

Tirage d'une pièce de monnaie et néguentropie de composition

Hugues Genvrin

02/09/2026

Résumé du chapitre. On prend le tirage d'une pièce comme cas d'école pour la néguentropie. On établit d'abord un résultat qui a l'allure d'un paradoxe : la pièce *équilibrée* n'a *aucune* néguentropie — elle est l'*étalon* du désordre maximal. On lève le paradoxe en montrant que la néguentropie n'apparaît qu'avec un *écart* : le *biais* ($p \neq 1/2$) ou la *mémoire* (tirages corrélés). On calcule les deux, on montre qu'ils produisent *numériquement la même* néguentropie, et l'on interprète cette néguentropie comme un *écart pur* (une divergence de Kullback–Leibler à la référence uniforme ou indépendante). On relie enfin le tout à la complexité de Kolmogorov : le support des suites porte n bits, mais c'est la *compressibilité* d'une suite particulière qui mesure sa structure. Chaque résultat est démontré, chiffré, et accompagné du code.

Sommaire

1.1	Le tirage d'une pièce : ce qu'un tirage apprend	2
1.2	Néguentropie de composition et équiprobabilité	2
1.3	Néguentropie de composition comme écart pur	3
1.3.1	Écart par le biais : la pièce truquée	3
1.3.2	Écart par la mémoire : la pièce collante	3
1.3.3	L'écart pur : biais et mémoire sont le même écart	4
1.4	Le tirage d'une pièce et l'information de Kolmogorov	4
1.5	Synthèse	5

1.1 Le tirage d'une pièce : ce qu'un tirage apprend

Une pièce est une source de Bernoulli : deux issues, pile (P) et face (F), de probabilités p et $1 - p$. L'information qu'un tirage *apprend* est son entropie de Shannon

$$\mathcal{H}(p) = -p \log_2 p - (1 - p) \log_2 (1 - p) \text{ bits.}$$

Proposition 1.1 (Ce qu'un tirage apprend). $\mathcal{H}(p)$ est maximale et vaut 1 bit pour $p = 1/2$, décroît symétriquement de part et d'autre, et tend vers 0 quand $p \rightarrow 0$ ou $p \rightarrow 1$.

Proof. $\mathcal{H}'(p) = -\log_2 p + \log_2 (1 - p) = \log_2 \frac{1-p}{p}$, qui s'annule en $p = 1/2$, est positif pour $p < 1/2$ et négatif pour $p > 1/2$: maximum unique en $p = 1/2$, où $\mathcal{H}(1/2) = -2 \cdot \frac{1}{2} \log_2 \frac{1}{2} = 1$. Aux bords, $\lim_{p \rightarrow 0^+} -p \log_2 p = 0$ (et de même en 1), donc $\mathcal{H} \rightarrow 0$. $\square$

Remarque 1.2 (Le socle, proprement posé). Un tirage équilibré apprend exactement 1 bit ; un tirage biaisé apprend *moins* (on « sait déjà » un peu le résultat). C'est la brique de base : la suite du chapitre mesure tout *écart* à ce bit maximal.

1.2 Néguentropie de composition et équiprobabilité

On pose d'abord le résultat sous forme de *question*, car il contredit une intuition tenace (« pile ou face, c'est le hasard, donc c'est structuré »).

Proposition 1.3 (La pièce équilibrée est l'étalon : $\mathcal{N} = 0$). La néguentropie (de biais comme de composition) d'une pièce équilibrée ($p = 1/2$, tirages indépendants) est nulle :

$$\mathcal{N} = \mathcal{H}_{\max} - \mathcal{H}_{\text{réel}} = \log_2 2 - 1 = 0.$$

Proof. Pour deux issues, $\mathcal{H}_{\max} = \log_2 2 = 1$. À $p = 1/2$, $\mathcal{H}_{\text{réel}} = \mathcal{H}(1/2) = 1$ (proposition 1.1). Donc $\mathcal{N} = 1 - 1 = 0$. Pour la composition, des tirages *indépendants* ont une loi jointe égale au produit des marginales, donc corrélation totale nulle. Les deux néguentropies s'annulent. $\square$

