Cumulant Update Equations for the β-Γ Bandit

1 Setup

Consider the two-armed bandit with binary outcomes. The agent maintains a log-likelihood ratio $\gamma \in \mathbb{Z}$ and selects actions according to:

$$p(a = +1 \mid \gamma) = \Phi(\beta\gamma), \tag{1}$$

where Φ is a sigmoid (e.g., logistic or probit). The outcome $y \in \{-1, +1\}$ is drawn from arm a with probability

$$p(y \mid a) = \frac{1}{2}(1 + y\eta_a), \quad \eta_a \in (-1, 1). \tag{2}$$

The state update is $\gamma' = \gamma + ya$.

Notation. Define:

$$\bar{\eta} = \frac{\eta_+ + \eta_-}{2}, \quad \Delta\eta = \frac{\eta_+ - \eta_-}{2}, \tag{3}$$

$$b(\gamma) = 2\Phi(\beta\gamma) - 1 = \mathbb{E}[a \mid \gamma]. \tag{4}$$

Throughout, Φ denotes the standard normal CDF and φ denotes the standard normal PDF:

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\, dt, \quad \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{5}$$

2 Master Equation

Let $p_n(\gamma)$ denote the probability of state γ at discrete time n. The master equation is:

$$p_{n+1}(\gamma) = \sum_{y,a \in \{\pm 1\}} p(y \mid a)\, p(a \mid \gamma - ya)\, p_n(\gamma - ya). \tag{6}$$

Expanding the four terms $(y, a) \in \{(+,+), (+,-), (-,+), (-,-)\}$:

$$\begin{aligned} p_{n+1}(\gamma) = {} & \frac{1+\eta_+}{2} \Phi(\beta(\gamma - 1))\, p_n(\gamma - 1) + \frac{1+\eta_-}{2} [1 - \Phi(\beta(\gamma + 1))]\, p_n(\gamma + 1) \\ & + \frac{1-\eta_+}{2} \Phi(\beta(\gamma + 1))\, p_n(\gamma + 1) + \frac{1-\eta_-}{2} [1 - \Phi(\beta(\gamma - 1))]\, p_n(\gamma - 1). \end{aligned} \tag{7}$$

Collecting terms at $\gamma \pm 1$:

$$p_{n+1}(\gamma) = R^+(\gamma - 1)\, p_n(\gamma - 1) + R^-(\gamma + 1)\, p_n(\gamma + 1), \tag{8}$$

where the transition rates are:

$$R^{+}(\gamma) = \frac{1+\eta_+}{2}\Phi(\beta\gamma) + \frac{1-\eta_-}{2}[1-\Phi(\beta\gamma)], \tag{9}$$

$$R^{-}(\gamma) = \frac{1-\eta_+}{2}\Phi(\beta\gamma) + \frac{1+\eta_-}{2}[1-\Phi(\beta\gamma)]. \tag{10}$$

After simplification:

$$R^{+}(\gamma) = \frac{1}{2}\big[1+\bar{\eta}+\Delta\eta\, b(\gamma)\big], \tag{11}$$

$$R^{-}(\gamma) = \frac{1}{2}\big[1-\bar{\eta}-\Delta\eta\, b(\gamma)\big]. \tag{12}$$

Note that $R^+ + R^- = 1$ (probability conservation).

3 Cumulant Generating Function

Let $\langle\cdot\rangle_n$ denote expectation with respect to p_n. Define the cumulant generating function:

$$K_n(\theta) = \log \mathbb{E}_n[e^{\theta\gamma}] = \log \sum_{\gamma} e^{\theta\gamma} p_n(\gamma). \tag{13}$$

The cumulants $\kappa_k^{(n)}$ are:

$$\kappa_k^{(n)} = \left.\frac{d^k K_n}{d\theta^k}\right|_{\theta=0}. \tag{14}$$

Explicitly: $\kappa_1 = \langle\gamma\rangle$ (mean), $\kappa_2 = \langle\gamma^2\rangle - \langle\gamma\rangle^2$ (variance), $\kappa_3 = \langle(\gamma-\langle\gamma\rangle)^3\rangle$ (skewness times σ^3), etc.

4 General Cumulant Update

Define the increment $\xi = ya$. Given γ, the conditional distribution of ξ is:

$$p(\xi \mid \gamma) = R^{+}(\gamma)\delta_{\xi,+1} + R^{-}(\gamma)\delta_{\xi,-1}. \tag{15}$$

The conditional cumulant generating function of ξ is:

$$K_{\xi|\gamma}(\theta) = \log\big[R^{+}(\gamma)e^{\theta} + R^{-}(\gamma)e^{-\theta}\big]. \tag{16}$$

Since $\gamma' = \gamma + \xi$, the CGF at time $n+1$ is:

$$K_{n+1}(\theta) = \log \mathbb{E}_n\big[e^{\theta\gamma}\cdot(R^{+}(\gamma)e^{\theta} + R^{-}(\gamma)e^{-\theta})\big]. \tag{17}$$

4.1 Cumulant Extraction

To find the update for κ_k, differentiate $K_{n+1}(\theta)$ and set $\theta = 0$. Using $R^{\pm}(\gamma) = \frac{1}{2}[1 \pm v(\gamma)]$ where

$$v(\gamma) \equiv R^{+}(\gamma) - R^{-}(\gamma) = \bar{\eta} + \Delta\eta\, b(\gamma), \tag{18}$$

we have:

$$R^{+}e^{\theta} + R^{-}e^{-\theta} = \cosh\theta + v(\gamma)\sinh\theta. \tag{19}$$

