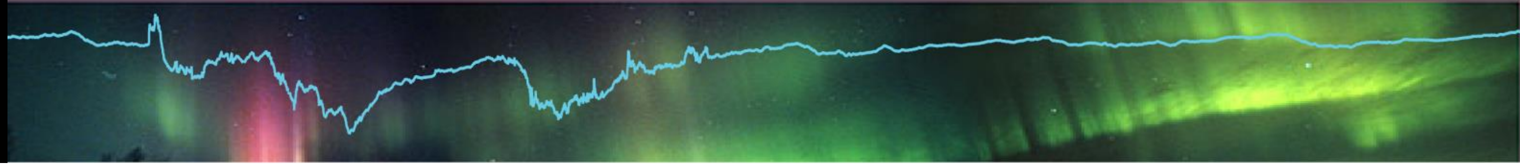
USGS
science for a changing world



Uncertainties in estimates of the occurrence rate of rare space weather events

Jeffrey J. Love

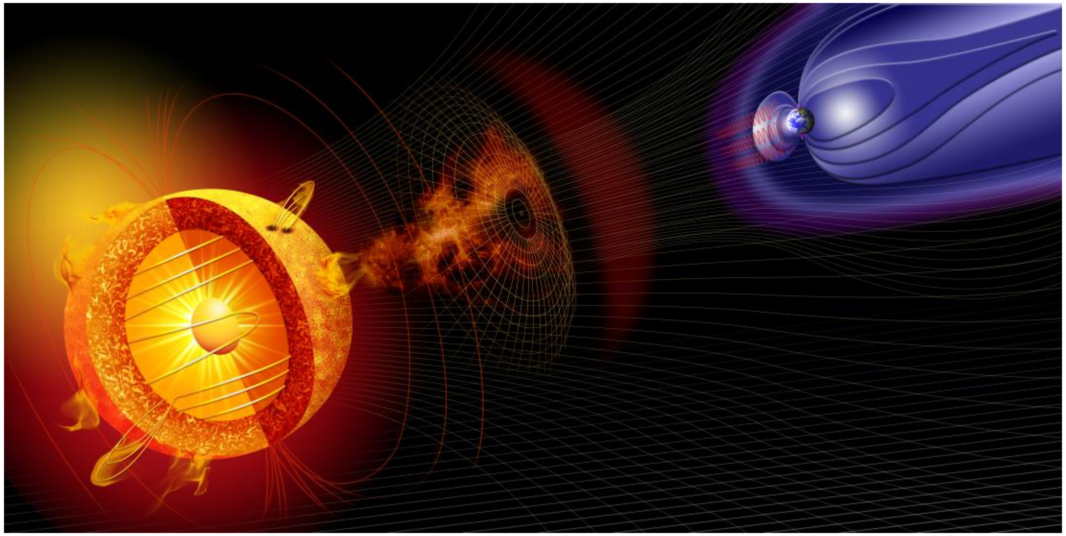
Geomagnetism Program

USGS Natural Hazards Mission

jlove@usgs.gov

Love, J. J., 2012. Credible occurrence probabilities for extreme geophysical events: Earthquakes, volcanic eruptions, magnetic storms, Geophys. Res. Lett., 39, L10301, doi:10.1029/2012GL051431.

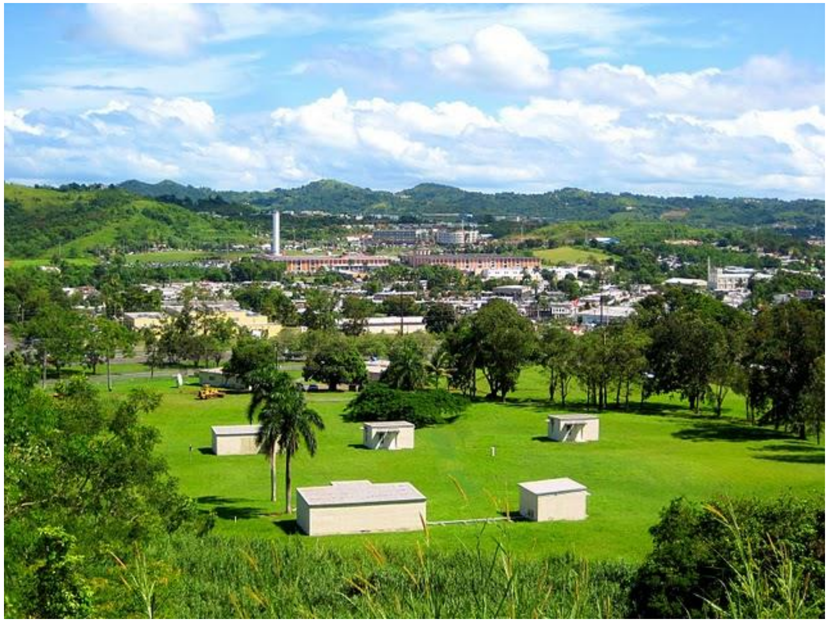
Rare-event analysis is important for lots of geophysical hazards



Magnetic observatory data

- Magnetic storms were discovered in the 19th century through ground-based measurement of geomagnetic activity.
- The only complete record of the 1859 Carrington event was collected through visual measurement at the Colaba magnetic observatory in India.
- Standard measures of magnetic storm intensity (Dst, AE, Kp, aa, etc.) are all derived from modern magnetic observatory data.
- In this study, we will count magnetic storm events defined by a maximum –Dst that exceeds a chosen threshold.