Remarque 1.4 (Résolution du paradoxe apparent). La pièce équilibrée n'est pas un exemple de néguentropie : c'est le cas de *désordre maximal*, le *sommet* d'entropie contre lequel toute néguentropie se mesure. Le hasard équiprobable est le *zéro de structure*, non un maximum de structure. C'est exactement pourquoi l'ancienne intuition « un tirage de pièce définit des sous-segments pile/face, d'où une néguentropie » est fausse : elle choisit le pire cas pour la thèse, celui où $\mathcal{N} = 0$. La pièce équilibrée *définit* le zéro, elle ne l'illustre pas comme un écart.

Remarque 1.5 (Ce qu'il faut pour une néguentropie non nulle). Il faut un *écart* à cet étalon. Deux écarts possibles, étudiés ci-dessous : le *biais* (§1.3, la pièce truquée $p \neq 1/2$) et la *mémoire* (§1.3, les tirages corrélés). Ce sont les deux seules sources de structure d'une source binaire.

1.3 Néguentropie de composition comme écart pur

On révèle maintenant la néguentropie par les deux écarts, on les chiffre, et l'on montre qu'ils sont *le même écart* mesuré de deux façons.

1.3.1 Écart par le biais : la pièce truquée

Proposition 1.6 (Néguentropie de biais de la pièce). *Pour une pièce de biais p , la néguentropie de biais est*

$$\mathcal{N}_{\text{biais}}(p) = 1 - \mathcal{H}(p) = 1 + p \log_2 p + (1 - p) \log_2 (1 - p) = D_{\text{KL}}((p, 1-p) \parallel (\frac{1}{2}, \frac{1}{2})) \geq 0,$$

nulle ssi $p = 1/2$, croissante vers 1 quand $p \rightarrow 0$ ou 1.

Proof. $\mathcal{N}_{\text{biais}} = \mathcal{H}_{\text{max}} - \mathcal{H}(p) = 1 - \mathcal{H}(p)$. La forme D_{KL} vient de la factorisation générale (chapitre précédent) : $1 - \mathcal{H}(p) = \sum_k p_k \log_2 \frac{p_k}{1/2}$ avec $(p_k) = (p, 1-p)$. Positivité et cas d'égalité par Gibbs. $\square$

p	$\mathcal{H}(p)$ (bits)	$\mathcal{N}_{\text{biais}} = 1 - \mathcal{H}(p)$
0,50	1,0000	0,0000
0,60	0,9710	0,0290
0,70	0,8813	0,1187
0,80	0,7219	0,2781
0,90	0,4690	0,5310
0,99	0,0808	0,9192

1.3.2 Écart par la mémoire : la pièce collante

On considère une pièce *sans biais* ($p = 1/2$) mais dont les tirages sont *corrélés* : une chaîne de Markov à deux états où $\Pr(P | P) = \Pr(F | F) = a > 1/2$ (« un pile appelle un pile »). Par symétrie, la loi stationnaire est $(\frac{1}{2}, \frac{1}{2})$: *aucun biais marginal*.

Proposition 1.7 (Néguentropie de composition de la pièce collante). *L'entropie marginale par tirage est 1 bit ; l'entropie conditionnelle (sachant le tirage précédent) est*

$$\mathcal{H}_{\text{cond}} = \mathcal{H}(a) = -a \log_2 a - (1 - a) \log_2 (1 - a),$$

et la néguentropie de composition vaut

$$\mathcal{N}_{\text{compo}} = \mathcal{H}_{\text{marg}} - \mathcal{H}_{\text{cond}} = 1 - \mathcal{H}(a) \geq 0,$$

strictement positive dès que $a \neq 1/2$, sans aucun biais.

