Statistical Forecasting using Poisson Regression

A few topics

- Parameter fitting
- Maximum likelihood method
- Poisson distribution
- Statistical Forecasting using Poisson Regression

Parameter Fitting

- What is parameter fitting?
  - Determine the distribution parameters using sample statistics
  - The mathematical form of a parametric distribution amounts to idealizations of real data by an abstract distribution model
- Why do we need parameter fitting, or why do we need to fit real data into a parametric distribution?
  - Compactness: allow us to characterize properties of the data with a small number of parameters
  - Smoothing and interpolation: Real data are subject to sampling variations that lead to gaps in the empirical distributions, which can be estimated using parametric distributions
  - Extrapolation: allow us to estimate probabilities for events outside the range of sample data

Wilks 2011: section 4.1.1

Likelihood function vs. probability distribution function

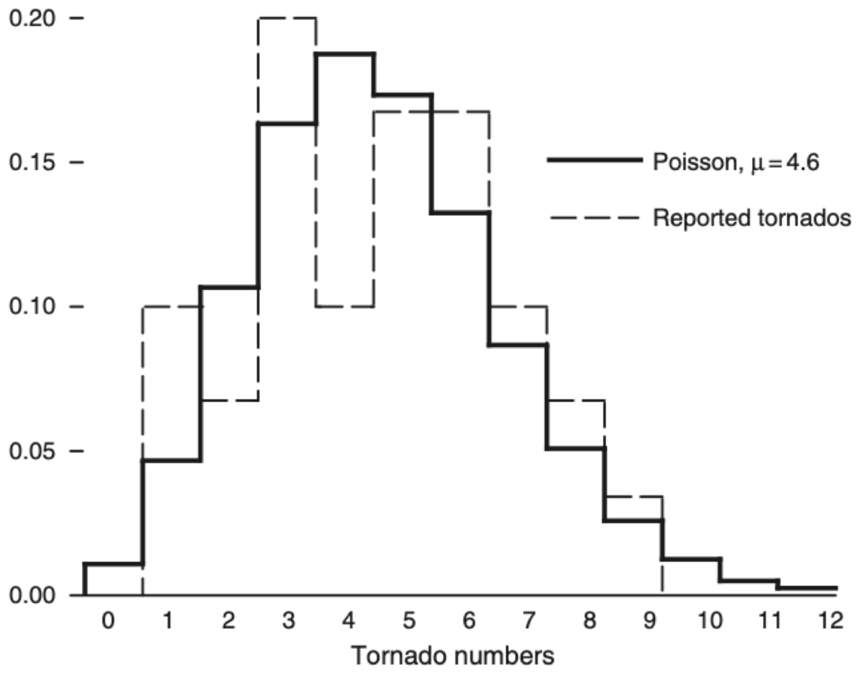


FIGURE 4.2 Histogram of number of tornados reported annually in New York state for 1959–1988 (dashed), and fitted Poisson distribution with $\mu = 4.6$ tornados/year (solid).

- Probability distribution function (PDF): the PDF is a function of the data for fixed values of the parameters.
- Likelihood function: the likelihood function is the distribution probability determined from observational data and is a function of the unknown parameters for fixed values of the already observed data

Wilks 2011: section 4.6

Parameter Fitting Using Maximum Likelihood

- The Maximum Likelihood method seeks to find values of the distribution parameters that maximize the likelihood function.
- For example, the likelihood function for the Gaussian parameters μ and σ, given a sample of n observations, xi, i = 1,… n, is

$$\Lambda(\mu, \sigma) = \sigma^{-n}(\sqrt{2\pi})^{-n} \prod_{i=1}^{n} \exp\left[-\frac{(x_i - \mu)^2}{2\sigma^2}\right]. \tag{4.67}$$

- The equation looks exactly the same as the joint PDF for n independent Gaussian variables, except that the parameters μ and σ are the variables, and x_i denotes fixed constants.
- Usually it is more convenient to work with the logarithm of the likelihood function, known as the log-likelihood.

$$L(\mu, \sigma) = \ln[\Lambda(\mu, \sigma)] = -n\ln(\sigma) - n\ln(\sqrt{2\pi}) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2, \tag{4.68}$$

Wilks 2011: section 4.6

Parameter Fitting Using Maximum Likelihood (cont'd)

- For the Gaussian distribution the maximization can be done analytically. Taking derivatives of Equation 4.68 with respect to the parameters yields

$$\frac{\partial L(\mu,\sigma)}{\partial \mu} = \frac{1}{\sigma^2}\left[\sum_{i=1}^{n} x_i - n\mu\right] \tag{4.69a}$$

and

$$\frac{\partial L(\mu,\sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2. \tag{4.69b}$$

- Setting each of these derivatives equal to zero and solving yields, respectively,