Define:

$$F(\theta, \gamma) = \log[\cosh\theta + v(\gamma)\sinh\theta]. \tag{20}$$

The derivatives are:

$$\partial_\theta F = \frac{\sinh\theta + v\cosh\theta}{\cosh\theta + v\sinh\theta}, \tag{21}$$

$$\partial_\theta^2 F = \frac{1 - v^2}{(\cosh\theta + v\sinh\theta)^2}. \tag{22}$$

At $\theta = 0$:

$$F(0, \gamma) = 0, \quad \partial_\theta F\big|_{\theta=0} = v(\gamma), \quad \partial_\theta^2 F\big|_{\theta=0} = 1 - v(\gamma)^2. \tag{23}$$

5 Mean Update (κ_1)

The mean satisfies:

$$\kappa_1^{(n+1)} = \left.\frac{dK_{n+1}}{d\theta}\right|_{\theta=0} = \left.\frac{dK_n}{d\theta}\right|_{\theta=0} + \langle v(\gamma)\rangle_n. \tag{24}$$

Hence:

$$\boxed{\kappa_1^{(n+1)} = \kappa_1^{(n)} + \langle v(\gamma)\rangle_n = \kappa_1^{(n)} + \bar{\eta} + \Delta\eta\,\langle b(\gamma)\rangle_n.} \tag{25}$$

This shows: the mean drifts by the expected velocity $\langle v(\gamma)\rangle$, which depends on the current distribution through $\langle b(\gamma)\rangle$.

6 Variance Update (κ_2)

For the variance, use the law of total variance. The increment $\xi = ya \in \{+1, -1\}$ satisfies $\xi^2 = 1$ always, so:

$$\mathrm{Var}(\gamma') = \mathrm{Var}(\gamma + \xi) = \mathrm{Var}(\gamma) + \mathrm{Var}(\xi) + 2\mathrm{Cov}(\gamma, \xi). \tag{26}$$

Since $\mathbb{E}[\xi \mid \gamma] = v(\gamma)$ and $\mathbb{E}[\xi^2] = 1$:

$$\mathbb{E}[\xi] = \langle v\rangle_n, \tag{27}$$

$$\mathrm{Var}(\xi) = \mathbb{E}[\xi^2] - \mathbb{E}[\xi]^2 = 1 - \langle v\rangle_n^2. \tag{28}$$

For the covariance, since ξ depends on γ only through $v(\gamma)$:

$$\mathrm{Cov}(\gamma, \xi) = \mathbb{E}[\gamma \cdot \mathbb{E}[\xi|\gamma]] - \mathbb{E}[\gamma]\mathbb{E}[\xi] = \mathbb{E}[\gamma\, v(\gamma)] - \mu_n\langle v\rangle_n = \mathrm{Cov}_n(\gamma, v). \tag{29}$$

Therefore:

$$\boxed{\kappa_2^{(n+1)} = \kappa_2^{(n)} + 1 - \langle v\rangle_n^2 + 2\,\mathrm{Cov}_n(\gamma, v(\gamma)).} \tag{30}$$

Expanding $v(\gamma) = \bar{\eta} + \Delta\eta\, b(\gamma)$:

$$\langle v\rangle = \bar{\eta} + \Delta\eta\langle b\rangle, \quad \mathrm{Cov}(\gamma, v) = \Delta\eta\,\mathrm{Cov}(\gamma, b). \tag{31}$$

So the variance update becomes:

$$\boxed{\sigma_{n+1}^2 = \sigma_n^2 + 1 - (\bar{\eta} + \Delta\eta\langle b\rangle_n)^2 + 2\Delta\eta\,\mathrm{Cov}_n(\gamma, b(\gamma)).} \tag{32}$$

Special case: $\bar{\eta} = 0$ (symmetric arms).

$$\sigma_{n+1}^2 = \sigma_n^2 + 1 - \Delta\eta^2 \langle b \rangle_n^2 + 2\Delta\eta \, \mathrm{Cov}_n(\gamma, b). \tag{33}$$

7 General k-th Cumulant Update

For arbitrary k, the update takes the form:

$$\kappa_k^{(n+1)} = \kappa_k^{(n)} + \Delta\kappa_k(p_n), \tag{34}$$

where $\Delta\kappa_k$ is a functional of the distribution p_n.

Define the "local cumulants" of ξ given γ:

$$c_1(\gamma) = v(\gamma), \tag{35}$$

$$c_2(\gamma) = 1 - v(\gamma)^2, \tag{36}$$

$$c_k(\gamma) = \frac{d^k}{d\theta^k} \log[R^+ e^{\theta} + R^- e^{-\theta}]\big|_{\theta=0}. \tag{37}$$

These are:

$$c_3(\gamma) = v(1 - v^2) - (1 - v^2)(-2v) = v(1 - v^2) + 2v(1 - v^2) = -v(1 - v^2)(1 - 2) \tag{38}$$

$$= 2v(1 - v^2) - 2v(1 - v^2) = -2v^3 + 2v. \tag{39}$$

Actually, let us compute more carefully. With $R^{\pm} = (1 \pm v)/2$:

$$c_3 = \frac{d^3}{d\theta^3} \log(\cosh\theta + v \sinh\theta)\big|_{\theta=0}. \tag{40}$$

Using $F''' = F'(1 - 2vF') - F''(2vF') + \ldots$, one finds:

$$c_3(\gamma) = 2v(v^2 - 1) = -2v(1 - v^2). \tag{41}$$

And in general, the local cumulants satisfy a recurrence.