Colaba observatory in Bombay India



USGS San Juan Puerto Rico observatory

Frequentist and Bayesian Statistical Analyses

Frequentist analysis (Classical): Model parameters are deterministic, proposed on the basis of independent theories, parameter values are subjected to tests of statistical significance by comparing model predictions against a null hypothesis of randomness.

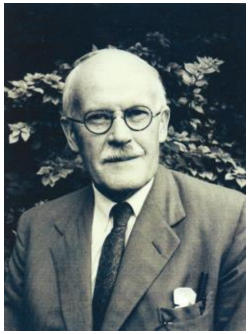
Ronald Fisher (1890--1962)

Bayesian analysis: Model parameters are treated as an idealized statistical realization of probability, inference of parameter values is made on the basis of data and prior beliefs. The Jeffreys prior preserves posterior probability under reparameterization.

Thomas Bayes (1701-1761)

Pierre-Simon Laplace (1749-1827)

Harold Jeffreys (1891--1989)



Frequentist and Bayesian Statistical Analyses

Frequentist analysis (Classical): Model parameters are deterministic, proposed on the basis of independent theories, parameter values are subjected to tests of statistical significance by comparing model predictions against a null hypothesis of randomness.

Ronald Fisher (1890--1962)

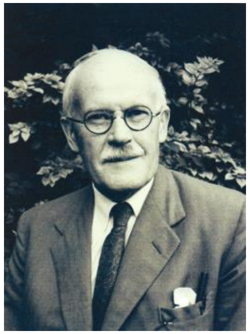
Maximum Likelihood: Model parameters are "deterministic", but they are estimated from a data set.

Bayesian analysis: Model parameters are treated as an idealized statistical realization of probability, inference of parameter values is made on the basis of data and prior beliefs. The Jeffreys prior preserves posterior probability under reparameterization.

Thomas Bayes (1701-1761)

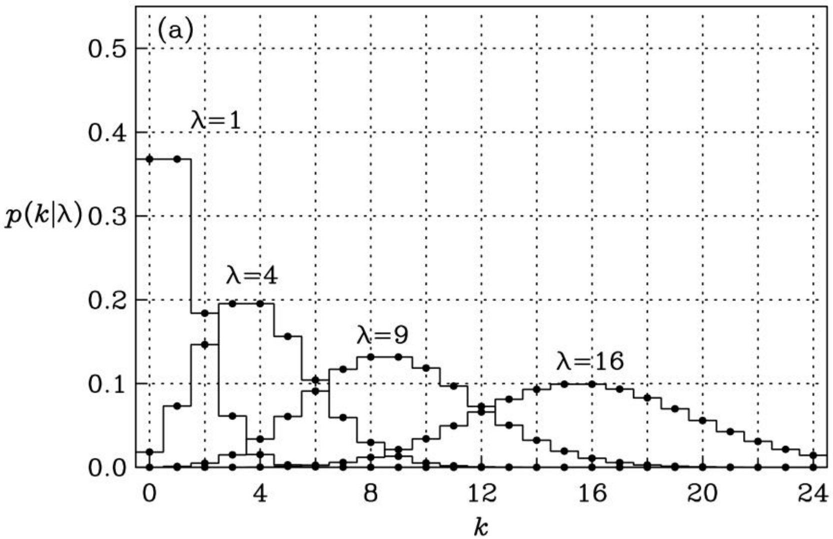
Pierre-Simon Laplace (1749-1827)

