

5.2: General Exponential Families

Basic Theory

Definition

We start with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as a model for a random experiment. So as usual, Ω is the set of outcomes, \mathcal{F} the σ -algebra of events, and \mathbb{P} the probability measure on the sample space (Ω, \mathcal{F}) . For the general formulation that we want in this section, we need two additional spaces, a measure space (S, \mathcal{S}, μ) (where the probability distributions will live) and a measurable space (T, \mathcal{T}) (serving the role of a *parameter space*). Typically, these spaces fall into our two standard categories. Specifically, the measure space is usually one of the following:

- *Discrete.* S is countable, \mathcal{S} is the collection of all subsets of S , and $\mu = \#$ is counting measure.
- *Euclidean.* S is a sufficiently nice Borel measurable subset of \mathbb{R}^n for some $n \in \mathbb{N}_+$, \mathcal{S} is the σ -algebra of Borel measurable subsets of S , and $\mu = \lambda_n$ is n -dimensional Lebesgue measure.

Similarly, the parameter space (T, \mathcal{T}) is usually either discrete, so that T is countable and \mathcal{T} the collection of all subsets of T , or Euclidean so that T is a sufficiently nice Borel measurable subset of \mathbb{R}^m for some $m \in \mathbb{N}_+$ and \mathcal{T} is the σ -algebra of Borel measurable subsets of T .

Suppose now that X is random variable defined on the probability space, taking values in S , and that the distribution of X depends on a parameter $\theta \in T$. For $\theta \in T$ we assume that the distribution of X has probability density function f_θ with respect to μ .

for $k \in \mathbb{N}_+$, the family of distributions of X is a k -parameter *exponential family* if

$$f_\theta(x) = \alpha(\theta) g(x) \exp\left(\sum_{i=1}^k \beta_i(\theta) h_i(x)\right); \quad x \in S, \theta \in T \quad (5.2.1)$$

where α and $(\beta_1, \beta_2, \dots, \beta_k)$ are measurable functions from T into \mathbb{R} , and where g and (h_1, h_2, \dots, h_k) are measurable functions from S into \mathbb{R} . Moreover, k is assumed to be the smallest such integer.

1. The parameters $(\beta_1(\theta), \beta_2(\theta), \dots, \beta_k(\theta))$ are called the *natural parameters* of the distribution.
2. the random variables $(h_1(X), h_2(X), \dots, h_k(X))$ are called the *natural statistics* of the distribution.

Although the definition may look intimidating, exponential families are useful because many important theoretical results in statistics hold for exponential families, and because many special parametric families of distributions turn out to be exponential families. It's important to emphasize that the representation of $f_\theta(x)$ given in the definition must hold for all $x \in S$ and $\theta \in T$. If the representation only holds for a set of $x \in S$ that depends on the particular $\theta \in T$, then the family of distributions is not a general exponential family.

The next result shows that if we sample from the distribution of an exponential family, then the distribution of the random sample is itself an exponential family with the same natural parameters.

Suppose that the distribution of random variable X is a k -parameter exponential family with natural parameters $(\beta_1(\theta), \beta_2(\theta), \dots, \beta_k(\theta))$, and natural statistics $(h_1(X), h_2(X), \dots, h_k(X))$. Let $\mathbf{X} = (X_1, X_2, \dots, X_n)$ be a sequence of n independent random variables, each with the same distribution as X . Then \mathbf{X} is a k -parameter exponential family with natural parameters $(\beta_1(\theta), \beta_2(\theta), \dots, \beta_k(\theta))$, and natural statistics

$$u_j(\mathbf{X}) = \sum_{i=1}^n h_j(X_i), \quad j \in \{1, 2, \dots, k\} \quad (5.2.2)$$

Proof

Let f_θ denote the PDF of X corresponding to the parameter value $\theta \in T$, so that $f_\theta(x)$ has the representation given in the [definition](#) for $x \in S$ and $\theta \in T$. Then for $\theta \in T$, $\mathbf{X} = (X_1, X_2, \dots, X_n)$ has PDF g_θ given by

$$g_\theta(x_1, x_2, \dots, x_n) = f_\theta(x_1) f_\theta(x_2) \cdots f_\theta(x_n), \quad (x_1, x_2, \dots, x_n) \in S^n \quad (5.2.3)$$

Substituting and simplifying gives the result.