Cumulant mixing formula. The update for the k-th cumulant involves mixed moments:

$$\Delta\kappa_k = \sum_{j=0}^{k} \binom{k}{j} \cdot (\text{joint cumulant of } j \text{ copies of } \gamma \text{ and 1 copy of } \xi) - \kappa_k^{(n)}. \tag{42}$$

For practical computation, we express everything in terms of the functionals:

$$\langle \gamma^m b(\gamma)^{\ell} \rangle_n, \tag{43}$$

which must be evaluated under the current distribution p_n.

7.1 Edgeworth Representation

If we represent the continuum distribution $q(\gamma)$ via an Edgeworth expansion:

$$q(\gamma) = \varphi\Big(\frac{\gamma-\mu}{\sigma}\Big)\left[1 + \frac{\kappa_3}{6\sigma^3}H_3(z) + \frac{\kappa_4}{24\sigma^4}H_4(z) + \cdots\right], \tag{44}$$

where $z = (\gamma - \mu)/\sigma$, φ is the standard normal density, and H_k are Hermite polynomials, then:

- $\kappa_1 = \mu$ (the mean),
- $\kappa_2 = \sigma^2$ (the variance),
- $\kappa_3, \kappa_4, \ldots$ appear as expansion coefficients.

The discrete-to-continuum matching condition is:

$$\kappa_k^{\text{discrete}}(n) = \kappa_k^{\text{Edgeworth}}(t = n\Delta t), \quad \forall k. \tag{45}$$

8 Summary: First Two Cumulant Updates

Mean:

$$\mu_{n+1} = \mu_n + \bar{\eta} + \Delta\eta\,\langle b(\gamma)\rangle_n. \tag{46}$$

Variance:

$$\sigma_{n+1}^2 = \sigma_n^2 + 1 - \langle v\rangle_n^2 + 2\Delta\eta\,\mathrm{Cov}_n(\gamma, b(\gamma)), \tag{47}$$

where $v(\gamma) = \bar{\eta} + \Delta\eta\, b(\gamma)$ and $b(\gamma) = 2\Phi(\beta\gamma) - 1$.

Remark. The expectations $\langle\cdot\rangle_n$ must be computed under the current distribution p_n. In the Edgeworth/Gaussian approximation, these become functions of $(\mu_n, \sigma_n^2, \kappa_3, \ldots)$, closing the system.

For a Gaussian approximation (truncating at κ_2), we need:

$$\langle b(\gamma)\rangle_n \approx \int_{-\infty}^{\infty} b(\gamma)\cdot\frac{1}{\sqrt{2\pi\sigma_n^2}}e^{-(\gamma-\mu_n)^2/(2\sigma_n^2)}\,d\gamma, \tag{48}$$

and similarly for $\langle b^2\rangle$, $\langle\gamma b\rangle$, etc.

9 Exact Gaussian-Probit Integrals

When Φ is the *probit* (standard normal CDF), the required expectations have exact closed forms. This follows from the fundamental identity:

9.1 The Owen-type Identity

For $\gamma \sim \mathcal{N}(\mu, \sigma^2)$, the expectation of $\Phi(a + b\gamma)$ is:

$$\boxed{\mathbb{E}[\Phi(a+b\gamma)] = \Phi\left(\frac{a+b\mu}{\sqrt{1+b^2\sigma^2}}\right).} \tag{49}$$

Proof. Write $\gamma = \mu + \sigma Z$ where $Z \sim \mathcal{N}(0,1)$. Then $a + b\gamma = (a + b\mu) + b\sigma Z$. We need $\mathbb{E}[\Phi((a + b\mu) + b\sigma Z)]$. Using the identity $\mathbb{E}[\Phi(\alpha + \beta Z)] = \Phi(\alpha/\sqrt{1+\beta^2})$ (which follows from $\Phi(\alpha+\beta Z) = P(W < \alpha+\beta Z)$ for independent $W \sim \mathcal{N}(0,1)$, so $P(W - \beta Z < \alpha) = \Phi(\alpha/\sqrt{1+\beta^2})$), we get the result with $\alpha = a + b\mu$ and $\beta = b\sigma$.

9.2 Application to the Mean Update

Since $b(\gamma) = 2\Phi(\beta\gamma) - 1$, we have:

$$\langle b(\gamma)\rangle_n = 2\,\mathbb{E}_n[\Phi(\beta\gamma)] - 1 = 2\Phi\left(\frac{\beta\mu_n}{\sqrt{1+\beta^2\sigma_n^2}}\right) - 1. \tag{50}$$

Define the **effective inverse temperature**:

$$\tilde{\beta}_n \equiv \frac{\beta}{\sqrt{1+\beta^2\sigma_n^2}}. \tag{51}$$

Then:

$$\boxed{\langle b(\gamma)\rangle_n = 2\Phi(\tilde{\beta}_n\mu_n) - 1 = \operatorname{erf}\left(\frac{\tilde{\beta}_n\mu_n}{\sqrt{2}}\right),} \tag{52}$$

where erf is the error function (for probit Φ).

The mean update becomes:

$$\boxed{\mu_{n+1} = \mu_n + \bar{\eta} + \Delta\eta \cdot \operatorname{erf}\left(\frac{\tilde{\beta}_n\mu_n}{\sqrt{2}}\right).} \tag{53}$$

Interpretation. The effective $\tilde{\beta}$ decreases as variance σ^2 grows. Large uncertainty "softens" the policy, reducing exploitation.