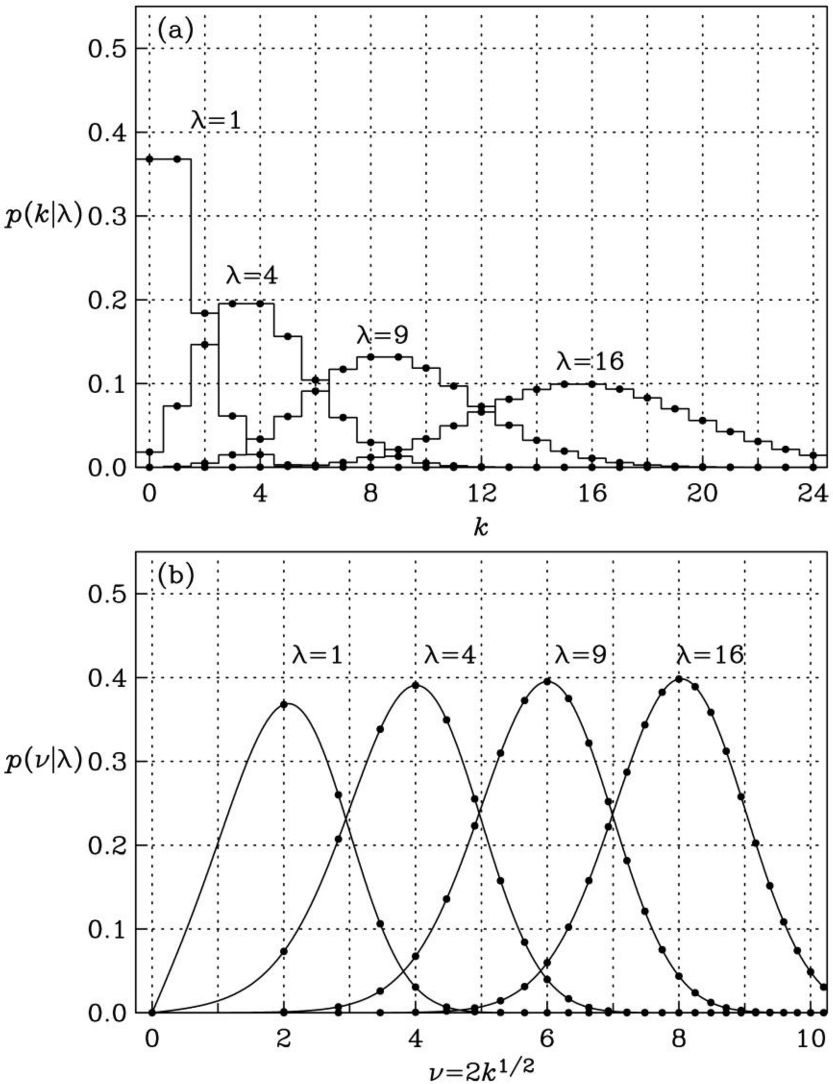
Harold Jeffreys (1891--1989)



Poisson probability of k events in time τ for a occurrence rate ρ

$$p(k|\lambda) = \frac{\lambda^k}{k!} \exp(-\lambda) \quad \text{where} \quad \lambda = \rho\tau$$





Poisson probability of k events in time τ for a occurrence rate ρ

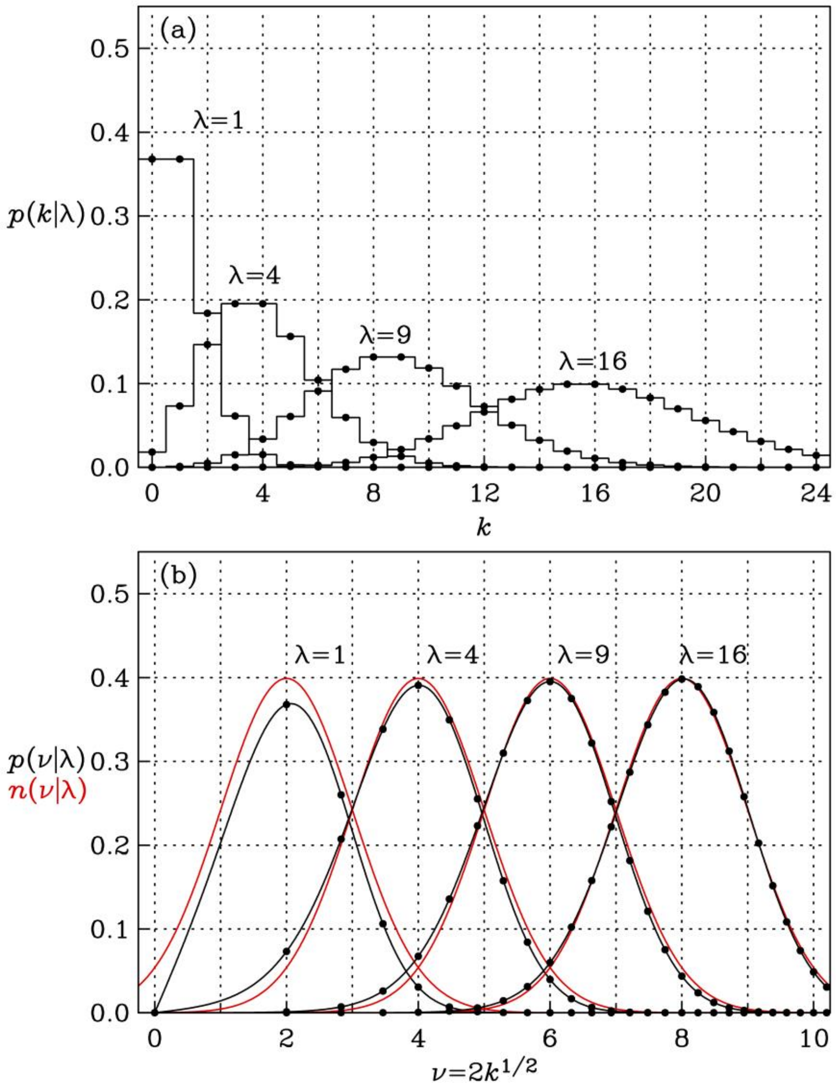
$$p(k|\lambda) = \frac{\lambda^k}{k!}\exp(-\lambda) \quad \text{where} \quad \lambda = \rho\tau$$

