Math Methodology — Information Geometry on the Simplex with KL-Bounded Updates

> *Goal: maximize **volatility (Variance)** while staying within **RTP/Hit guardrails**. Treat the paytable as a point on the probability simplex. Use **mirror descent** / **natural gradient** with a **KL trust region** for small steps; when a step violates guardrails, apply a **KL (Bregman) projection**; validate with **Monte Carlo** and confidence intervals. The entire flow is config-driven and auditable.*

0) Setup

Payout multipliers: $a_1, \dots, a_n \in \mathbb{R}_{\geq 0}$.
Probability vector: $p = (p_1, \dots, p_n) \in \Delta_n := \{p_i > 0,\ \sum_i p_i = 1\}$ (open simplex).
Win index set: $W \subset \{1, \dots, n\}$ (e.g., $i \in W$ means prize i counts as a hit).

KPIs

$$\begin{aligned} \mathrm{RTP}(p) &= \sum_i a_i\, p_i, \\ \mathrm{Hit}(p) &= \sum_{i \in W} p_i, \\ \mu(p) &= \mathrm{RTP}(p), \\ \mathrm{Var}(p) &= \sum_i (a_i - \mu)^2 p_i \quad \Big(= \sum_i a_i^2 p_i - \mu^2 \Big). \end{aligned} \tag{1}$$

Guardrails (feasible set)

$$\mathcal{C} = \Big\{ p \in \Delta_n :\ \mathrm{RTP}(p) \in [R_-, R_+],\ \mathrm{Hit}(p) \in [H_-, H_+] \Big\}. \tag{2}$$

1) Objective and loss

We minimize a penalized loss:

$$L(p) = -\lambda_{\text{var}} \cdot \operatorname{Var}(p) \; + \; \rho_{\text{rtp}} \cdot \phi\big(\mathrm{RTP}(p)\big) \; + \; \rho_{\text{hit}} \cdot \psi\big(\mathrm{Hit}(p)\big), \tag{3}$$

where ϕ, ψ are **band penalties** (zero inside the band, positive outside, e.g. squared). Example: $\phi(x) = \big[\max\{0, x - R_+\}\big]^2 + \big[\max\{0, R_- - x\}\big]^2$; ψ analogous.

Gradients (w.r.t. p_i)

$$\frac{\partial \operatorname{Var}}{\partial p_i} = a_i^2 - 2\,\mu\, a_i, \qquad \frac{\partial \mathrm{RTP}}{\partial p_i} = a_i, \qquad \frac{\partial \mathrm{Hit}}{\partial p_i} = \mathbf{1}_{i \in W}. \tag{4}$$

Hence

$$\nabla_{p_i} L = -\lambda_{\text{var}}(a_i^2 - 2\mu a_i) \; + \; \rho_{\text{rtp}} \phi'(\mathrm{RTP})\, a_i \; + \; \rho_{\text{hit}} \psi'(\mathrm{Hit})\, \mathbf{1}_{i \in W}. \tag{5}$$

---

2) Statistical manifold and coordinates

- **Mixture coordinates**: $p_1, \ldots, p_{n-1}$ with $p_n = 1 - \sum_{i<n} p_i$.
- **Exponential (natural) coordinates**: $\theta_i = \log \frac{p_i}{p_n}$, mapping back to p via softmax.

KL divergence $D_{\mathrm{KL}}(q \| p) = \sum_i q_i \log \frac{q_i}{p_i}$ is the Bregman divergence of negative entropy. **Fisher information** induces the Riemannian metric on the statistical manifold; **small KL** steps approximate short geodesic motion.

---

3) Mirror descent / natural-gradient step (KL trust region)

Let $g = \nabla_p L(p)$. With negative entropy as the mirror map, a single step is

$$\tilde{p}_i \; \propto \; p_i \, \exp(-\eta\, g_i), \qquad \tilde{p} \leftarrow \frac{\tilde{p}}{\sum_j \tilde{p}_j}. \tag{6}$$

Trust region: accept only if

$$D_{\mathrm{KL}}(\tilde{p} \| p) \; \leq \; \epsilon. \tag{7}$$

Otherwise **backtrack** η (or shrink ϵ) and retry.

> *Intuition: this is multiplicative weights / exponentiated-gradient; on the statistical manifold the first-order approximation matches a natural-gradient step under the Fisher metric.*

---

4) KL (Bregman) projection back to guardrails

If $\tilde{p} \notin \mathcal{C}$, compute

$$p^{+} = \arg\min_{q \in \mathcal{C}} D_{\mathrm{KL}}(q \| \tilde{p}) \tag{8}$$

to return to feasibility. For **linear constraints** (normalization, RTP/Hit bands), KKT yields a closed-form **exponential tilting**:

$$q_i \;\propto\; \tilde{p}_i \,\exp\big(-\alpha\, a_i - \beta\, \mathbf{1}_{i \in W}\big), \tag{9}$$

with α, β chosen to meet active bounds; then renormalize.