9.3 Higher Moments via Hermite Polynomials

For $\gamma \sim \mathcal{N}(\mu, \sigma^2)$, define $z = (\gamma - \mu)/\sigma$. Then $\gamma^j = \sum_{k=0}^{j} \binom{j}{k}\mu^{j-k}\sigma^k z^k$, and we need:

$$I_k(a,b) \equiv \int_{-\infty}^{\infty} z^k \varphi(z)\Phi(a + bz)\,dz, \quad |b| < 1. \tag{54}$$

The key recurrence (integration by parts) is:

$$I_k(a,b) = (k-1)I_{k-2}(a,b) + b \cdot J_{k-1}(a,b), \tag{55}$$

where $J_k(a,b) = \int z^k\varphi(z)\varphi(a+bz)\,dz$.

Base cases.

$$I_0(a,b) = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right), \tag{56}$$

$$I_1(a,b) = \frac{b}{\sqrt{1+b^2}}\,\varphi\left(\frac{a}{\sqrt{1+b^2}}\right). \tag{57}$$

Second moment. Using the recurrence:

$$I_2(a,b) = I_0(a,b) + b \cdot J_1(a,b). \tag{58}$$

Since $J_1(a,b) = -\frac{ab}{(1+b^2)^{3/2}}\varphi\left(\frac{a}{\sqrt{1+b^2}}\right)$, we get:

$$I_2(a,b) = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right) - \frac{ab^2}{(1+b^2)^{3/2}}\varphi\left(\frac{a}{\sqrt{1+b^2}}\right). \tag{59}$$

9.4 Closed Form for Variance Update

We need $\langle b \rangle$ and $\mathrm{Cov}(\gamma, b)$ under $\gamma \sim \mathcal{N}(\mu_n, \sigma_n^2)$.

Define $\tilde{a} = \beta\mu_n$, $\tilde{b} = \beta\sigma_n$, and $\rho = 1/\sqrt{1+\tilde{b}^2}$ (so $\tilde{\beta}_n = \beta\rho$).

1. Expectation of $b(\gamma)$:

$$\langle b \rangle = 2\mathbb{E}[\Phi(\beta\gamma)] - 1 = 2\Phi(\tilde{\beta}\mu) - 1. \tag{60}$$

2. Covariance $\mathrm{Cov}(\gamma, b)$:

$$\mathrm{Cov}(\gamma, b) = 2\mathrm{Cov}(\gamma, \Phi(\beta\gamma)) = 2\big(\mathbb{E}[\gamma\Phi(\beta\gamma)] - \mu\mathbb{E}[\Phi(\beta\gamma)]\big). \tag{61}$$

Using $\gamma = \mu + \sigma z$ and the I_1 integral:

$$\mathbb{E}[\gamma\Phi(\beta\gamma)] = \mu\Phi(\tilde{\beta}\mu) + \sigma \cdot I_1(\tilde{a}, \tilde{b}) = \mu\Phi(\tilde{\beta}\mu) + \frac{\sigma\tilde{b}}{\sqrt{1+\tilde{b}^2}}\varphi(\tilde{\beta}\mu). \tag{62}$$

Thus:

$$\boxed{\mathrm{Cov}(\gamma, b) = \frac{2\beta\sigma^2}{\sqrt{1+\beta^2\sigma^2}}\varphi(\tilde{\beta}\mu).} \tag{63}$$

9.5 Closed-Form Cumulant Dynamics

Under the Gaussian ansatz, the cumulant updates become a 2D map $(\mu, \sigma^2) \mapsto (\mu', \sigma'^2)$.

Define $\tilde{\beta} = \beta/\sqrt{1+\beta^2\sigma^2}$ and $\langle b \rangle = 2\Phi(\tilde{\beta}\mu) - 1$.

$$\mu' = \mu + \bar{\eta} + \Delta\eta \cdot (2\Phi(\tilde{\beta}\mu) - 1), \tag{64}$$

$$\sigma'^2 = \sigma^2 + 1 - (\bar{\eta} + \Delta\eta\langle b \rangle)^2 + \frac{4\Delta\eta^2\beta\sigma^2}{\sqrt{1+\beta^2\sigma^2}}\varphi(\tilde{\beta}\mu). \tag{65}$$

Simplification for $\bar{\eta} = 0$:

$$\mu' = \mu + \Delta\eta \cdot (2\Phi(\tilde{\beta}\mu) - 1), \tag{66}$$

$$\sigma'^2 = \sigma^2 + 1 - \Delta\eta^2(2\Phi(\tilde{\beta}\mu) - 1)^2 + \frac{4\Delta\eta^2\beta\sigma^2}{\sqrt{1+\beta^2\sigma^2}}\varphi(\tilde{\beta}\mu). \tag{67}$$

Interpretation. The variance update has three terms:

- $+1$: diffusion from the random walk
- $-\langle v \rangle^2$: reduction from mean drift (deterministic component)
- $+2\Delta\eta\,\mathrm{Cov}(\gamma, b)$: amplification when γ and policy are correlated

The covariance term $\propto \varphi(\tilde{\beta}\mu)$ is maximal when $\mu \approx 0$ (agent uncertain) and vanishes when $|\mu|$ is large (agent committed).

10 Exact Cumulant Updates via Probit-Polynomial-Gaussian Integrals

Under the Gaussian ansatz, all cumulant updates have exact closed forms.