Make "variance stabilizing" transformation $k \to \nu$, such that

$$\nu = 2\sqrt{k} \quad \text{and} \quad k = \tfrac{1}{4}\nu^2 \quad \text{and}$$

conserving unit quantities of probability by changing variables

$$p(\nu|\lambda) = p(k(\nu)|\lambda) \cdot \left|\frac{\partial k}{\partial \nu}\right|$$



Poisson probability of k events in time τ for a occurrence rate ρ

$$p(k|\lambda) = \frac{\lambda^k}{k!}\exp(-\lambda) \quad \text{where} \quad \lambda = \rho\tau$$

Make "variance stabilizing" transformation $k \to \nu$, such that

$$\nu = 2\sqrt{k} \quad \text{and} \quad k = \tfrac{1}{4}\nu^2 \quad \text{and}$$

conserving unit quantities of probability by changing variables

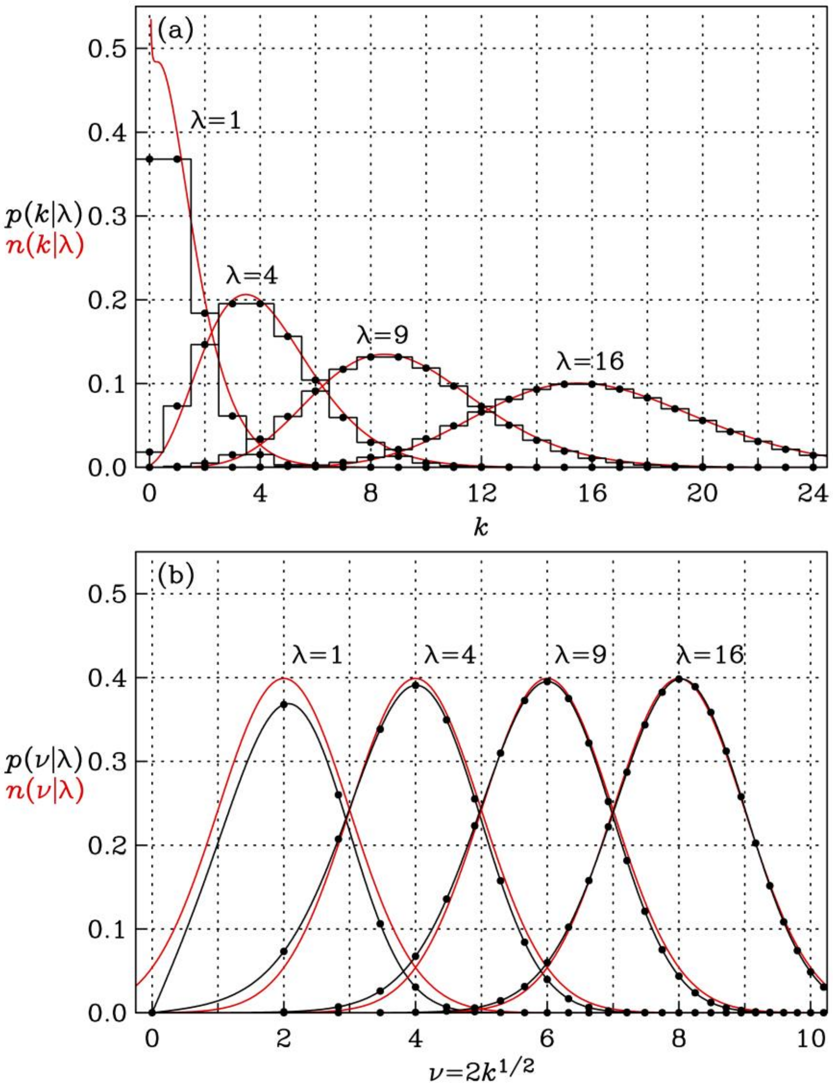
$$p(\nu|\lambda) = p(k(\nu)|\lambda) \cdot \left|\frac{\partial k}{\partial \nu}\right|$$

The resulting distribution $p(\nu|\lambda)$ is approximately Normal

$$n(\nu|\mu) = \frac{1}{\sqrt{2\pi}}\exp\left[-\tfrac{1}{2}(\nu-\mu)^2\right] \quad \text{where} \quad \sigma^2 = 1$$

From this we obtain maximum likelihood and confidence intervals

$$\mu_{ML} = \nu \quad \text{and} \quad C_z(\mu|\nu) = [(\nu - z)\,,(\nu + z)]$$



Poisson probability of k events in time τ for a occurrence rate ρ

$$p(k|\lambda) = \frac{\lambda^k}{k!}\exp(-\lambda) \quad \text{where} \quad \lambda = \rho\tau$$

Make "variance stabilizing" transformation $k \to \nu$, such that

$$\nu = 2\sqrt{k} \quad \text{and} \quad k = \tfrac{1}{4}\nu^2 \quad \text{and}$$

conserving unit quantities of probability by changing variables

$$p(\nu|\lambda) = p(k(\nu)|\lambda) \cdot \left|\frac{\partial k}{\partial \nu}\right|$$

