Discrete Distributions

DASC 512

Overview

- Uniform Distribution
- Bernoulli Distribution
- Geometric Distribution
- Binomial Distribution
- Poisson Distribution

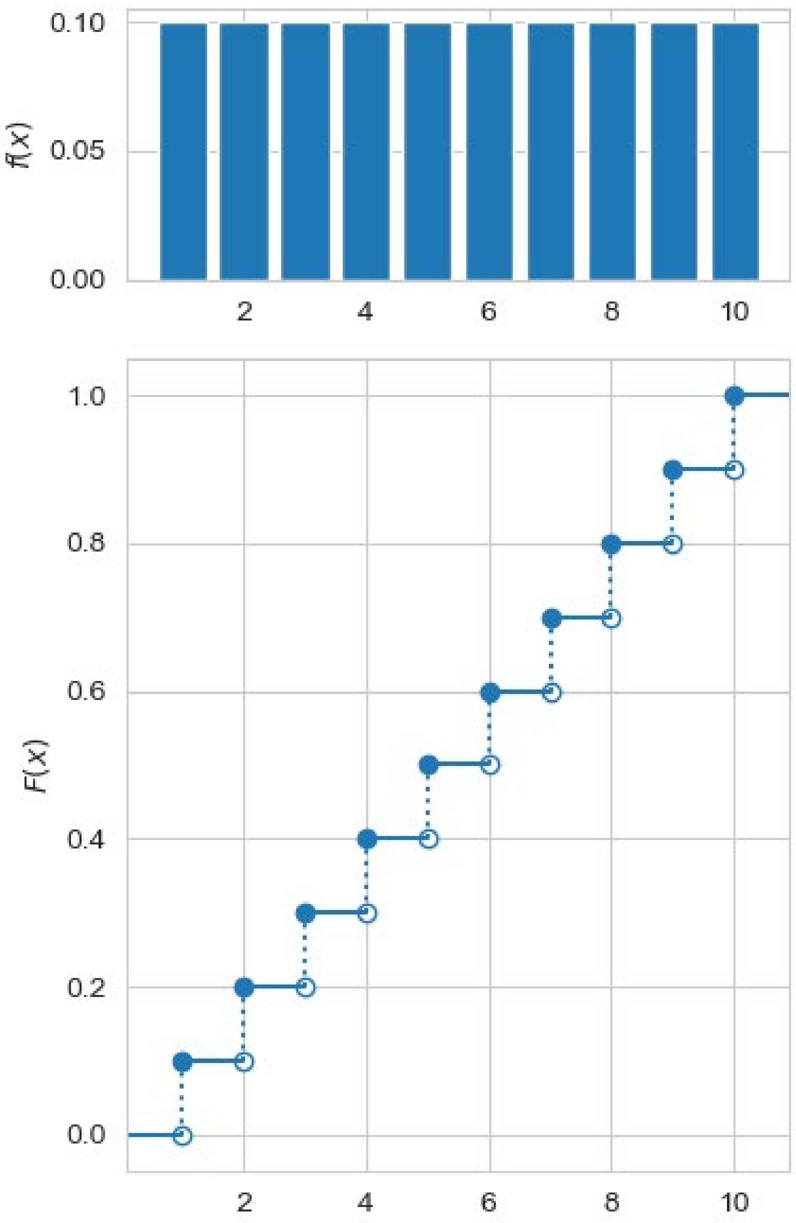
Discrete Uniform Distribution

In a uniform distribution, every outcome is equally probable

$$X \sim \text{Uniform}(n)$$

where n is the upper limit.

$$f(k) = \begin{cases} \frac{1}{n} & \text{if } k = 1,2,\ldots,n \\ 0 & \text{otherwise} \end{cases}$$

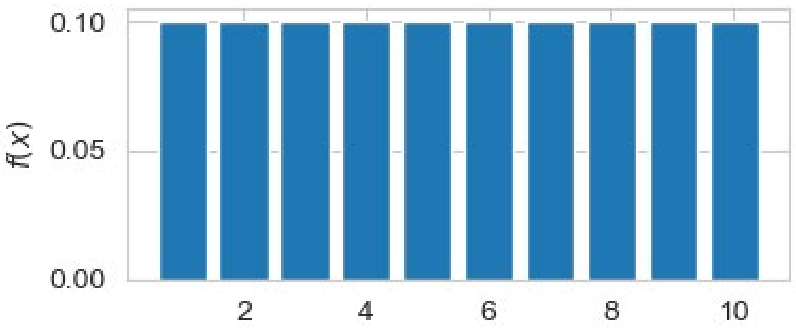


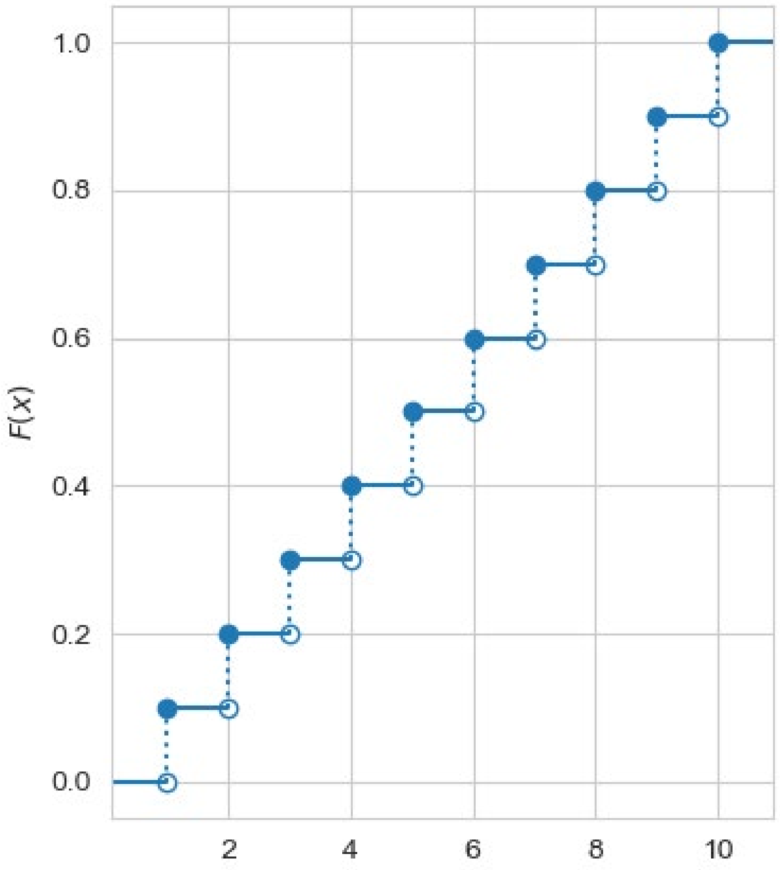
Discrete Uniform Distribution

Notable characteristics:

$$\mu = \frac{n+1}{2}$$

$$\sigma^2 = \frac{n^2-1}{12}$$



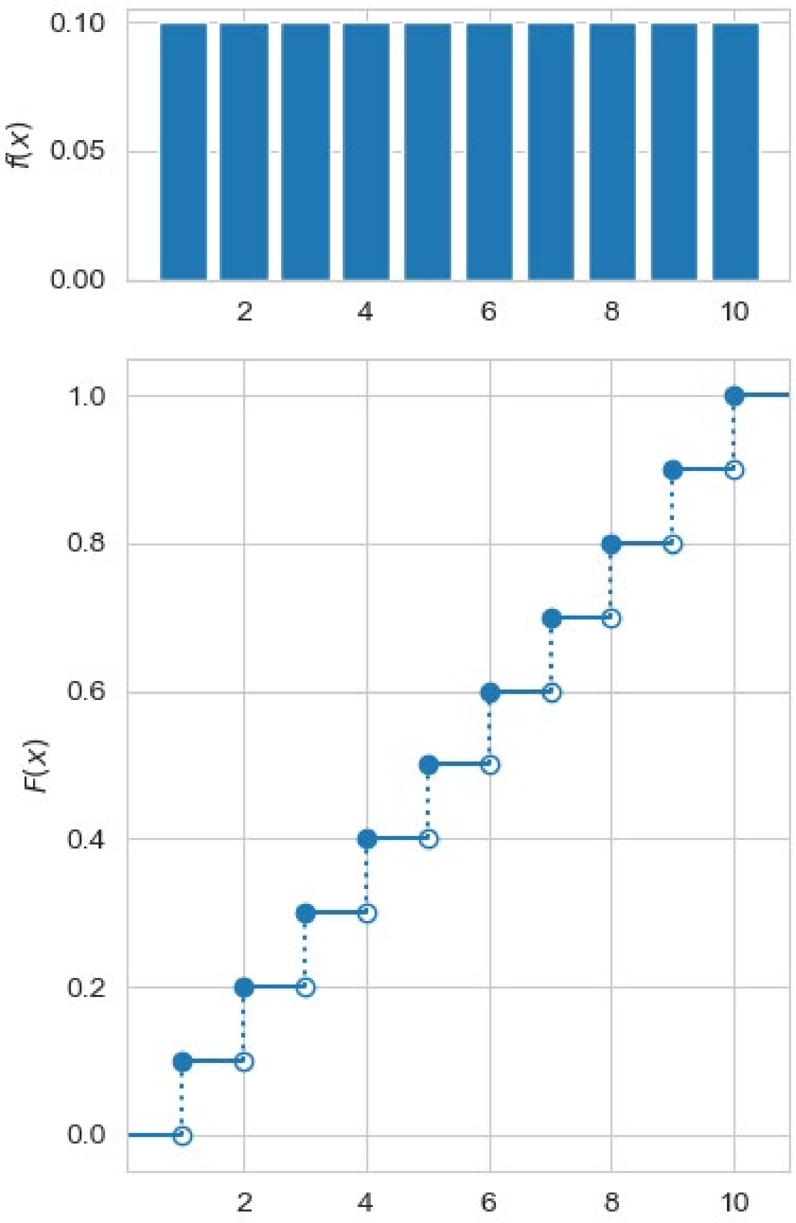


Discrete Uniform Distribution

In SciPy:

scipy.stats.randint(k, low, high, loc=μ)

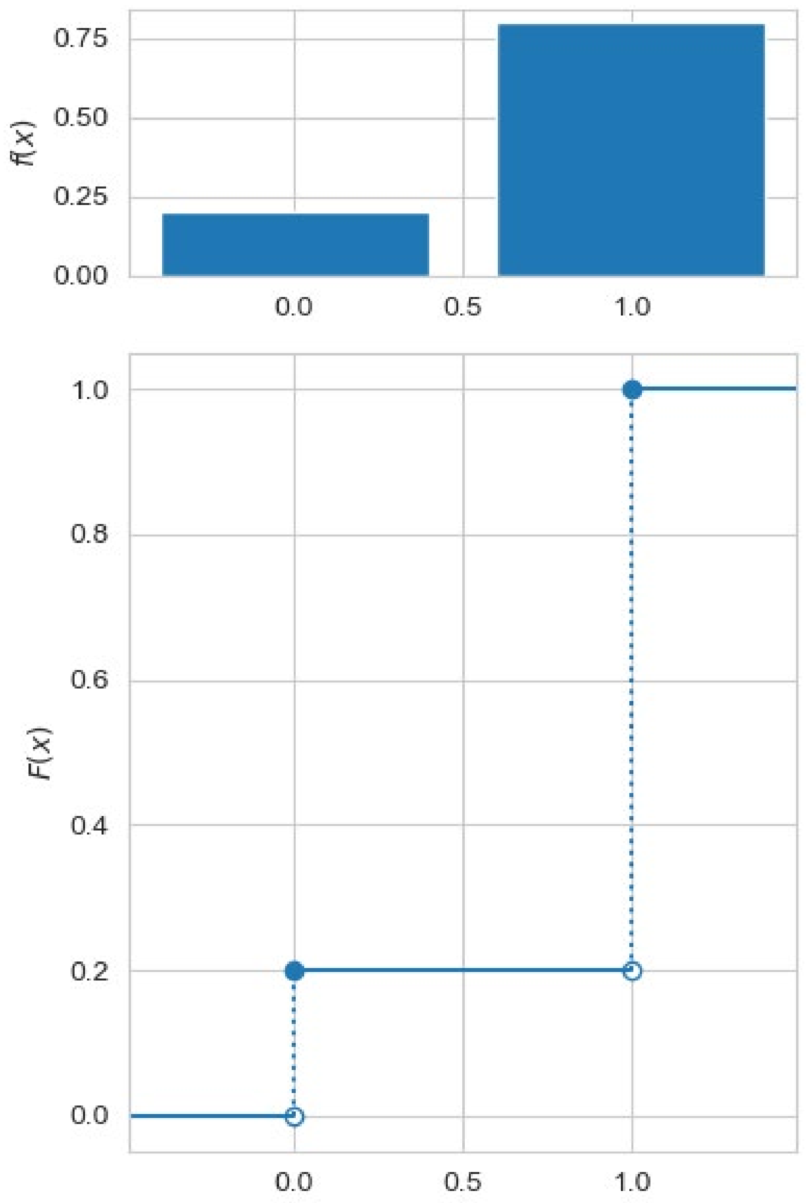
- Discrete uniform from low to (high-1) like range
- μ here refers to the location family of uniform distributions – another way to set low/high



Bernoulli Distribution

Let an outcome of interest ("success") be denoted by $X = 1$ and any other outcome ("failure") be denoted by $X = 0$.

In a Bernoulli trial, success occurs with probability p. The Bernoulli distribution represents a single Bernoulli trial.