10.1 Fundamental Integrals

For $\gamma \sim \mathcal{N}(\mu, \sigma^2)$ with $z = (\gamma - \mu)/\sigma$, define:

$$\rho = \frac{1}{\sqrt{1 + \beta^2\sigma^2}}, \quad u = \frac{\beta\mu}{\sqrt{1 + \beta^2\sigma^2}} = \beta\rho\mu, \quad \lambda = \beta\sigma\rho. \tag{68}$$

The probit-polynomial-Gaussian integrals are:

$$I_k \equiv \int_{-\infty}^{\infty} z^k \varphi(z)\Phi(u + \lambda z)\, dz. \tag{69}$$

Exact formulas via Hermite polynomials. The probabilist's Hermite polynomials $\mathrm{He}_k(u)$ satisfy $\mathrm{He}_k(u) = u\,\mathrm{He}_{k-1}(u) - (k-1)\mathrm{He}_{k-2}(u)$ with $\mathrm{He}_0 = 1$, $\mathrm{He}_1 = u$, $\mathrm{He}_2 = u^2 - 1$, $\mathrm{He}_3 = u^3 - 3u$, $\mathrm{He}_4 = u^4 - 6u^2 + 3$.

The integrals have the structure:

$$I_0 = \Phi(u), \tag{70}$$
$$I_1 = \lambda\varphi(u), \tag{71}$$
$$I_2 = \Phi(u) - \lambda^2\mathrm{He}_1(u)\varphi(u), \tag{72}$$
$$I_3 = \lambda[3 + \lambda^2\mathrm{He}_2(u)]\varphi(u), \tag{73}$$
$$I_4 = 3\Phi(u) - \lambda^2[3\mathrm{He}_1(u) + \lambda^2\mathrm{He}_3(u)]\varphi(u), \tag{74}$$
$$I_5 = \lambda[15 + 6\lambda^2\mathrm{He}_2(u) + \lambda^4\mathrm{He}_4(u)]\varphi(u). \tag{75}$$

Recurrence. Using $\partial_u\Phi(u) = \varphi(u)$ and $\partial_u[\mathrm{He}_j(u)\varphi(u)] = -\mathrm{He}_{j+1}(u)\varphi(u)$:

$$I_k = (k-1)I_{k-2} + \lambda\frac{\partial I_{k-1}}{\partial u}. \tag{76}$$

This shows that I_k is built from $\{\Phi(u), \varphi(u)\}$ with coefficients that are polynomials in (λ^2, u), where the u-dependence enters through Hermite polynomials.

Structure. For k even: $I_k = (k-1)!!\Phi(u) + \lambda \cdot (\text{polynomial in } \lambda^2) \cdot \varphi(u)$. For k odd: $I_k = \lambda \cdot (\text{polynomial in } \lambda^2, u) \cdot \varphi(u)$. The leading Hermite polynomial in I_k is $\mathrm{He}_{k-1}(u)$ with coefficient λ^{k-1}.

10.2 Notation: Mean Policy

Since the combination $2\Phi(u) - 1$ appears throughout, we define:

$$\mathcal{B} \equiv 2\Phi(u) - 1 = \operatorname{erf}\left(\frac{u}{\sqrt{2}}\right) = \langle b \rangle. \tag{77}$$

This is the **mean policy**: the expected value of $b(\gamma) = 2\Phi(\beta\gamma) - 1$ under the Gaussian ansatz. Explicitly,

$$\mathcal{B} = 2\Phi\left(\frac{\beta\mu}{\sqrt{1+\beta^2\sigma^2}}\right) - 1. \tag{78}$$

We also write $\mathcal{Q} = \varphi(u)$ for the Gaussian density at the effective argument.

10.3 Moments of the Policy

Since $b(\gamma) = 2\Phi(\beta\gamma) - 1$:

$$\langle b \rangle = 2I_0 - 1 = \mathcal{B}, \tag{79}$$
$$\langle z\, b \rangle = 2I_1 = 2\lambda\mathcal{Q}, \tag{80}$$
$$\langle z^2 b \rangle = 2I_2 - 1 = \mathcal{B} - 2u\lambda^2\mathcal{Q}, \tag{81}$$
$$\langle z^k b \rangle = 2I_k - \langle z^k \rangle_{\text{Gaussian}}. \tag{82}$$

Converting to centered moments: with $\tilde{\gamma} = \gamma - \mu = \sigma z$,

$$\langle \tilde{\gamma}^k b \rangle = \sigma^k \langle z^k b \rangle = \sigma^k (2I_k - \delta_{k,\text{even}}(k-1)!!). \tag{83}$$

10.4 Increment Statistics

The increment $\xi = ya \in \{+1, -1\}$ has conditional moments:

$$\mathbb{E}[\xi^k|\gamma] = \begin{cases} 1 & k \text{ even} \\ v(\gamma) & k \text{ odd} \end{cases}, \quad v(\gamma) = \bar{\eta} + \Delta\eta\, b(\gamma). \tag{84}$$

Define the centered increment $\tilde{\xi} = \xi - \langle v \rangle$ where $\langle v \rangle = \bar{\eta} + \Delta\eta\langle b \rangle$.

