

Exponential Families

Last modified: 2026-10-27 23:24:14 (PDT)

Many of the distributions in this course, such as the [Bernoulli](#), [Poisson](#) and [Normal](#) distributions, share one algebraic form for their [probability mass function](#) or [density](#). Distributions with that form are called an [exponential family](#), and they share properties that make estimation easier. This page defines the form and its [natural parameter](#), and checks the form for the [Poisson distribution](#) on data.

The exponential family form

Definition 0.1 (Exponential family). Let X be a random variable whose distribution depends on one parameter η , and let $f(x; \eta)$ be its [probability mass function](#) or [density](#). The distributions of X form an **exponential family** if there are a nonnegative [function](#) h and a function T of x alone, and a function A of η alone, such that

$$f(x; \eta) = h(x) \exp\{\eta T(x) - A(\eta)\} \quad (1)$$

for every value x and every allowed η .

i Remark (Source: Exponential family)

Remark 0.1 (Source: Exponential family). Casella and Berger [1, Sec. 3.4], Eq. 3.4.1 write the family as a sum of k terms inside the exponential, each a [function](#) of the parameter times a function of x ; Dobson and Barnett [2, Sec. 3.2], Eq. 3.3 write it in the form $\exp\{a(y)b(\theta) + c(\theta) + d(y)\}$. Equation 1 is the case of one term. It is Dobson and Barnett's canonical form when $T(x) = x$, and the [natural parameter](#) η plays the role of their $b(\theta)$. See the [notation choices](#) for how the symbols line up.

Example 0.1 (The Poisson distribution is an exponential family). Let $X \sim \text{Pois}(\lambda)$, the [Poisson distribution](#) with rate $\lambda > 0$. Its [probability mass function](#) has one factor free of

λ , and the factors λ^x and $\exp\{-\lambda\}$. We rewrite it in the form of Equation 1 by writing λ^x as an exponential of a logarithm:

$$\begin{aligned}
 f(x; \lambda) &= \frac{\lambda^x \exp\{-\lambda\}}{x!} && \text{(Poisson mass function)} \\
 &= \frac{1}{x!} \exp\{\log\{\lambda^x\}\} \exp\{-\lambda\} && \text{(a positive number is the exponential of its logarithm)} \\
 &= \frac{1}{x!} \exp\{x \log\{\lambda\}\} \exp\{-\lambda\} && \text{(the logarithm of a power is the exponent times the logarithm)} \\
 &= \frac{1}{x!} \exp\{x \log\{\lambda\} - \lambda\} && \text{(the product of exponentials is the exponential of the sum)}
 \end{aligned}$$

So the Poisson distributions are an **exponential family** with $h(x) = 1/x!$, $T(x) = x$, **natural parameter** $\eta = \log\{\lambda\}$, and $A(\eta) = \lambda = \exp\{\eta\}$.

To check the algebra on data, we use the number of tropical cyclones in northeastern Australia in each of 13 seasons [2, Sec. 1.6.5], with the rate set to the **sample mean**. We evaluate both sides of Equation 1 at every observed count.

R

```

cyclones <- c(6, 5, 4, 6, 6, 3, 12, 7, 4, 2, 6, 7, 4)
rate <- mean(cyclones)

base_measure <- function(x) 1 / factorial(x)
natural_parameter <- log(rate)
log_partition <- function(eta) exp(eta)

exponential_form <- base_measure(cyclones) *
  exp(natural_parameter * cyclones - log_partition(natural_parameter))

data.frame(
  count = cyclones,
  poisson = dpois(cyclones, rate),
  exponential_form = exponential_form
)
#>      count    poisson exponential_form
#> 1      6 0.15764459      0.15764459
#> 2      5 0.17078164      0.17078164
#> 3      4 0.15417787      0.15417787
#> 4      6 0.15764459      0.15764459
#> 5      6 0.15764459      0.15764459
#> 6      3 0.11135069      0.11135069
#> 7     12 0.00683927      0.00683927
#> 8      7 0.12472979      0.12472979
#> 9      4 0.15417787      0.15417787
#> 10     2 0.06031496      0.06031496
#> 11     6 0.15764459      0.15764459
#> 12     7 0.12472979      0.12472979
#> 13     4 0.15417787      0.15417787
max(abs(dpois(cyclones, rate) - exponential_form))
#> [1] 1.11022e-16

```

Python

```

import numpy as np
import pandas as pd
from scipy.special import factorial
from scipy.stats import poisson

cyclones = np.array([6, 5, 4, 6, 6, 3, 12, 7, 4, 2, 6, 7, 4])
rate = cyclones.mean()

natural_parameter = np.log(rate)
exponential_form = (1 / factorial(cyclones)) * np.exp(
    natural_parameter * cyclones - np.exp(natural_parameter)
)

table = pd.DataFrame(
    {
        "count": cyclones,
        "poisson": poisson.pmf(cyclones, rate),
        "exponential_form": exponential_form,
    }
)
print(table)
#>      count  poisson  exponential_form
#> 0         6  0.157645          0.157645
#> 1         5  0.170782          0.170782
#> 2         4  0.154178          0.154178
#> 3         6  0.157645          0.157645
#> 4         6  0.157645          0.157645
#> 5         3  0.111351          0.111351
#> 6        12  0.006839          0.006839
#> 7         7  0.124730          0.124730
#> 8         4  0.154178          0.154178
#> 9         2  0.060315          0.060315
#> 10        6  0.157645          0.157645
#> 11        7  0.124730          0.124730
#> 12        4  0.154178          0.154178
print(np.max(np.abs(poisson.pmf(cyclones, rate) - exponential_form)))
#> 5.551115123125783e-17

```

The two columns agree to within rounding error: the largest difference is 1.1e-16.

Definition 0.2 (Natural parameter). In the form of Equation 1 of an **exponential family**, the parameter η that multiplies the **statistic** $T(x)$ inside the exponential is the **natural parameter**, also called the **canonical parameter** [2, Sec. 3.2].

Example 0.2 (The natural parameter of the Poisson distribution is the log rate). In Example 0.1, the **Poisson** rate λ and the **natural parameter** η are related by

$$\eta = \log\{\lambda\} \quad \text{and} \quad \lambda = \exp\{\eta\}$$

For the cyclone data, the rate set to the **sample mean** is 5.538 cyclones per season, so the natural parameter is $\eta = 1.712$. The natural parameter can be any real number, while the rate must be greater than zero.

Further reading

- The **Fall 2024 course site** for UC Berkeley's Stat 210A, *Theoretical Statistics* (Will Fithian) has a course reader on exponential families, sufficiency, completeness and **unbiased** estimation.

References

- [1] G. Casella and R. Berger, *Statistical inference*, 2nd ed. Cengage Learning, 2002. Available: <https://www.cengage.com/c/statistical-inference-2e-casella-berger/9780534243128/>
- [2] A. J. Dobson and A. G. Barnett, *An introduction to generalized linear models*, 4th ed. CRC press, 2018. doi: [10.1201/9781315182780](https://doi.org/10.1201/9781315182780).