b , C .
- 3) Choisir la carte $A \cdot K$ et la calibrer (rétroanalyse).
- 4) Construire le tableau de bord $S = (A \cdot I \cdot R \cdot U) / (C \cdot P)$ et publier τ .
- 5) En exploitation : agréger les signaux ($\sum \log K_i$), évaluer EVSI de la *prochaine* mesure, décider si $EVSI > \text{Coût}$.
- 6) Journaliser toutes les décisions et réajuster les calibrations trimestriellement.

10. Limites et garde-fous

- Risque de surinvestigation si B est surestimé ; imposer des bornes et audits externes.
- Cartes $A \cdot K$ non universelles ; documenter le contexte d'application.
- Dépendances entre signaux ; préférer des modèles graphiques quand c'est critique.
- Éthique : publier la décomposition de τ et les journaux de décision pour relecture par les pairs.

Références & lectures utiles

- Notes internes « Paradoxe de Thib » (2025).
- Litterature standard sur décision bayésienne, utilité espérée et EVSI.
- Études de cas ISO (1I/Oumuamua, 2I/Borisov, 3I/ATLAS) pour calibration a posteriori.