Key expectations:

$$\mathbb{E}[\tilde{\xi}] = 0, \tag{85}$$
$$\mathbb{E}[\tilde{\xi}^2] = 1 - \langle v \rangle^2, \tag{86}$$
$$\mathbb{E}[\tilde{\xi}^3] = \langle v \rangle - 3\langle v \rangle + 2\langle v \rangle^3 = -2\langle v \rangle(1 - \langle v \rangle^2), \tag{87}$$
$$\mathbb{E}[\tilde{\xi}^4] = 1 - 6\langle v \rangle^2 + 6\langle v \rangle^4 - \langle v \rangle^4 + 6\langle v \rangle^2 = 1 - \langle v \rangle^4 + 6\langle v \rangle^2(\langle v \rangle^2 - 1). \tag{88}$$

Cross terms: Since $\mathbb{E}[\tilde{\gamma}^j \xi|\gamma] = \tilde{\gamma}^j v(\gamma)$ for odd powers of ξ,

$$\mathbb{E}[\tilde{\gamma}^j \tilde{\xi}] = \mathbb{E}[\tilde{\gamma}^j v] - \langle v \rangle \mathbb{E}[\tilde{\gamma}^j] = \Delta\eta\big(\langle \tilde{\gamma}^j b \rangle - \langle b \rangle \kappa_j\big) \equiv \Delta\eta\, C_j, \tag{89}$$

where $C_j = \sigma^j(2I_j - \delta_{j,\text{even}}(j-1)!!) - \mathcal{B}\,\kappa_j$ is the **policy-cumulant covariance**.

10.5 General Cumulant Update

The central moments of $\gamma' = \gamma + \xi$ expand via the binomial theorem. After converting to cumulants, the exact update is:

$$\boxed{\Delta\kappa_j = \sum_{k=0}^{j} \binom{j}{k} M_{j-k,k} - \kappa_j} \tag{90}$$

where $M_{m,n} = \mathbb{E}[\tilde{\gamma}^m \tilde{\xi}^n]$ with $\tilde{\gamma} = \gamma - \mu$, $\tilde{\xi} = \xi - \langle v \rangle$.

General formula for $M_{m,n}$. Expand via binomial theorem:

$$M_{m,n} = \mathbb{E}[\tilde{\gamma}^m \tilde{\xi}^n] = \mathbb{E}\left[\tilde{\gamma}^m \sum_{\ell=0}^{n} \binom{n}{\ell} \xi^\ell (-\langle v \rangle)^{n-\ell}\right] = \sum_{\ell=0}^{n} \binom{n}{\ell} (-\langle v \rangle)^{n-\ell} \mathbb{E}[\tilde{\gamma}^m \xi^\ell]. \tag{91}$$

Since $\xi^2 = 1$, we have $\xi^\ell = 1$ for ℓ even and $\xi^\ell = \xi$ for ℓ odd. Using $\mathbb{E}[\xi|\gamma] = v(\gamma)$:

$$\mathbb{E}[\tilde{\gamma}^m \xi^\ell] = \begin{cases} \kappa_m & \ell \text{ even} \\ \mathbb{E}[\tilde{\gamma}^m v(\gamma)] = \bar{\eta}\kappa_m + \Delta\eta \langle \tilde{\gamma}^m b \rangle & \ell \text{ odd} \end{cases}. \tag{92}$$

Therefore:

$$\boxed{M_{m,n} = \sum_{\ell=0}^{n} \binom{n}{\ell} (-\langle v \rangle)^{n-\ell} \begin{cases} \kappa_m & \ell \text{ even} \\ \bar{\eta}\kappa_m + \Delta\eta \langle \tilde{\gamma}^m b \rangle & \ell \text{ odd} \end{cases}} \tag{93}$$

where $\langle \tilde{\gamma}^m b \rangle = 2\sigma^m I_m - \mathcal{B}\,\kappa_m$ from the probit integrals.

Simplification. Define $V_m = \bar{\eta}\kappa_m + \Delta\eta \langle \tilde{\gamma}^m b \rangle = \mathbb{E}[\tilde{\gamma}^m v]$. Separating even and odd terms:

$$M_{m,n} = \kappa_m \sum_{\ell \text{ even}} \binom{n}{\ell} (-\langle v \rangle)^{n-\ell} + V_m \sum_{\ell \text{ odd}} \binom{n}{\ell} (-\langle v \rangle)^{n-\ell}. \tag{94}$$

Using $(1-x)^n \pm (-1-x)^n = 2\sum_{\ell \text{ even/odd}} \binom{n}{\ell}(-x)^{n-\ell}$:

$$\boxed{M_{m,n} = \frac{\kappa_m + V_m}{2}(1 - \langle v \rangle)^n + \frac{\kappa_m - V_m}{2}(-1 - \langle v \rangle)^n} \tag{95}$$

Special cases:

$$M_{m,0} = \kappa_m, \tag{96}$$
$$M_{m,1} = V_m - \langle v \rangle \kappa_m = \Delta\eta(\langle \tilde{\gamma}^m b \rangle - \langle b \rangle \kappa_m), \tag{97}$$
$$M_{m,2} = \kappa_m(1 + \langle v \rangle^2) - 2\langle v \rangle V_m = (1 - \langle v \rangle^2)\kappa_m + 2\Delta\eta \langle v \rangle(\langle b \rangle \kappa_m - \langle \tilde{\gamma}^m b \rangle). \tag{98}$$

10.6 Explicit Updates: κ_1 through κ_4

With $\mathcal{B} = 2\Phi(u) - 1$, $\mathcal{Q} = \varphi(u)$, and $\langle v \rangle = \bar{\eta} + \Delta\eta\, \mathcal{B}$:

Mean:

$$\boxed{\Delta\kappa_1 = \bar{\eta} + \Delta\eta\, \mathcal{B}} \tag{99}$$

Variance:

$$\boxed{\Delta\kappa_2 = 1 - \langle v \rangle^2 + 4\Delta\eta\lambda\sigma\mathcal{Q}} \tag{100}$$