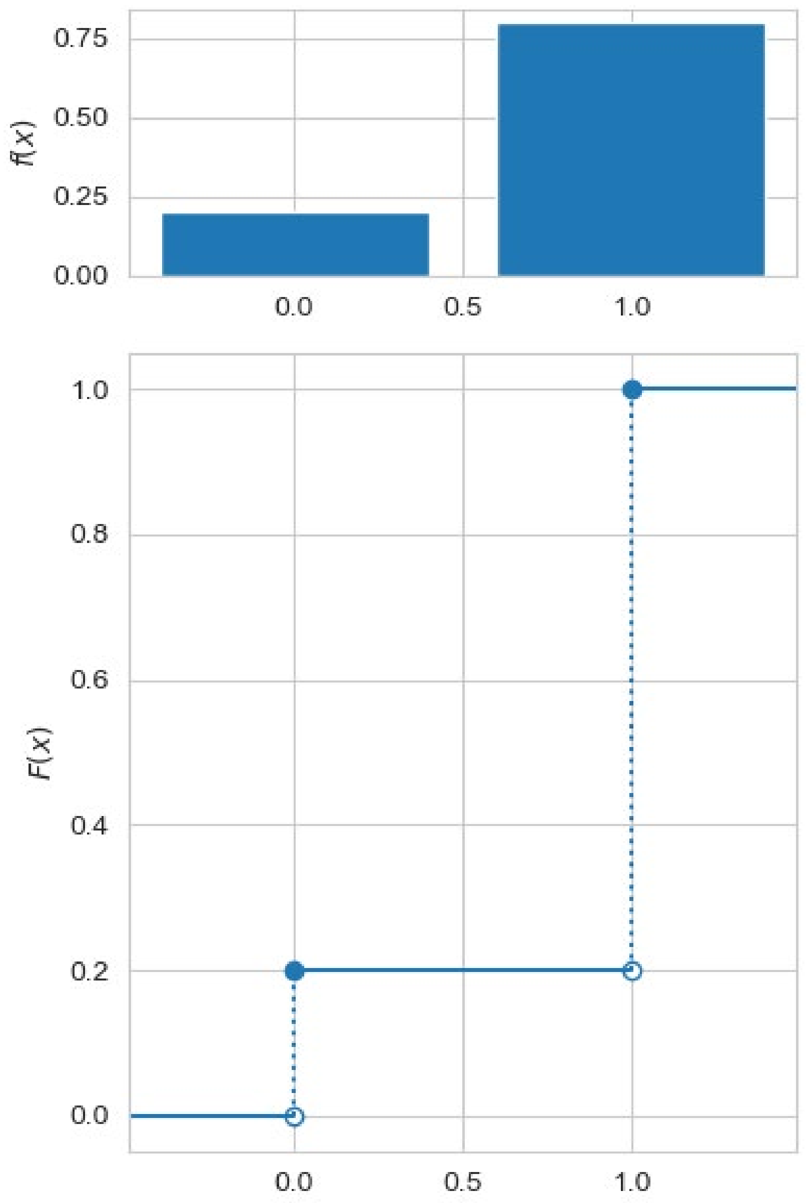


Bernoulli Distribution

$$X \sim \text{Bernoulli}(p)$$

$$f(k) = \begin{cases} p & k = 1 \\ 1 - p & k = 0 \end{cases}$$

$$F(k) = \begin{cases} 0 & k < 0 \\ 1 - p & 0 \leq k < 1 \\ 1 & k \geq 1 \end{cases}$$

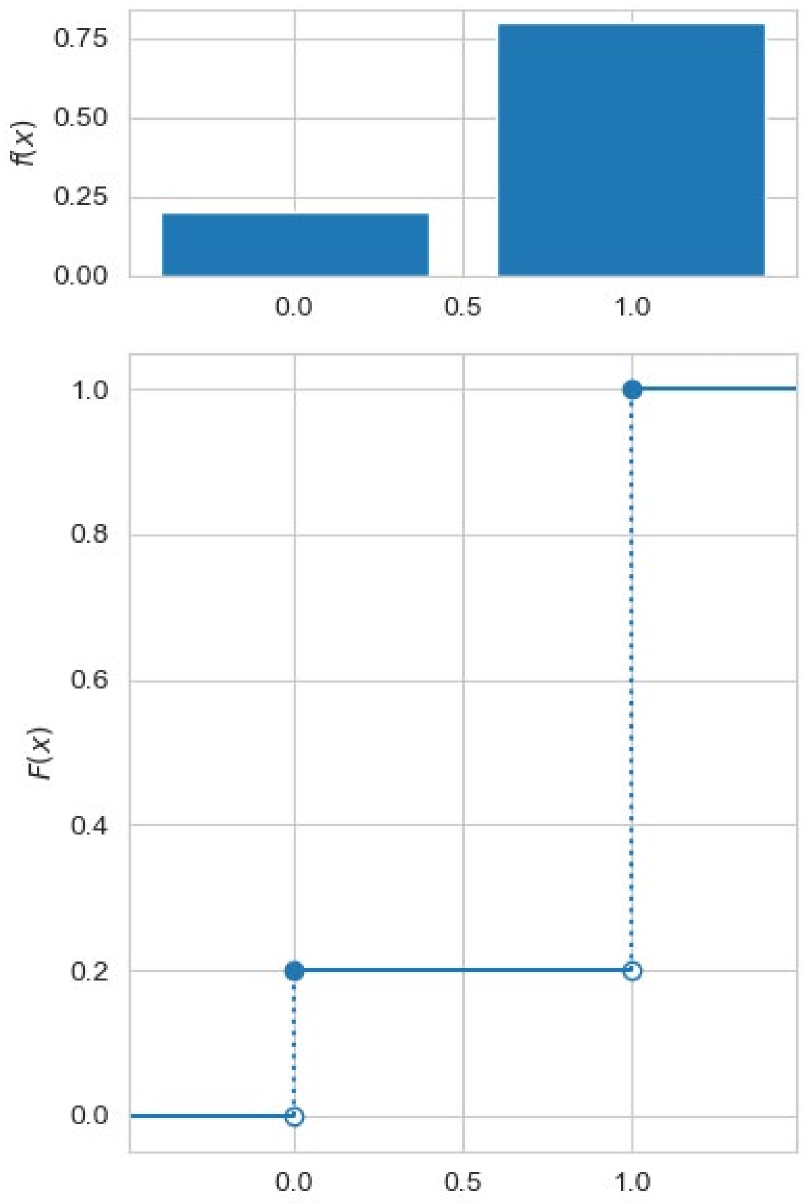


Bernoulli Distribution

Notable characteristics:

$$\mu = p$$

$$\sigma^2 = p(1 - p)$$

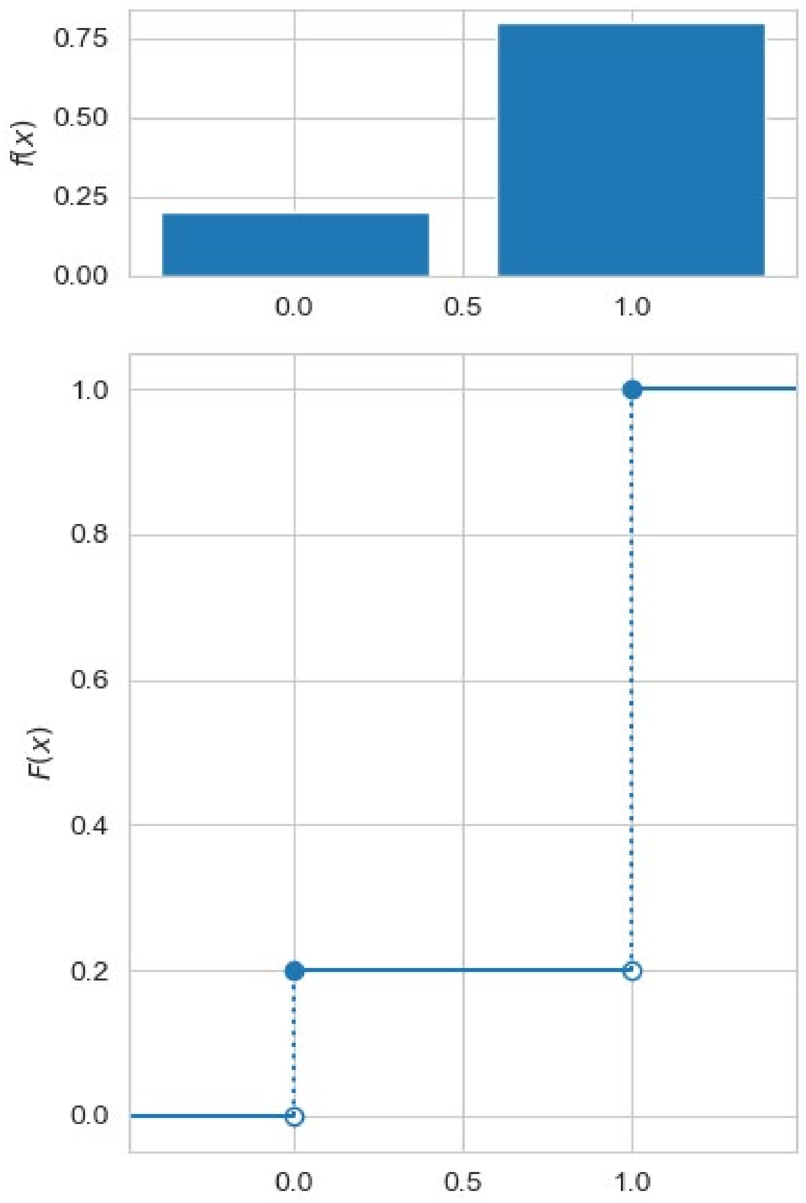


Bernoulli Distribution

In SciPy:

scipy.stats.bernoulli(k, p=p, loc=μ)