Proof. La stationnaire π vérifie $\pi_P = a\pi_P + (1-a)\pi_F$ et $\pi_P + \pi_F = 1$; par symétrie $\pi_P = \pi_F = 1/2$, d'où $\mathcal{H}_{\text{marg}} = \mathcal{H}(1/2) = 1$. L'entropie conditionnelle d'une chaîne de Markov stationnaire est $\mathcal{H}_{\text{cond}} = \sum_i \pi_i \mathcal{H}(\text{ligne } i) = \frac{1}{2}\mathcal{H}(a, 1-a) + \frac{1}{2}\mathcal{H}(1-a, a) = \mathcal{H}(a)$. La néguentropie de composition est l'écart marginal moins conditionnel, $1 - \mathcal{H}(a)$. $\square$

$a = \Pr(P P)$	$\mathcal{H}_{\text{marg}}$	$\mathcal{H}_{\text{cond}} = \mathcal{H}(a)$	$\mathcal{N}_{\text{compo}} = 1 - \mathcal{H}(a)$
0,60	1,0000	0,9710	0,0290
0,70	1,0000	0,8813	0,1187
0,80	1,0000	0,7219	0,2781
0,90	1,0000	0,4690	0,5310

1.3.3 L'écart pur : biais et mémoire sont le même écart

Théorème 1.8 (Identité numérique biais/mémoire). *Les deux tableaux coïncident terme à terme : pour tout $t \in (1/2, 1)$,*

$$\mathcal{N}_{\text{biais}}(p=t) = \mathcal{N}_{\text{compo}}(a=t) = 1 - \mathcal{H}(t).$$

Biais et mémoire produisent exactement la même négentropie ; ce sont deux incarnations d'un unique écart pur $1 - \mathcal{H}(t) = D_{\text{KL}}((t, 1-t) \| (\frac{1}{2}, \frac{1}{2}))$.

Proof. Les propositions 1.6 et 1.7 donnent la même expression $1 - \mathcal{H}(t)$, l'une avec $t = p$ (biais marginal), l'autre avec $t = a$ (probabilité conditionnelle). L'égalité est donc une identité de fonctions. $\square$

Remarque 1.9 (La négentropie est un écart, pas une substance). Ce théorème est le cœur du chapitre. La négentropie de composition n'est pas une « quantité de structure » stockée quelque part : c'est un *écart* à une référence (l'uniforme pour le biais, l'indépendance pour la mémoire). Que les deux écarts coïncident numériquement montre qu'il s'agit du *même* objet — une divergence de Kullback–Leibler — lu selon deux axes (le biais est l'axe « horizontal à une position » ; la mémoire est l'axe « corrélation entre positions » du chapitre précédent). L'« écart pur » $1 - \mathcal{H}(t)$ est l'invariant commun.

Le code des deux tableaux et de l'identité :

```
import math
log2 = math.log2
def H(ps): return -sum(p*log2(p) for p in ps if p > 0)
Hb = lambda t: H([t, 1-t]) # entropie binaire

print(" t Nbiais(p=t) Ncompo(a=t) 1-H(t)")
for t in [0.5, 0.6, 0.7, 0.8, 0.9, 0.99]:
    Nbiais = 1 - Hb(t) # piece truquee : ecart a l'uniforme
    # piece collante Markov P(P|P)=P(F|F)=t, stationnaire 1/2
    Hcond = 0.5*H([t, 1-t]) + 0.5*H([1-t, t]) # = H(t)
    Ncompo = 1 - Hcond
    print("%.2f %.4f %.4f %.4f" % (t, Nbiais, Ncompo, 1 - Hb(t)))
# t=0.80 -> 0.2781, 0.2781, 0.2781 : les trois colonnes coïncident
```

1.4 Le tirage d'une pièce et l'information de Kolmogorov

On change de mesure : de l'entropie (propriété d'une *source*) à la complexité de Kolmogorov (propriété d'une *suite* individuelle).

Définition 1.10 (Support et structure). Pour n tirages, le *support* est l'ensemble des 2^n suites possibles ; son information de Hartley est $I_{\text{support}} = \log_2 2^n = n$ bits, *identique* pour toutes les suites. La *structure* d'une suite particulière s est sa *complexité de Kolmogorov* $K(s)$: la longueur du plus court programme qui la produit.