Third cumulant:

$$\boxed{\Delta\kappa_3 = 6\Delta\eta\lambda\sigma^2(1 - u\lambda)\mathcal{Q} - 2\langle v \rangle(1 - \langle v \rangle^2)} \tag{101}$$

Fourth cumulant:

$$\begin{aligned}\Delta\kappa_4 = &-2(1 - \langle v \rangle^2)(1 - 3\langle v \rangle^2) \\ &+ 8\Delta\eta\lambda\sigma^3(3 - 3u\lambda + u^2\lambda^2 - \lambda^2)\mathcal{Q} \\ &+ 6\sigma^2\big[(1 - \langle v \rangle^2) + \Delta\eta^2(1 - \mathcal{B}^2 - 4\lambda^2\mathcal{Q}^2)\big].\end{aligned} \tag{102}$$

10.7 Compact Notation

Define the **policy moments**:

$$B_k = 2\sigma^k I_k, \quad k = 0, 1, 2, \ldots \tag{103}$$

Then the cumulant updates are polynomials in $\{B_k, \bar{\eta}, \Delta\eta, \sigma\}$:

$$\Delta\kappa_j = P_j(B_0, B_1, \ldots, B_j; \bar{\eta}, \Delta\eta, \sigma), \tag{104}$$

where P_j is explicitly computable from the binomial expansion.

10.8 Continuum Limit

For large t, treating $\Delta\kappa_j$ as a rate gives the ODE system:

$$\frac{d\kappa_j}{dt} = \Delta\kappa_j(\mu, \sigma^2), \tag{105}$$

which is closed since each $\Delta\kappa_j$ depends only on (μ, σ^2) through $(u, \lambda, \mathcal{B}, \mathcal{Q})$.

10.9 Power-Law Decay

For $\Delta\eta = 0$ (identical arms), CLT applies and standardized cumulants decay as:

$$\frac{\kappa_j}{\sigma^j} \sim t^{-(j-2)/2}, \quad j \geq 3. \tag{106}$$

11 Edgeworth-K Truncation

The Gaussian ansatz ($K = 2$) tracks only mean and variance. For higher accuracy, we use the Edgeworth expansion truncated at order K, tracking cumulants $\kappa_1, \ldots, \kappa_K$.

11.1 Edgeworth Expansion

The K-truncated Edgeworth expansion is:

$$p_K(\gamma) = \frac{1}{\sigma}\varphi(z)\left[1 + \sum_{j=3}^{K} \frac{\kappa_j}{j!\,\sigma^j}\mathrm{He}_j(z)\right], \tag{107}$$

where $z = (\gamma - \mu)/\sigma$ and He_j are probabilist's Hermite polynomials.

Normalization. The expansion satisfies $\int p_K \, d\gamma = 1$ since $\int \varphi(z)\mathrm{He}_j(z)\, dz = 0$ for $j \geq 1$.

Cumulant recovery. Under p_K, the cumulants are exactly $\kappa_1, \ldots, \kappa_K$ (to leading order in the expansion), with higher cumulants treated as zero.

11.2 Expectations Under p_K

For any function $f(\gamma)$, the expectation under p_K is:

$$\langle f \rangle_K = \langle f \rangle_{\mathrm{Gauss}} + \sum_{j=3}^{K} \frac{\kappa_j}{j!\,\sigma^j}\langle f(\gamma)\,\mathrm{He}_j(z)\rangle_{\mathrm{Gauss}}. \tag{108}$$

For the policy $b(\gamma) = 2\Phi(\beta\gamma) - 1$, we need:

$$\langle b \rangle_K = \mathcal{B} + \sum_{j=3}^{K} \frac{\kappa_j}{j!\,\sigma^j} \cdot 2J_{0,j}, \tag{109}$$

$$\langle \tilde{\gamma}^m b \rangle_K = 2\sigma^m I_m - \mathcal{B}\kappa_m + \sum_{j=3}^{K} \frac{\kappa_j}{j!\,\sigma^j} \cdot 2\sigma^m J_{m,j}, \tag{110}$$

where $J_{m,j}$ are the **probit-Hermite-Gaussian integrals**:

$$\boxed{J_{m,j}(u,\lambda) \equiv \int_{-\infty}^{\infty} z^m \varphi(z)\mathrm{He}_j(z)\Phi(u + \lambda z)\, dz.} \tag{111}$$

11.3 Analytic Form of $J_{m,j}$

Key identity. The Hermite polynomials satisfy:

$$\mathrm{He}_j(z)\varphi(z) = (-1)^j \frac{d^j\varphi}{dz^j}(z). \tag{112}$$

Therefore:

$$J_{m,j} = (-1)^j \int z^m \frac{d^j\varphi}{dz^j}(z)\Phi(u + \lambda z)\, dz. \tag{113}$$

Integration by parts. Applying integration by parts j times, moving all derivatives onto $\Phi(u + \lambda z)$:

$$J_{m,j} = \int z^m \varphi(z) \frac{d^j}{dz^j}\Phi(u + \lambda z)\, dz = \lambda^j \int z^m \varphi(z)\mathrm{He}_{j-1}(\lambda z + u)\varphi(u + \lambda z)\, dz, \tag{114}$$

using $\frac{d^j\Phi}{dz^j}(u + \lambda z) = \lambda^j \frac{d^{j-1}\varphi}{du^{j-1}}(u + \lambda z) = \lambda^j (-1)^{j-1}\mathrm{He}_{j-1}(u + \lambda z)\varphi(u + \lambda z)$.