Examples and Special Cases

Special Distributions

Many of the special distributions studied in this chapter are general exponential families, at least with respect to some of their parameters. On the other hand, most commonly, a parametric family fails to be a general exponential family because the support set depends on the parameter. The following theorems give a number of examples. Proofs will be provided in the individual sections.

The Bernoulli distribution is a one parameter exponential family in the success parameter $p \in [0, 1]$

The beta distribution is a two-parameter exponential family in the shape parameters $a \in (0, \infty)$, $b \in (0, \infty)$.

The beta prime distribution is a two-parameter exponential family in the shape parameters $a \in (0, \infty)$, $b \in (0, \infty)$.

The binomial distribution is a one-parameter exponential family in the success parameter $p \in [0, 1]$ for a fixed value of the trial parameter $n \in \mathbb{N}_+$.

The chi-square distribution is a one-parameter exponential family in the degrees of freedom $n \in (0, \infty)$.

The exponential distribution is a one-parameter exponential family (appropriately enough), in the rate parameter $r \in (0, \infty)$.

The gamma distribution is a two-parameter exponential family in the shape parameter $k \in (0, \infty)$ and the scale parameter $b \in (0, \infty)$.

The geometric distribution is a one-parameter exponential family in the success probability $p \in (0, 1)$.

The half normal distribution is a one-parameter exponential family in the scale parameter $\sigma \in (0, \infty)$

The Laplace distribution is a one-parameter exponential family in the scale parameter $b \in (0, \infty)$ for a fixed value of the location parameter $a \in \mathbb{R}$.

The Lévy distribution is a one-parameter exponential family in the scale parameter $b \in (0, \infty)$ for a fixed value of the location parameter $a \in \mathbb{R}$.

The logarithmic distribution is a one-parameter exponential family in the shape parameter $p \in (0, 1)$

The lognormal distribution is a two parameter exponential family in the shape parameters $\mu \in \mathbb{R}$, $\sigma \in (0, \infty)$.

The Maxwell distribution is a one-parameter exponential family in the scale parameter $b \in (0, \infty)$.

The k -dimensional multinomial distribution is a k -parameter exponential family in the probability parameters (p_1, p_2, \dots, p_k) for a fixed value of the trial parameter $n \in \mathbb{N}_+$.

The k -dimensional multivariate normal distribution is a $\frac{1}{2}(k^2 + 3k)$ -parameter exponential family with respect to the mean vector $\boldsymbol{\mu}$ and the variance-covariance matrix \mathbf{V} .

The negative binomial distribution is a one-parameter exponential family in the success parameter $p \in (0, 1)$ for a fixed value of the stopping parameter $k \in \mathbb{N}_+$.

The normal distribution is a two-parameter exponential family in the mean $\mu \in \mathbb{R}$ and the standard deviation $\sigma \in (0, \infty)$.

The Pareto distribution is a one-parameter exponential family in the shape parameter for a fixed value of the scale parameter.

The Poisson distribution is a one-parameter exponential family.

The Rayleigh distribution is a one-parameter exponential family.

The U-power distribution is a one-parameter exponential family in the shape parameter, for fixed values of the location and scale parameters.

The Weibull distribution is a one-parameter exponential family in the scale parameter for a fixed value of the shape parameter.

The zeta distribution is a one-parameter exponential family.

The Wald distribution is a two-parameter exponential family.

This page titled [5.2: General Exponential Families](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist](#) ([Random Services](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.