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4.70a}$$

and

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2}. \tag{4.70b}$$

Wilks 2011: section 4.6

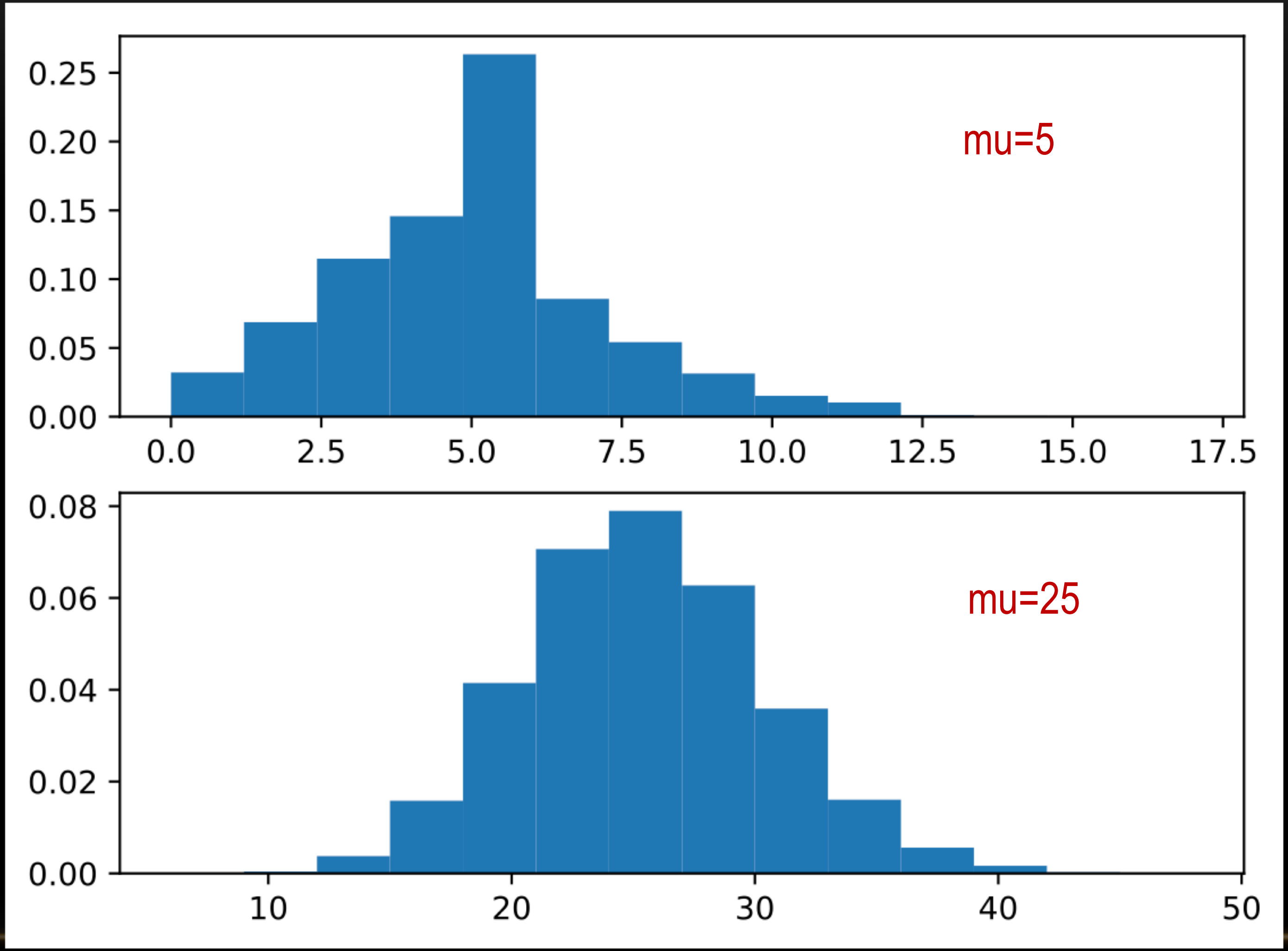
Poisson Distribution

- The Poisson distribution describes the numbers of discrete events occurring in a series, or a sequence of occurrence counts. It thus pertains to data on counts that can take on only nonnegative integer values.
- Assumption: The individual events being counted are independent in the sense that they do not depend on whether or how many other events may have occurred elsewhere in the sequence.
- The Poisson distribution has a single parameter, μ, that specifies the average occurrence rate. The probability of count = x is given below

$$\Pr\{X = x\} = \frac{\mu^x e^{-\mu}}{x!} x = 0, 1, 2, \ldots,$$

Wilks 2011

Examples: Poisson Distributions



The Poisson distribution allocates probability smoothly (given the constraint that the data are discrete) among the possible outcomes, and the most probable count is near the mean (Wilks 2011)

Statistical Forecasting using Poisson Regression

- Poisson regression can be used for predictands that can take on only nonnegative integers, such as counts of tornadoes or hurricanes.
- If the outcomes to be predicted by a regression are Poisson-distributed counts, and the Poisson parameter, μ, may depend on one or more predictor variables, we can structure a regression to specify the Poisson mean as a nonlinear function of those predictors,

$$\mu_i = \exp[b_0 + b_1 x_1 + \cdots + b_K x_K], \tag{6.32a}$$

or

$$\ln(\mu_i) = b_0 + b_1 x_1 + \cdots + b_K x_K. \tag{6.32b}$$

where xi denotes predictors.

- The parameter fitting can be determined by maximizing the Poisson log-likelihood below using iteration.

$$L(\mathbf{b}) = \sum_{i=1}^{n} \{ y_i (b_0 + b_1 x_1 + \cdots + b_K x_K) - \exp(b_0 + b_1 x_1 + \cdots + b_K x_K) \}, \tag{6.33}$$

Wilks 2011, section 6.3.2

References

- *Wilks, D., 2011: Statistical Methods in the Atmospheric Sciences . 3rd ed. Oxford ; Academic Press, 2011. Section 4.1.1, 4.2.4, 4.6, and 6.3.2*