- μ here refers to the location family of Bernoulli distributions (not commonly used)



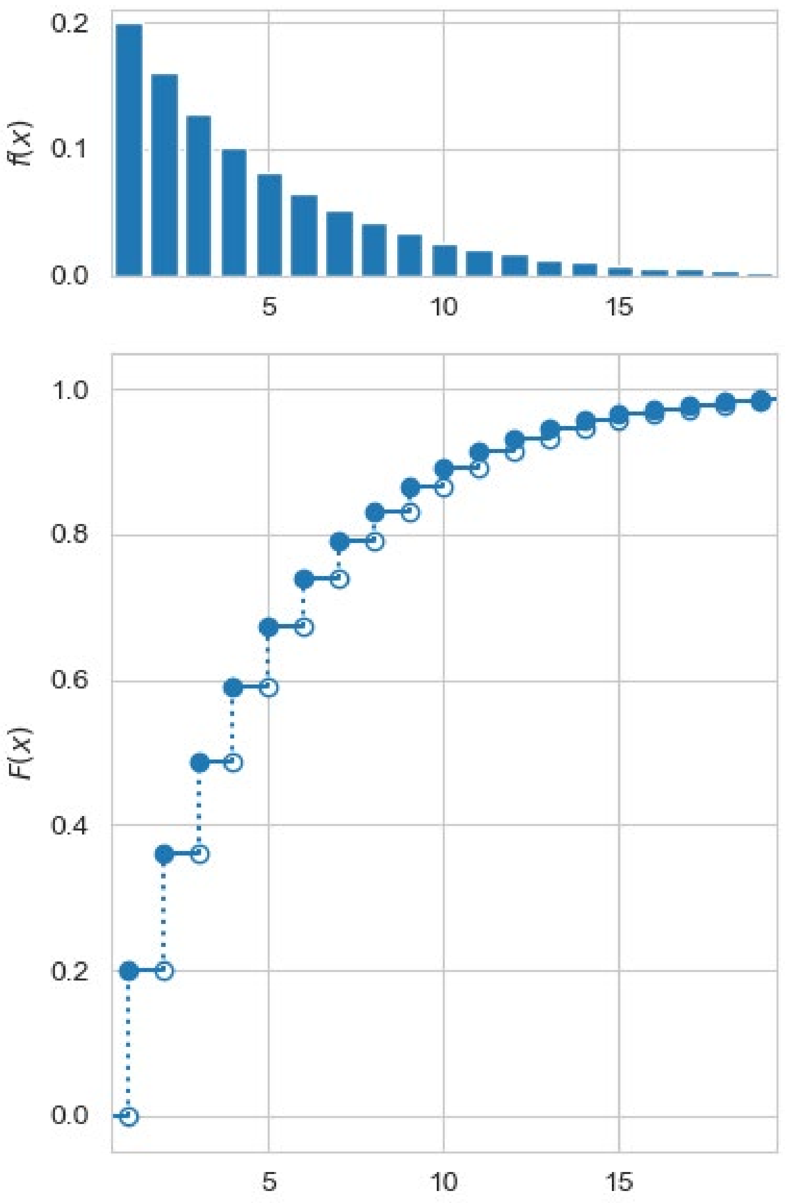
Geometric Distribution

Now let X be the number of iid Bernoulli trials with probability of success p, until the first success is achieved

$$X \sim \mathrm{Geom}(p)$$

$$f(k) = (1-p)^{k-1}p, \qquad k \geq 1$$

$$F(k) = 1 - (1-p)^{\lfloor k \rfloor}$$

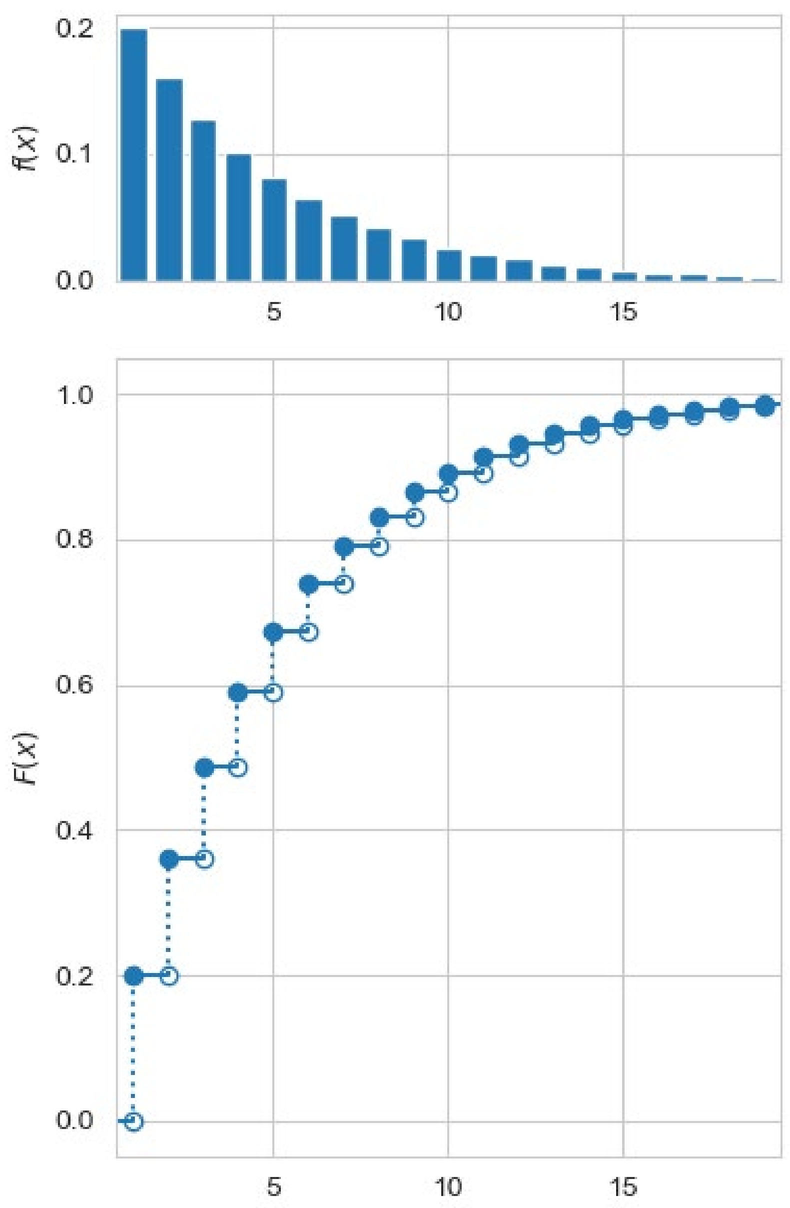


Geometric Distribution

Notable characteristics:

$$\mu = \frac{1}{p}$$

$$\sigma^2 = \frac{1-p}{p^2}$$

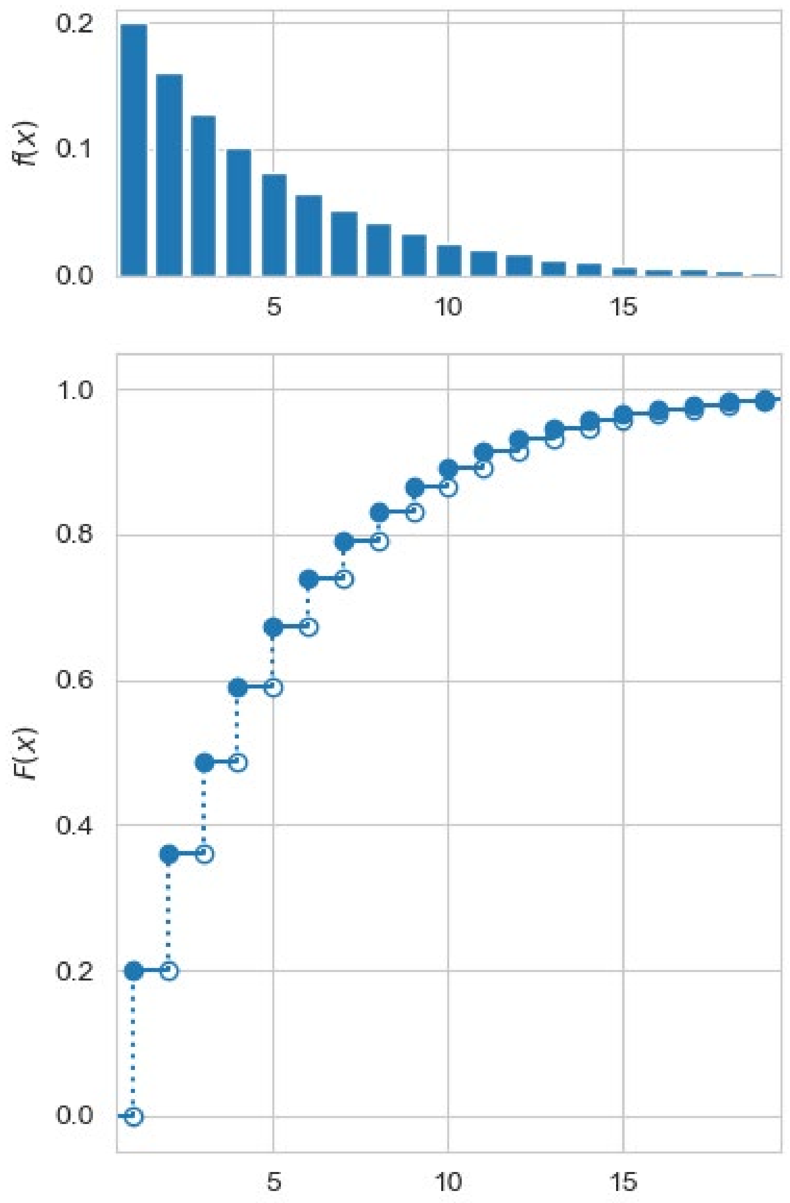


Geometric Distribution

In SciPy:

scipy.stats.geom(k, p=p, loc=μ)

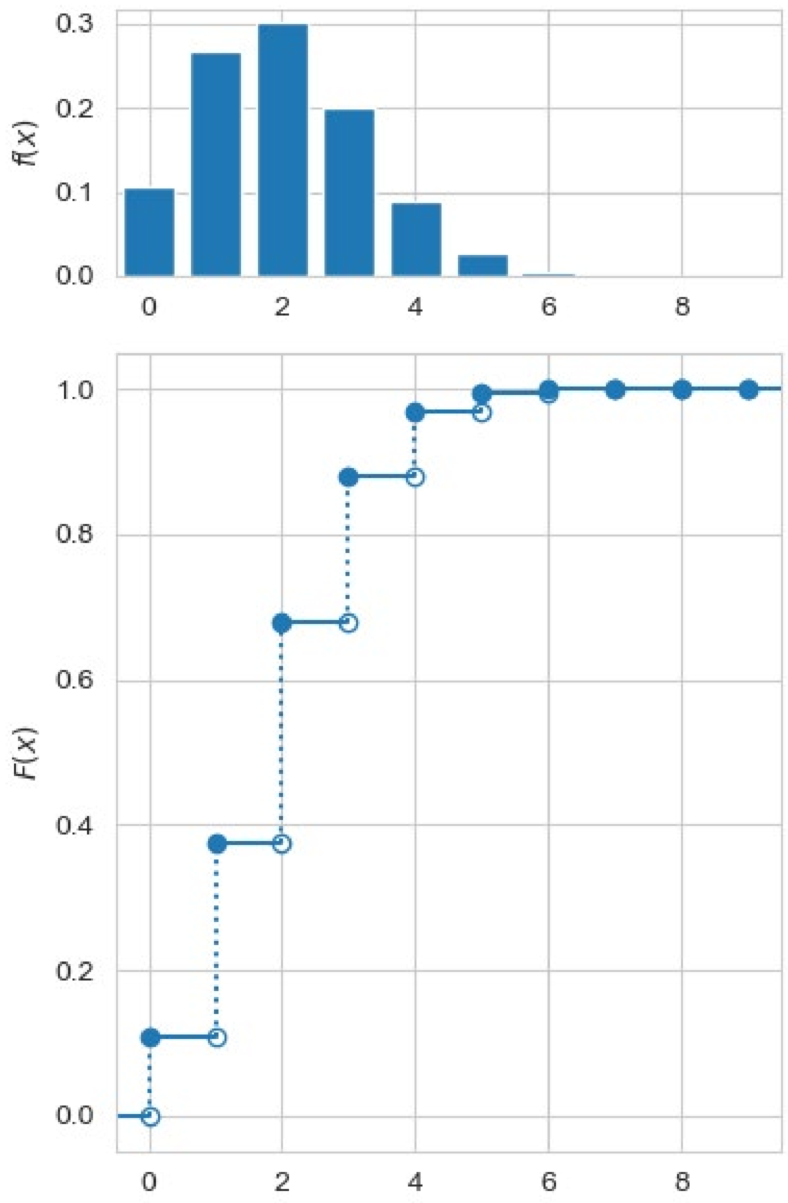
- μ here refers to the location family of geometric distributions (not commonly used)



Binomial Distribution

Now consider a fixed number n of iid Bernoulli trials are conducted. Let X be the total number of successes.

