Exponential Families

Statistics · Module 2 — Statistical Models

Node 2.6

Position in Knowledge Graph

System:
Statistics

Structural Domain:
Statistical Models and Likelihood

Node:
2.6

Formal Definition

A statistical model belongs to the **exponential family** if its density can be written as:

$$f_\theta(x) = h(x) \exp\left(\eta(\theta)^\top T(x) - A(\theta)\right),$$

where:

- $T(x)$ is the sufficient statistic,
- $\eta(\theta)$ is the natural parameter,
- $A(\theta)$ is the log-partition function,
- $h(x)$ is the base measure.

Intuition

Exponential families separate:

- Data structure $T(x)$
- Parameter structure $\eta(\theta)$

All parameter dependence flows through a low-dimensional summary $T(x)$.

They are maximally structured statistical models.

Structural Role

Exponential families:

- Guarantee existence of sufficient statistics
- Simplify likelihood analysis
- Enable conjugate priors in Bayesian inference

- Underlie generalized linear models
- Connect to information geometry

They are the algebraic backbone of classical statistics.

Mathematical Development

Log-Likelihood Form

For i.i.d. data:

$$\ell(\theta) = \eta(\theta)^\top \sum_{i=1}^{n} T(x_i) - nA(\theta) + \sum_{i=1}^{n} \log h(x_i).$$

Parameter dependence appears only through:

$$\sum_{i=1}^{n} T(x_i).$$

This implies sufficiency.

Mean and Variance Structure

The log-partition function satisfies:

$$\nabla_\theta A(\theta) = \mathbb{E}_\theta[T(X)].$$

Second derivative:

$$\nabla_\theta^2 A(\theta) = \mathrm{Var}_\theta(T(X)).$$

Thus, curvature encodes variance.

Examples

Bernoulli

$$f_p(x) = \exp\left(x \log \frac{p}{1-p} + \log(1-p)\right).$$

Natural parameter:

$$\eta = \log \frac{p}{1-p}.$$

Normal (Known Variance)

$$f_\mu(x) = \exp\left(\frac{\mu x}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right) h(x).$$

Cross-Links

- **Linear Algebra:** Natural parameter space as vector space.
- **AI:** Softmax, logistic regression are exponential family models.
- **Economics:** Logit models in discrete choice.
- **Quantum:** Gibbs states resemble exponential family structure.

Structural Compression

Invariant:
Parameter enters linearly in sufficient statistic.

Mechanism:
Log-partition function controls normalization and variance.

Cross-System Echo:
In physics, Boltzmann distributions are exponential family models.

Exercises

1. Show Poisson distribution is exponential family.
2. Derive sufficient statistic for Normal mean.
3. Prove that expectation of sufficient statistic equals gradient of $A(\theta)$.
4. Explain why exponential families are closed under i.i.d. sampling.

Statistics · Module 2 — Statistical Models · Node 2.6
Concepts: exponential-families, sufficient-statistics, likelihood

hbar.university — own.your.cognition