The resulting distribution $p(\nu|\lambda)$ is approximately Normal

$$n(\nu|\mu) = \frac{1}{\sqrt{2\pi}}\exp\left[-\tfrac{1}{2}(\nu-\mu)^2\right] \quad \text{where} \quad \sigma^2 = 1$$

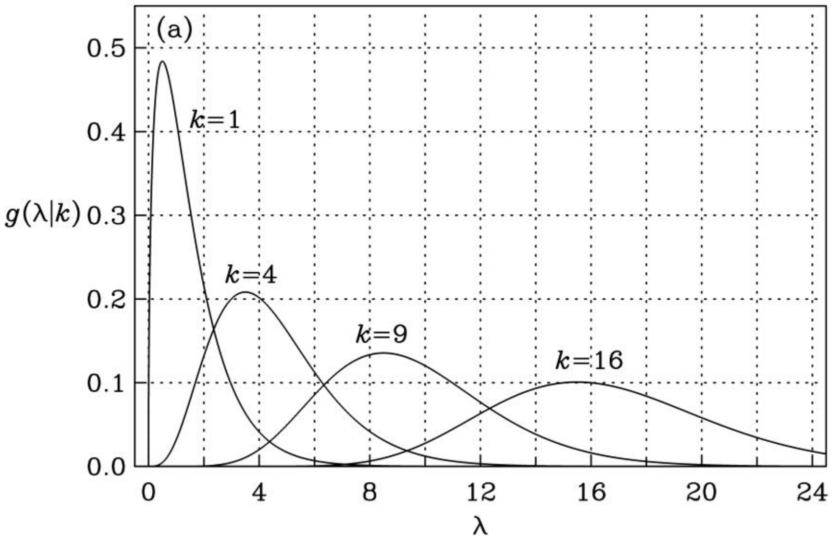
From this we obtain maximum likelihood and confidence intervals

$$\mu_{ML} = \nu \quad \text{and} \quad C_z(\mu|\nu) = [(\nu - z), (\nu + z)]$$

Upon inverse transformation, we obtain these in their usual form

$$\lambda_{ML} = k \quad \text{and} \quad C_z(\lambda|k) = \left[\left(\sqrt{k} - \tfrac{1}{2}z\right)^2, \left(\sqrt{k} + \tfrac{1}{2}z\right)^2\right]$$

where z specifies the number of standard deviations. Under standard interpretations, the confidence interval is difficult to interpret if all we have are k events for one duration of time τ.

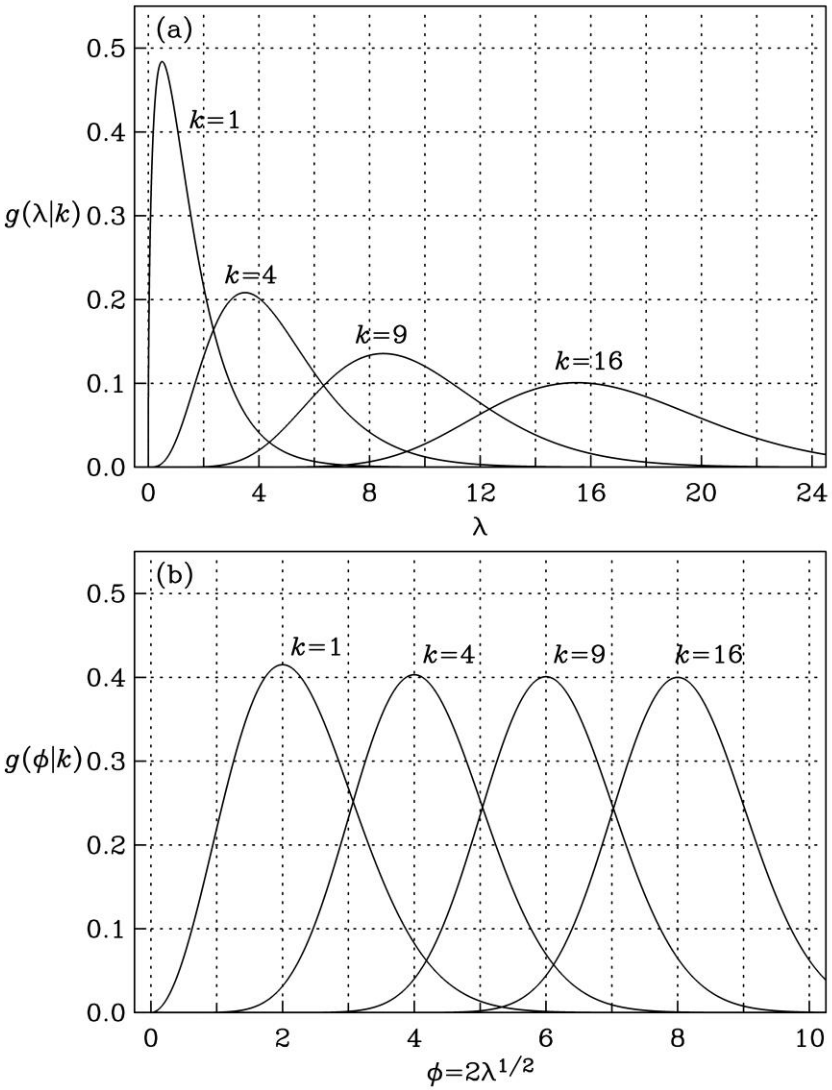


The relationship between data k and rate parameter λ

$$g(\lambda|k) \propto p(k|\lambda) \times \pi(\lambda)$$

Many priors $\pi(\lambda)$ can be considered. The Jeffreys prior ensures conservation of probability under changes of variables. For Poisson