$$X \sim \mathrm{Binom}(n, p)$$

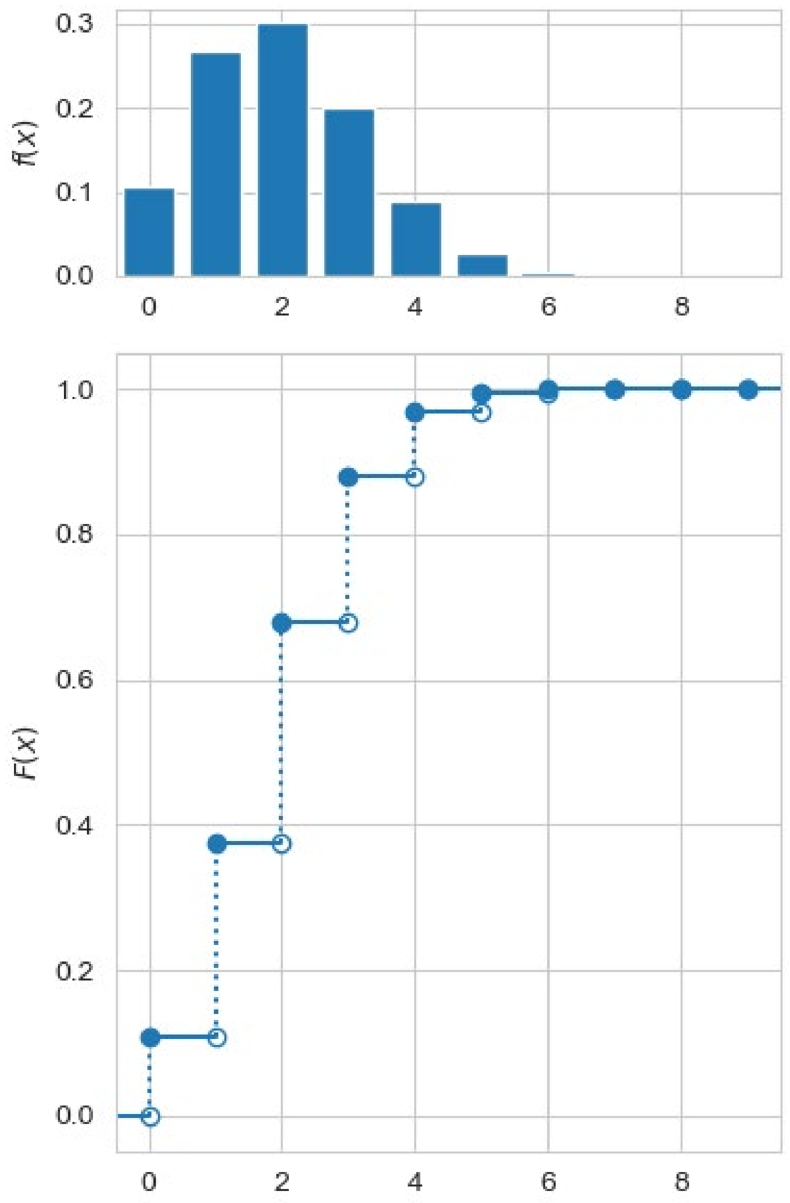


Binomial Distribution

$$X \sim \text{Binom}(n, p)$$

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$F(k) = \sum_{i=0}^{k} f(i)$$

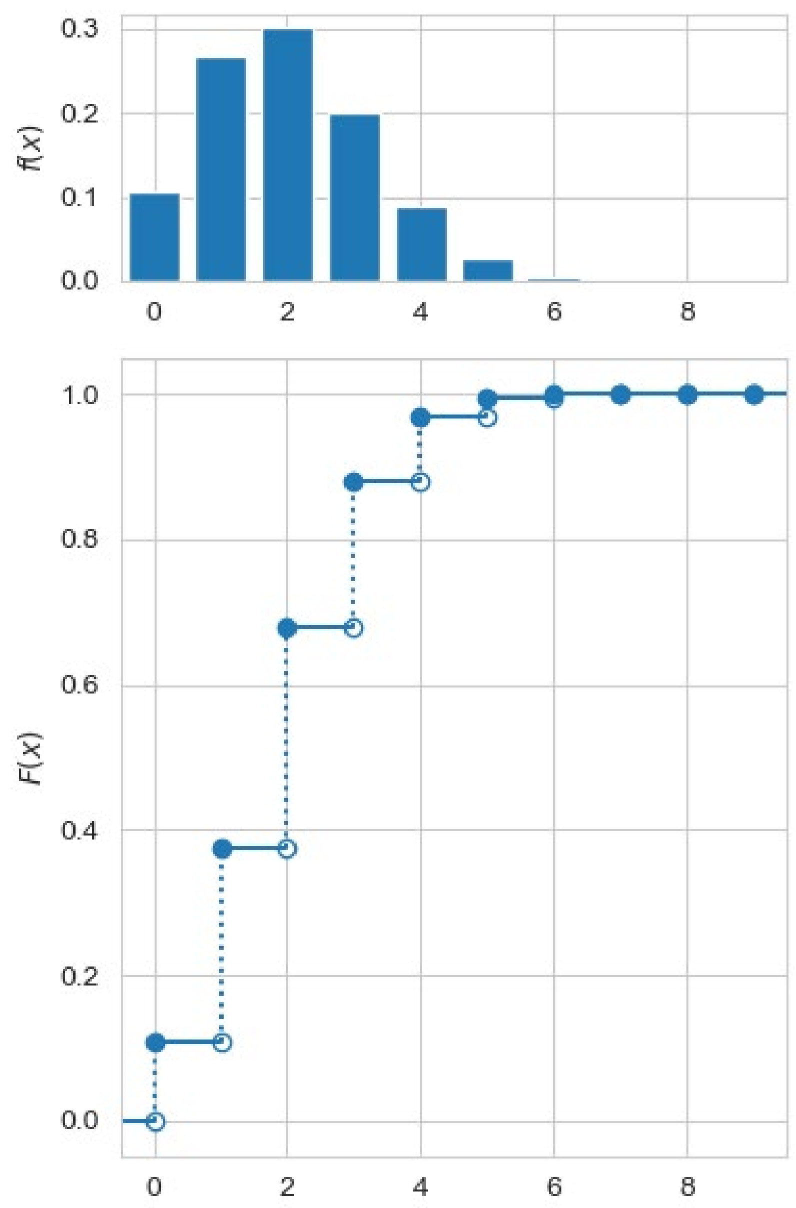


Binomial Distribution

Notable characteristics:

$$\mu = np$$

$$\sigma^2 = np(1 - p)$$

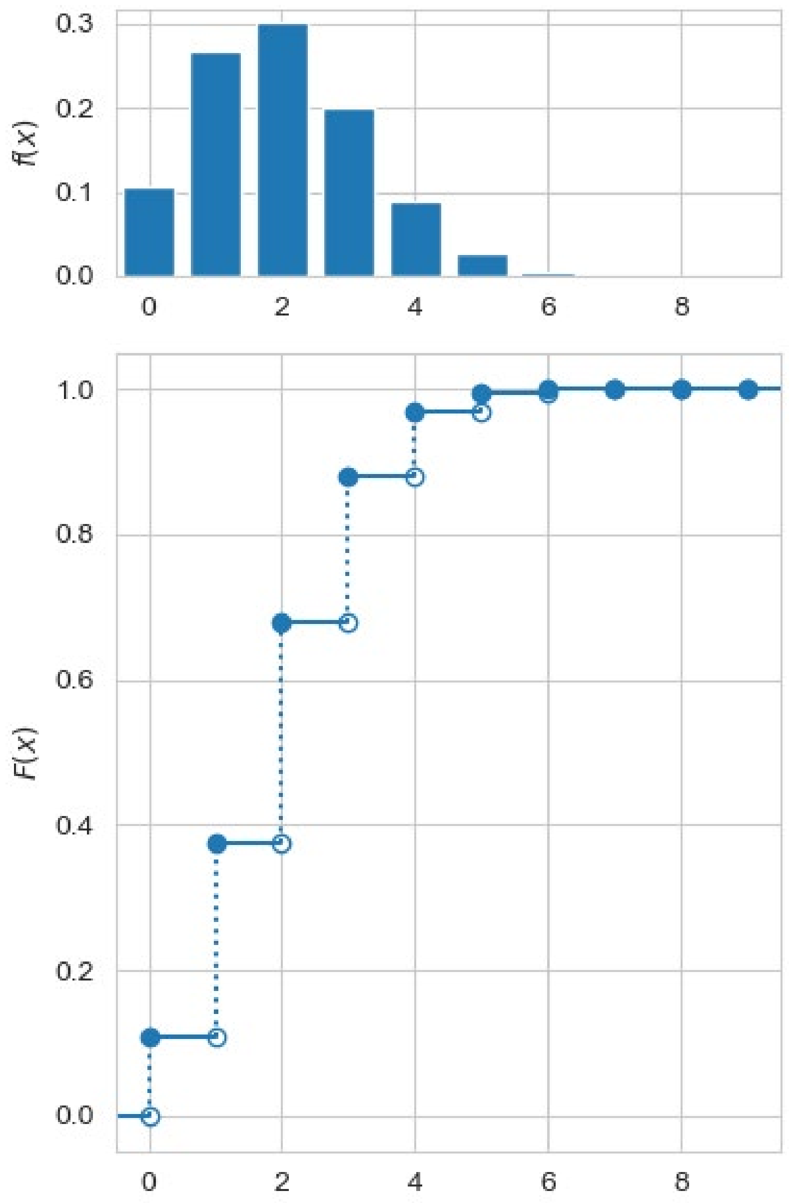


Binomial Distribution

In SciPy:

scipy.stats.binom(k, n=n, p=p, loc=μ)

- μ here refers to the location family of geometric distributions (not commonly used)

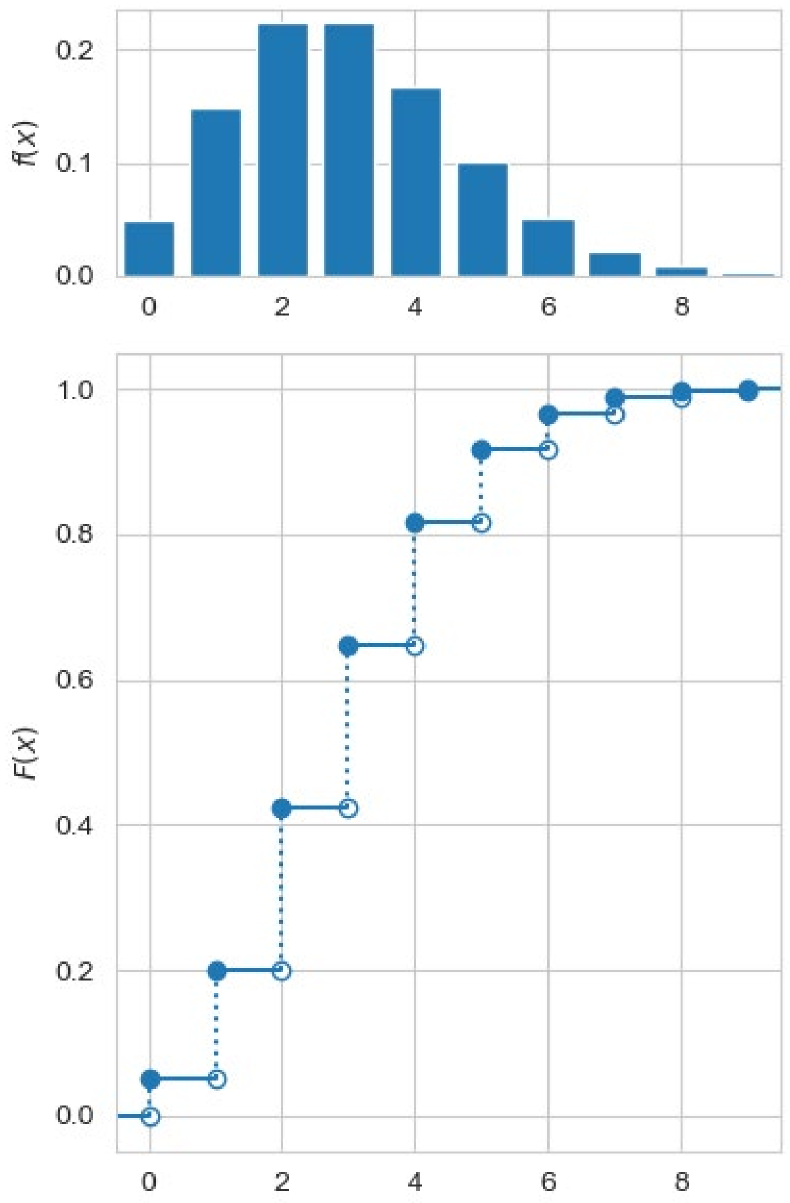


Poisson Distribution

Now for something not related to Bernoulli.

Let X be the number of occurrences of some event within a specific interval, where that events occurs on average λ (lambda) times per interval

$$X \sim \mathrm{Poisson}(\lambda)$$

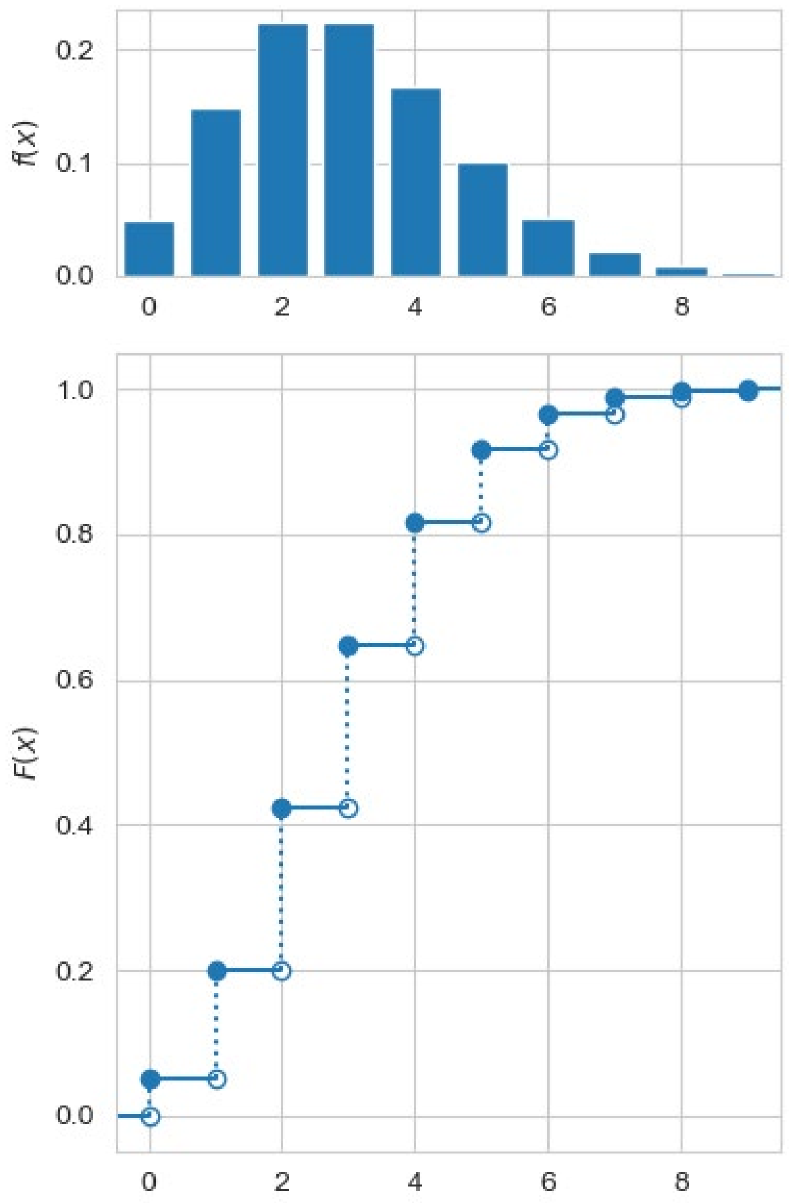


Poisson Distribution

$$X \sim \text{Poisson}(\lambda)$$

$$f(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$F(k) = e^{-\lambda} \sum_{i=0}^{\lfloor k \rfloor} \frac{\lambda^i}{i!}$$