> *If neither band is touched, no projection is needed; if only one band is violated, one of α, β is zero.*

---

5) Convergence and stopping

- Small objective change: $|L(p^{t+1}) - L(p^{t})| < \delta$
- Small KL step: $D_{\mathrm{KL}}(p^{t+1} \| p^{t}) < \delta_{\mathrm{KL}}$
- Or max iterations reached

---

6) Monte Carlo validation (with CIs)

Given final p^{*}, simulate N i.i.d. plays to obtain rewards $\{X_k\}$.

- $\widehat{\mathrm{RTP}} = \bar{X} = \frac{1}{N} \sum X_k$, 95% CI: $\bar{X} \pm 1.96\, \hat{\sigma}/\sqrt{N}$, with sample variance $\hat{\sigma}^2$.
- $\widehat{\mathrm{Hit}} = \frac{1}{N} \sum \mathbf{1}_{\{X_k > 0\}}$, 95% CI: $\hat{p} \pm 1.96 \sqrt{\hat{p}(1 - \hat{p})/N}$.

- $\widehat{\mathrm{Var}}$ compares to the analytic Var (for a CI use large-sample approximation or bootstrap).

Validation rule: analytic KPIs should lie within MC 95% CIs. Fix random seed for reproducibility.

---

7) Reproducibility and audit

- **Configs**: η, ϵ, max iters, RTP/Hit targets and bands, Var weight, MC N, seed.
- **Per-step logs**: KPIs and KL distance each step; export `metrics_history.csv` and final `summary.json`.
- **Excel mirror (optional)**: recompute RTP/Hit/Var and updates via formulas for non-engineering audit.

---

8) Limits and extensions

- **Locality**: small-KL steps are local; schedule step sizes and trust region to avoid oscillation.
- **Nonconvexity**: objectives with $-\mathrm{Var}$ are nonconvex; guarantees are first-order stationarity (local optima).
- **Extensions**:
    a. **Natural-gradient Newton**: Fisher preconditioning or proximal quadratic models for speed.
    b. **Multi-objective Pareto**: optimize Var and retention proxies jointly and plot the frontier.
    c. **Bayesian online updates**: use real play logs to update beliefs about the optimal p.
    d. **Production mechanics**: incorporate reels/lines combinatorics and approximations.

---

9) Algorithm skeleton (reference)

```
p ← p0                  # initial paytable distribution (in Δn)
for t in 1..T:
    g ← ∇_p L(p)      # use §1 gradient formulas
    η ← backtracking_until_KL_ok(η0, g, p, ε)
    p_tilde ← normalize(p * exp(-η * g))
    if not feasibility(p_tilde):
        p_next ← KL_project_to_bands(p_tilde, RTP_band, Hit_band)
# §4
    else:
        p_next ← p_tilde
    log_metrics(p_next, KL(p_next||p))
    if stop_rule_met(): break
    p ← p_next
# MC validation: simulate with p, output CIs and compare to
analytics
```

Appendix — Selected facts used (one-liners)

(F1) KL/Bregman I-projection (existence, uniqueness, shape)
For a convex nonempty linear constraint set $\mathcal{C}$, $\arg\min_{q\in\mathcal{C}} D_{\mathrm{KL}}(q\|\tilde{p})$ exists uniquely and has **exponential tilting** form
$q_i \propto \tilde{p}_i \exp(-\alpha a_i - \beta \mathbf{1}_{i\in W})$.

(F2) Bregman Pythagorean identity
If p^* is the Bregman projection of $\tilde{p}$ onto $\mathcal{C}$, then for any $q \in \mathcal{C}$:
$D_\phi(q\|\tilde{p}) = D_\phi(q\|p^*) + D_\phi(p^*\|\tilde{p})$.

(F3) Mirror descent = exponentiated-gradient; first-order equivalence to natural gradient
With negative entropy, $\tilde{p}_i \propto p_i \exp(-\eta \nabla_i L)$ matches multiplicative weights; on the statistical manifold the first-order step equals a Fisher-metric natural gradient step.

(F4) Fisher–KL second-order approximation (local geometry)
$2\, D_{\mathrm{KL}}(q\|p) \approx \|q-p\|^2_{F(p)}$, where $\|\cdot\|_{F(p)}$ is induced by the Fisher metric at p.

(F5) LLN/CLT for MC validation

Sample means converge to analytic KPIs (LLN) and admit $O(1/\sqrt{N})$ 95% CIs (CLT).

(F6) Trust region + backtracking ensures descent

For smooth L, a KL-prox step with backtracking yields acceptable descent and convergence to a first-order stationary point (local for nonconvex cases).

Scope / conditions

Work on the open simplex (use an ε-floor to avoid $p_i \to 0$ in practice).
Guardrails are linear bands (RTP/Hit); if adding nonlinear KPIs, (F1)(F2) still hold for convex sets, otherwise use numerical projections.
With nonconvex objectives we guarantee only local optima and Pareto frontiers.

One-liner.

On the statistical manifold, use **KL-bounded mirror/natural-gradient** steps; when violated, **KL-project** back to RTP/Hit bands; validate with **MC + CIs**; keep the flow **config-driven and auditable**.