$$\pi(\lambda) \propto \frac{1}{\sqrt{\lambda}}$$



The relationship between data k and rate parameter λ

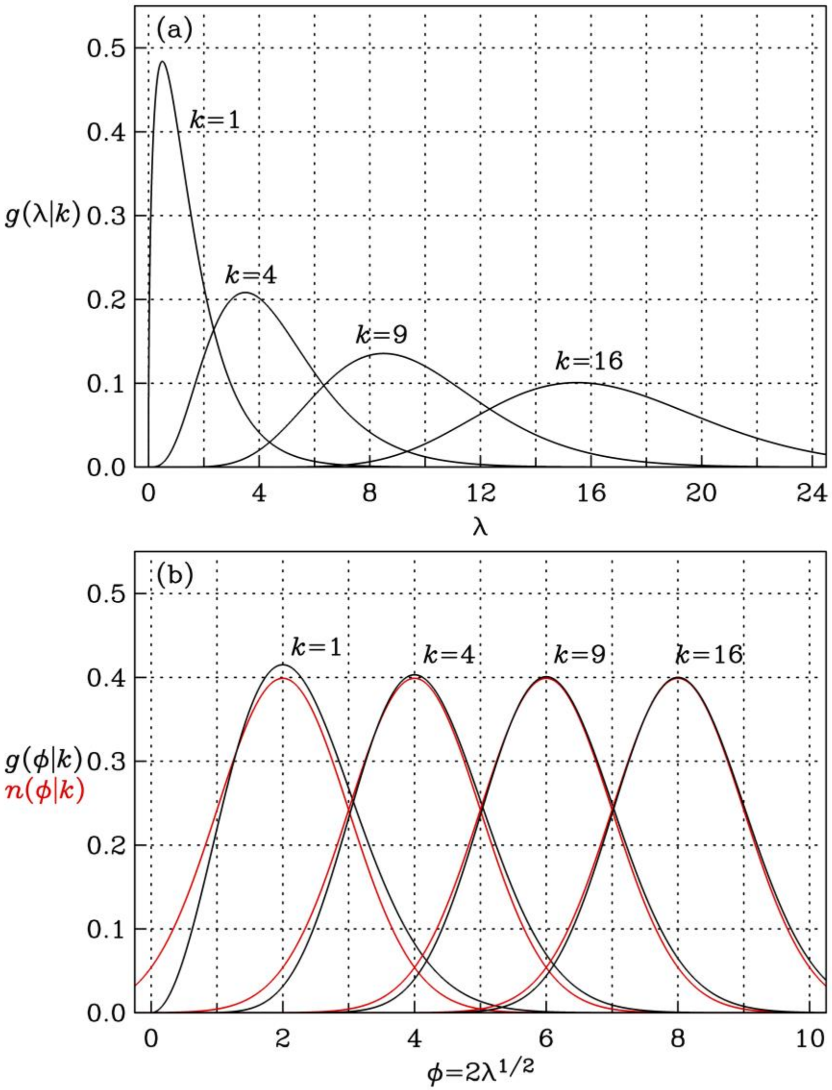
$$g(\lambda|k) \propto p(k|\lambda) \times \pi(\lambda)$$

Many priors $\pi(\lambda)$ can be considered. The Jeffreys prior ensures conservation of probability under changes of variables. For Poisson

$$\pi(\lambda) \propto \frac{1}{\sqrt{\lambda}}$$

Make "variance stabilizing" transformation $\lambda \to \phi$, such that

$$\phi = 2\sqrt{\lambda} \ \text{ and } \ \lambda = \tfrac{1}{4}\phi^2$$



The relationship between data k and rate parameter λ

$$g(\lambda|k) \propto p(k|\lambda) \times \pi(\lambda)$$

Many priors $\pi(\lambda)$ can be considered. The Jeffreys prior ensures conservation of probability under changes of variables. For Poisson