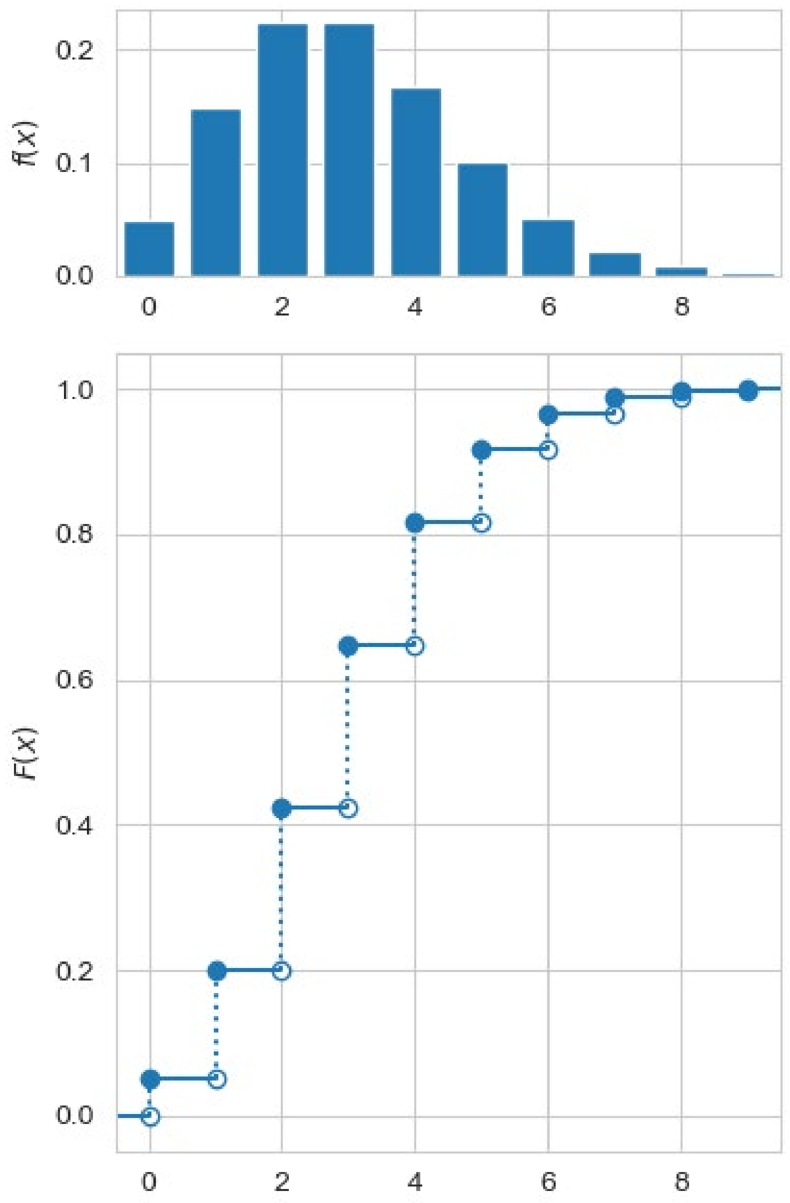
Poisson Distribution

Notable characteristics:

$$\mu = \lambda$$

$$\sigma^2 = \lambda$$

Related to the continuous exponential distribution – the time between events is exponentially distributed

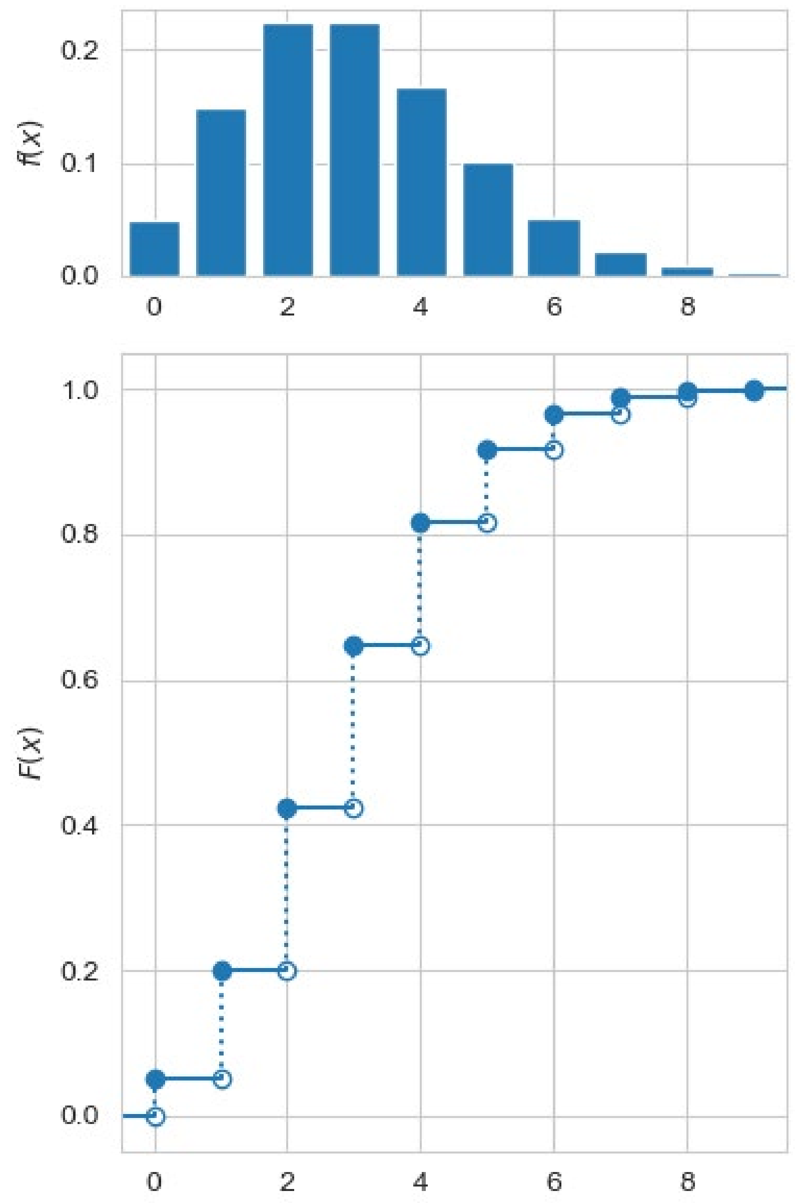


Poisson Distribution

In SciPy:

scipy.stats.poisson(k, mu=λ, loc=μ)

- μ here refers to the location family of geometric distributions (not commonly used)



Resources

Wikipedia

- https://en.Wikipedia.org

SciPy.Stats Reference

- https://docs.scipy.org/doc/scipy/reference/stats.html

For deep theory, the STAT 601/602 textbook

- Casella, G., & Berger, R. L. (2002). *Statistical inference*. Cengage Learning.

Recap

- Uniform Distribution
- Bernoulli Distribution
- Geometric Distribution
- Binomial Distribution
- Poisson Distribution