Result. $J_{m,j}$ is expressible in terms of derivatives of I_m with respect to u:

$$\boxed{J_{m,j} = \frac{\partial^j I_m}{\partial u^j}.} \tag{115}$$

Proof. Starting from $I_m(u,\lambda) = \int z^m \varphi(z)\Phi(u+\lambda z)\,dz$,

$$\frac{\partial^j I_m}{\partial u^j} = \int z^m \varphi(z)\frac{\partial^j \Phi}{\partial u^j}(u+\lambda z)\,dz = \int z^m \varphi(z)\varphi^{(j-1)}(u+\lambda z)\,dz, \tag{116}$$

where $\varphi^{(k)}$ denotes the k-th derivative of φ.

But $\varphi^{(k)}(x) = (-1)^k \mathrm{He}_k(x)\varphi(x)$, so:

$$\frac{\partial^j I_m}{\partial u^j} = (-1)^{j-1}\int z^m \varphi(z)\mathrm{He}_{j-1}(u+\lambda z)\varphi(u+\lambda z)\,dz. \tag{117}$$

Comparing with the Hermite identity applied to the original $J_{m,j}$ integral, after careful tracking of signs, we get $J_{m,j} = \partial_u^j I_m$.

11.4 Explicit $J_{m,j}$ Formulas

Using $\partial_u \Phi(u) = \varphi(u)$ and $\partial_u \varphi(u) = -u\varphi(u) = -\mathrm{He}_1(u)\varphi(u)$, and the general rule $\partial_u[\mathrm{He}_k(u)\varphi(u)] = -\mathrm{He}_{k+1}(u)\varphi(u)$:

$j = 0$ (standard integrals):

$$J_{m,0} = I_m. \tag{118}$$

$j = 1$:

$$J_{0,1} = \partial_u I_0 = \partial_u \Phi(u) = \varphi(u) = \mathcal{Q}, \tag{119}$$

$$J_{1,1} = \partial_u I_1 = \partial_u[\lambda\varphi(u)] = -\lambda u\varphi(u) = -\lambda \mathrm{He}_1(u)\mathcal{Q}, \tag{120}$$

$$\begin{aligned} J_{2,1} &= \partial_u I_2 = \partial_u[\Phi(u) - \lambda^2 u\varphi(u)] = \varphi(u) - \lambda^2[\varphi(u) - u^2\varphi(u)] \\ &= (1-\lambda^2)\mathcal{Q} + \lambda^2 u^2 \mathcal{Q} = [1 - \lambda^2 + \lambda^2 \mathrm{He}_2(u) + \lambda^2]\mathcal{Q} \\ &= [1 + \lambda^2 \mathrm{He}_2(u)]\mathcal{Q}. \end{aligned} \tag{121}$$

$j = 2$:

$$J_{0,2} = \partial_u^2 I_0 = \partial_u \varphi(u) = -u\varphi(u) = -\mathrm{He}_1(u)\mathcal{Q}, \tag{122}$$

$$J_{1,2} = \partial_u J_{1,1} = \partial_u[-\lambda u\varphi(u)] = -\lambda[\varphi - u^2\varphi] = \lambda(u^2-1)\mathcal{Q} = \lambda \mathrm{He}_2(u)\mathcal{Q}. \tag{123}$$

General pattern. $J_{m,j}$ is a polynomial in (u,λ) times $\mathcal{Q} = \varphi(u)$, with the polynomial expressible in terms of Hermite polynomials.

11.5 Cumulant Updates Under p_K

The general update $\Delta\kappa_j$ under the K-truncated Edgeworth expansion takes the form:

$$\Delta\kappa_j^{(K)} = \Delta\kappa_j^{(\text{Gauss})} + \sum_{\ell=3}^{K} \frac{\kappa_\ell}{\ell!} \cdot C_{j,\ell}(\mu, \sigma; \beta, \bar{\eta}, \Delta\eta), \tag{124}$$

where $\Delta\kappa_j^{(\text{Gauss})}$ is the Gaussian ($K = 2$) update derived earlier, and $C_{j,\ell}$ are correction coefficients involving the $J_{m,\ell}$ integrals.

$K = 1$ **(mean only).** Only $\kappa_1 = \mu$ is tracked. The variance and all higher cumulants are treated as fixed (or evolved by some auxiliary rule). This gives a 1D map for μ.

$K = 2$ **(Gaussian).** This is the case already derived: track (μ, σ^2), compute updates using I_0, I_1, I_2.

$K = 5$ **(includes skewness and kurtosis).** Track $(\kappa_1, \kappa_2, \kappa_3, \kappa_4, \kappa_5)$. The corrections involve $J_{m,3}, J_{m,4}, J_{m,5}$ for $m = 0, \ldots, 5$.

$K = 10$**.** Track 10 cumulants. Requires $J_{m,j}$ for $m, j = 0, \ldots, 10$.

11.6 Error Analysis

The truncation error at order K comes from:

1. Neglecting cumulants $\kappa_{K+1}, \kappa_{K+2}, \ldots$ in the distribution ansatz.
2. Cross-coupling: higher cumulants affect the dynamics of lower ones.

Expected scaling. For distributions approaching Gaussian (CLT regime), the standardized cumulants κ_j/σ^j decay as $t^{-(j-2)/2}$. Thus, the error from truncating at K should scale as:

$$\text{Error}(K, t) \sim O\big(\kappa_{K+1}/\sigma^{K+1}\big) \sim O\big(t^{-(K-1)/2}\big). \tag{125}$$

Higher K gives faster error decay, but requires tracking more cumulants and computing more $J_{m,j}$ integrals.