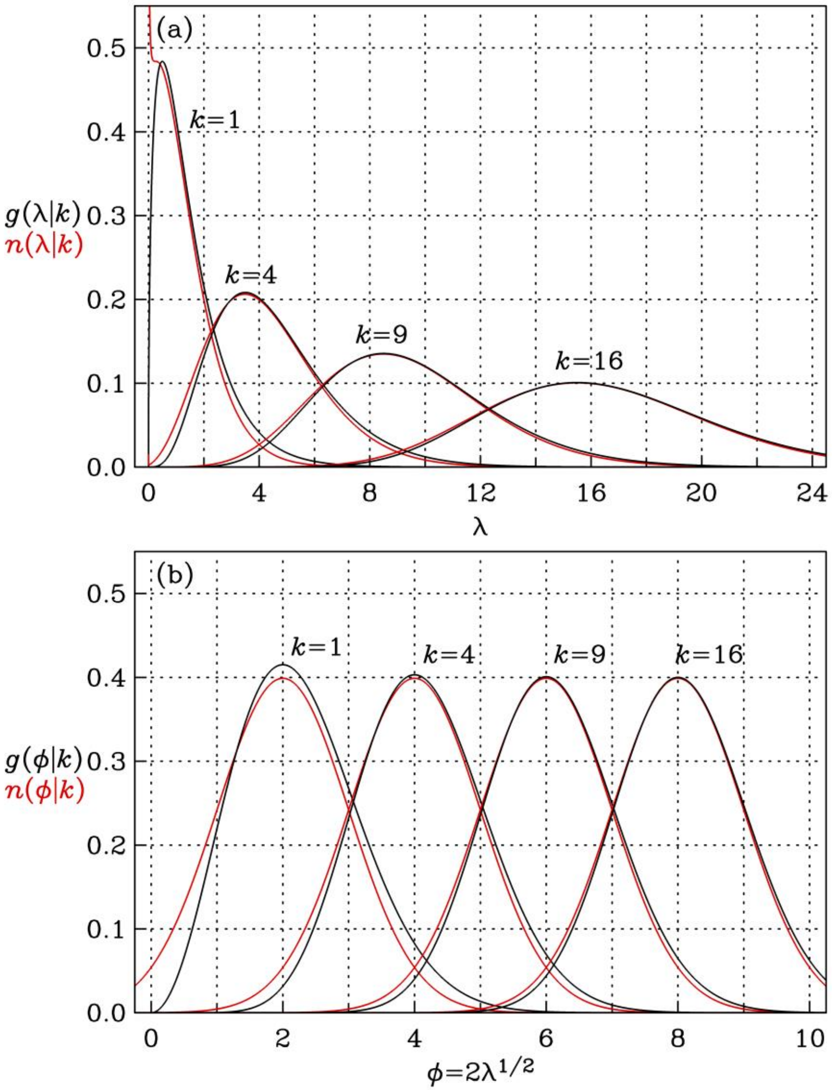
$$\pi(\lambda) \propto \frac{1}{\sqrt{\lambda}}$$

Make "variance stabilizing" transformation $\lambda \to \phi$, such that

$$\phi = 2\sqrt{\lambda} \quad \text{and} \quad \lambda = \tfrac{1}{4}\phi^2$$

This yields a distribution $g(\phi|\mu)$ that is approximately Normal

$$n\,(\phi|\mu) = \frac{1}{\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}(\phi-\mu)^2\right] \quad \text{where} \quad \sigma^2 = 1$$



The relationship between data k and rate parameter λ

$$g(\lambda|k) \propto p(k|\lambda) \times \pi(\lambda)$$

Many priors $\pi(\lambda)$ can be considered. The Jeffreys prior ensures conservation of probability under changes of variables. For Poisson

$$\pi(\lambda) \propto \frac{1}{\sqrt{\lambda}}$$

Make "variance stabilizing" transformation $\lambda \to \phi$, such that

$$\phi = 2\sqrt{\lambda} \quad \text{and} \quad \lambda = \tfrac{1}{4}\phi^2$$

This yields a distribution $g(\phi|\mu)$ that is approximately Normal

$$n\,(\phi|\mu) = \frac{1}{\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}(\phi-\mu)^2\right] \quad \text{where} \quad \sigma^2 = 1$$