Proposition 1.11 (La structure distingue ce que le support égalise). *Le support ne distingue pas les suites (toutes à n bits). La complexité K , elle, les sépare : une suite régulière est compressible ($K \ll n$), une suite aléatoire est incompressible ($K \approx n$).*

Argument. La suite constante P^n est produite par « répéter P n fois », programme de longueur $O(\log n)$ (il suffit d'encoder n) : $K(P^n) = O(\log n) \ll n$. La suite périodique $(PF)^{n/2}$ est produite par « répéter PF $n/2$ fois », de longueur $O(\log n)$ également. Une suite *aléatoire* typique n'admet aucun programme plus court qu'elle-même : par comptage, il y a 2^n suites mais seulement $2^n - 1$ programmes de longueur $< n$, donc au moins une suite (en fait la plupart) vérifie $K(s) \geq n$. $\square$

Suite ($n = 16$)	I_{support}	compressible ?	K (ordre de grandeur)
$PPPPPPPPPPPPPPPP$	16 bits	oui (motif P)	$O(\log n)$ très petit
$PFPFPFPFPFPFPFP$	16 bits	oui (motif PF)	$O(\log n)$ petit
suite aléatoire	16 bits	non	≈ 16 maximal

Remarque 1.12 (Prudence sur la mesure de compressibilité). Un codage *run-length* naïf ne capture pas toute la structure : la suite périodique $PFPF \dots$ a 16 « runs » (aucune répétition immédiate) et paraît incompressible *pour ce codage*, alors qu'elle est très compressible *par motif* (K petit). La complexité de Kolmogorov est le *minimum sur tous les programmes*, pas sur un codage particulier ; elle est d'ailleurs non calculable en général. Les valeurs du tableau sont des ordres de grandeur, non des mesures exactes.

Remarque 1.13 (Lien avec le support/structure du programme). On retrouve la décomposition $I = I_{\text{support}} + I_{\text{struct}}$: le support (n bits) est commun ; la structure est ce que K mesure. La pièce *équilibrée et indépendante* produit des suites *typiquement incompressibles* ($K \approx n$) : c'est le sens algorithmique de « désordre maximal » — cohérent avec $\mathcal{N} = 0$ de la proposition 1.3. Biais ou mémoire *abaissent* K (les suites deviennent compressibles), et cette baisse est la contrepartie algorithmique de la néguentropie $\mathcal{N} = 1 - \mathcal{H}(t)$ du théorème 1.8.

Code (illustration des ordres de grandeur) :

```
import random, zlib
random.seed(1)
def bits(s): return s.encode()
seqs = {
    "constante P^16": "P"*16,
    "periodique (PF)^8": "PF"*8,
    "aleatoire": "".join(random.choice("PF") for _ in range(16)),
}
for name, s in seqs.items():
    comp = len(zlib.compress(bits(s))) # proxy grossier de K (borne SUP)
    print("%-20s I_support=16 bits taille_compressée=%d octets" % (name, comp))
# la constante se comprime le plus ; l'aleatoire résiste (proxy, non exact)
```

1.5 Synthèse

Remarque 1.14 (Le fil du chapitre). La pièce parcourt les trois régimes en une seule variable t :

- $t = 1/2$: pièce équilibrée indépendante — *étalon*, $\mathcal{N} = 0$, suites incompressibles ($K \approx n$) ;
- $t \neq 1/2$ par *biais* : $\mathcal{N}_{\text{biais}} = 1 - \mathcal{H}(t) > 0$, suites compressibles ;
- $t \neq 1/2$ par *mémoire* : $\mathcal{N}_{\text{compo}} = 1 - \mathcal{H}(t) > 0$, même valeur, suites compressibles.

Biais et mémoire sont le *même écart pur* $1 - \mathcal{H}(t)$; Kolmogorov en donne la lecture algorithmique (structure = compressibilité). La pièce équilibrée n'est pas un exemple de néguentropie : elle en est le *zéro*, et c'est ce zéro qui rend les écarts mesurables.

Remarque 1.15 (Limite épistémique). Toutes les quantités sont des capacités de distinction / compression (en bits). La néguentropie de la pièce ne dit rien d'une « énergie » ni d'un « ordre physique » : c'est un écart informationnel à une référence, mesuré et rien de plus. L'invocation de Schrödinger reste une *analogie de forme* ($\sum p \log p$), sans le $k_B T$ d'une entropie thermodynamique.