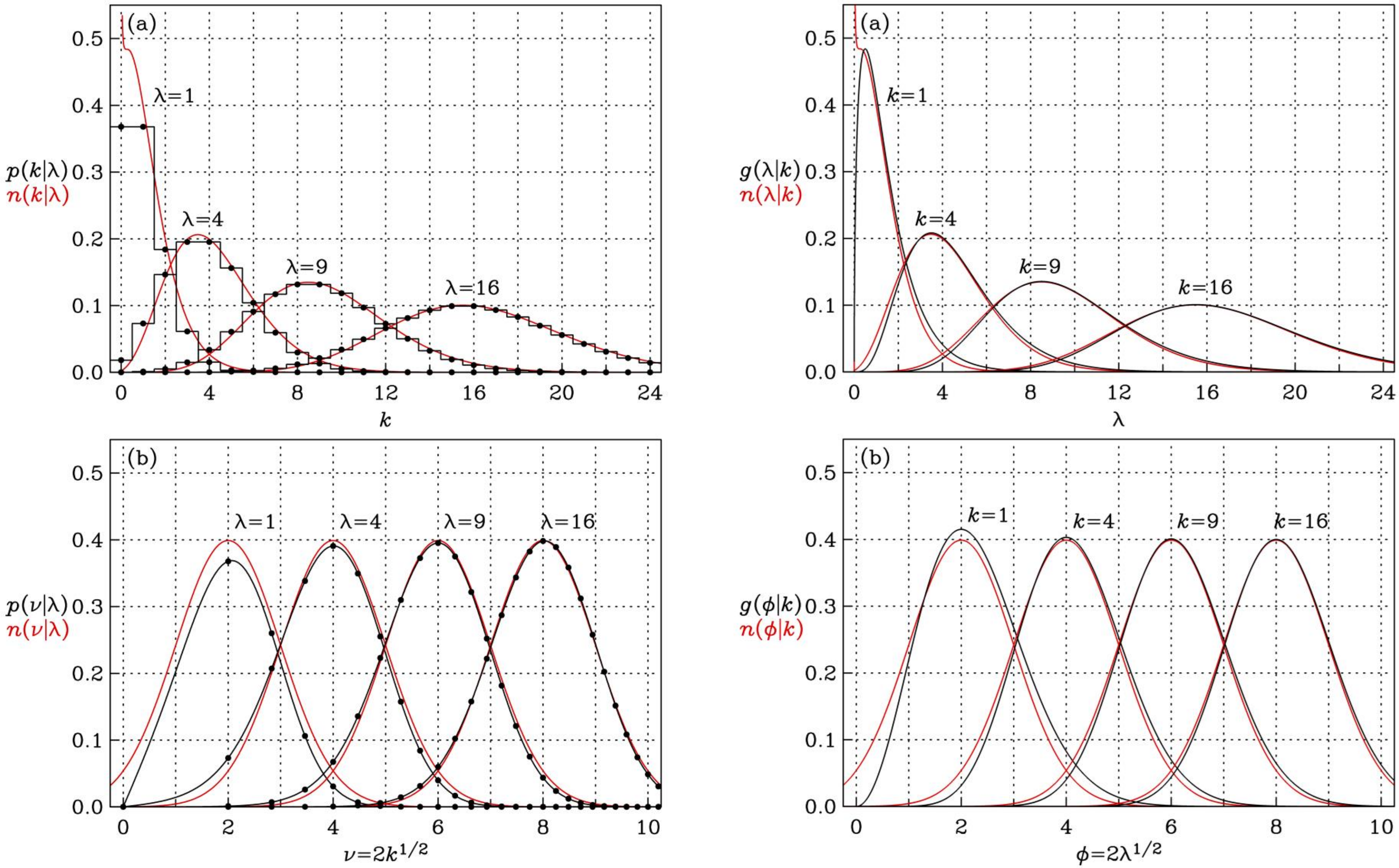
Upon inverse transformation, the maximum likelihood and credible intervals are

$$\lambda_{ML} = k \quad \text{and} \quad C_z(\lambda|k) = \left[\left(\sqrt{k} - \tfrac{1}{2}z\right)^2, \left(\sqrt{k} + \tfrac{1}{2}z\right)^2\right]$$

Same as the mathematical approximations obtained via frequentist interpretation, but more amenable to interpretation by scientists!

Frequentist

Bayesian with Jeffreys Prior



- Three magnetic storm events with –Dst ≥ 589 nT since 1859

  Average occurrence rate: **1.961** /century
  95% credibility interval: [**0.350, 4.878**]/century

  Most likely 10-yr occurrence probability (N ≥ 1): **0.178**
  95% credibility interval: [**0.034, 0.386**]

- One magnetic storm event with –Dst ≥ 1760 nT since 1859 (Carrington)

  Average occurrence rate: **0.654** /century
  95% credibility interval: [**0.163, 1.471**]/century

  Most likely 10-yr occurrence probability (N ≥ 1): **0.063**
  95% credibility interval: [<**0.001, 0.230**]

- **Riley**: 10-yr occurrence probability for Carrington event: **0.120**

How might statistical estimations be improved?

- Old magnetic observatory data sets might reveal other extreme events measured in terms of Dst or related absolute measures of disturbance.
- Satellite (SOHO, STEREO, etc.) monitoring and counting of extreme coronal mass ejections that haven't arrived at Earth:

  Baker, D. N. et al., A major solar eruptive event in July 2012: Defining extreme space weather scenarios, Space Weather, 2013.

  Ngwira, C., et al. Simulation of the 23 July 2012 extreme space weather event: What if this extremely rare CME was Earth directed?, Space Weather, 2013.

- Statistical models need to be improved, perhaps by combining the Bayesian Poisson models (presented here) with the power-law models:

  Love, J. J., Credible occurrence probabilities for extreme geophysical events: Earthquakes, volcanic eruptions, magnetic storms, Geophys. Res. Lett., 2012.

  Riley, P., On the probability of occurrence of extreme space weather events, Space Weather, 2012.

- Ice core records of transient enhancements in nitrate, proposed to be related to proton events (not magnetic storms), now appear to be unreliable:

  Wolff, E. W. et al., The Carrington event not observed in most ice core nitrate records, Geophys. Res. Lett., 2012.

It is the mark of an educated mind to rest satisfied with the degree of precision which the nature of the subject admits and not to seek exactness where only an approximation is possible. *Aristotle (384 -322 BC)*