MATH 205: Statistical methods

Lab 6: Central limit theorem

Goals: Generate random data

- built-in distributions
- central limit theorem

Useful distributions

In the lectures:

- Uniform distribution
- Normal distribution
- Bernoulli distribution
- Binomial distribution

In the lab:

- Geometric distribution
- Poisson distribution
- Beta distribution
- Gamma distribution
- Exponential distribution

Bernoulli random variables

Definition

A Bernoulli random variable takes the value 1 with probability p and 0 with probability $1 - p$.
An experiment associated with a Bernoulli random variable is called a Bernoulli trial. p is also called the probability of success.

The Binomial Probability Distribution

Definition

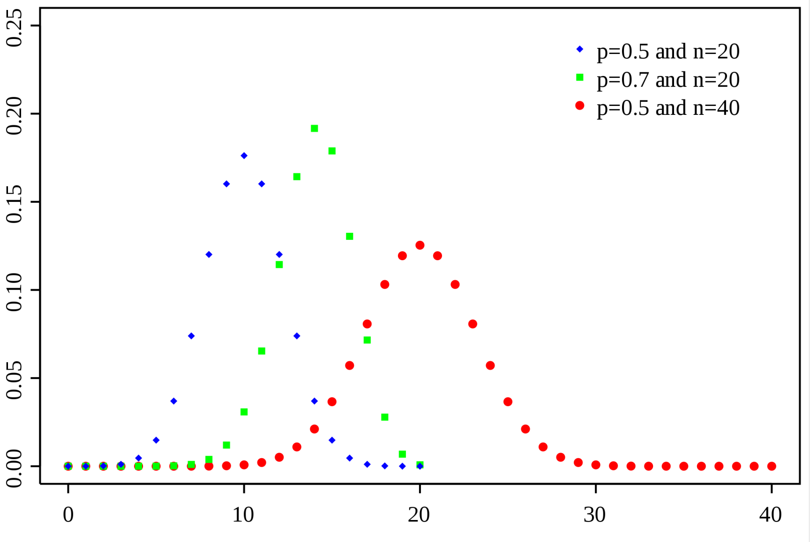
The binomial distribution with parameters N and p is the discrete probability distribution of the number of successes in a sequence of N Bernoulli trials.

$$P(X = n) = \binom{N}{n} p^n (1-p)^{N-n}$$

Recall that:

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

The Binomial probability distribution



The Binomial probability distribution

Alternative definition: If $\{X_1, X_2, \ldots, X_N\}$ is a sequence of independent Bernoulli random variables with probability p. Then

$$Y = X_1 + X_2 + \ldots X_N$$

follows binomial probability distribution $B(N, p)$.

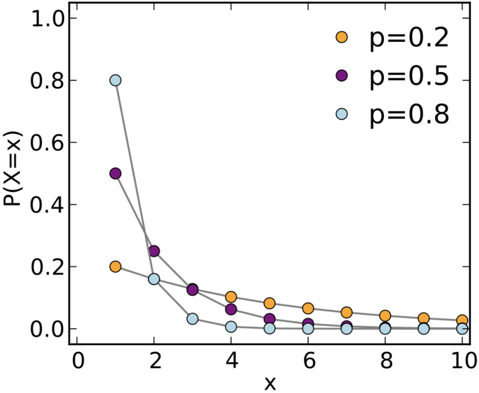
The geometric distribution

Definition

The geometric distribution is the probability distribution of the number X of Bernoulli trials needed to get one success, supported on the set $\{1, 2, 3, \ldots\}$

$$P(X = n) = p(1 - p)^{n-1}$$

The geometric distribution



A geometric distribution with parameter p has mean $1/p$ and variance $(1-p)/p^2$.

Poisson Distribution

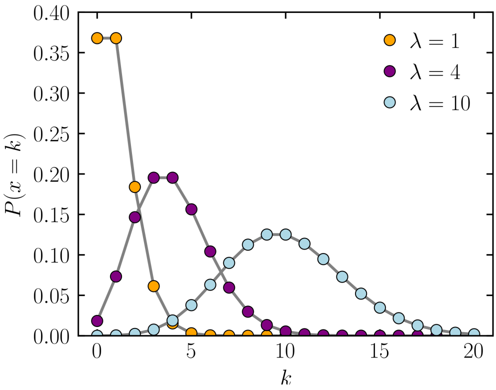
Definition

A non-negative, integer valued random variable X has a Poisson distribution when its probability distribution takes the form

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

where $\lambda > 0$ is a parameter often known as the intensity of the distribution.

Poisson Distribution



A Poisson distribution with intensity λ has mean λ and variance λ.

Poisson Distribution

Usually used to model counts that occur in an interval of time or space that

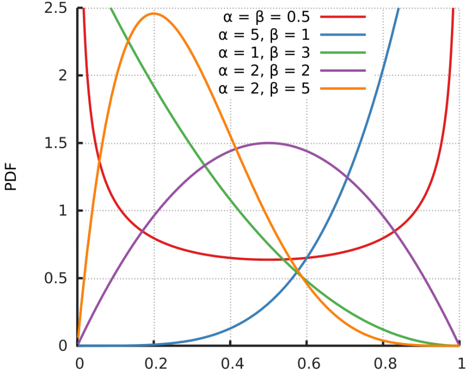
- occur with some fixed average rate
- observation occurs on disjoint interval are independent

Examples:

- the marketing phone calls you receive during the day time
- number of Prussian soldiers killed by horse-kicks each year
- the number of raisins in a loaf/slice of raisin bread

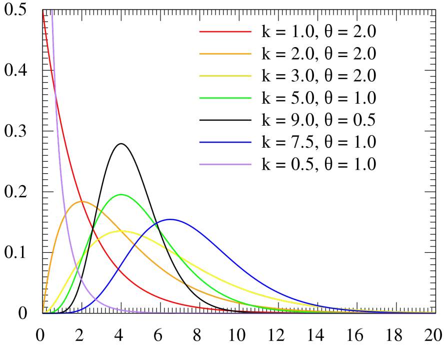
Beta distributions

The Beta distribution is a family of continuous probability distributions defined on the interval $[0, 1]$ parameterized by two positive shape parameters, denoted by α and β, that control the shape of the distribution.



Gamma distributions

The Beta distribution is a family of continuous probability distribution for a non-negative continuous random variable, parameterized by two positive shape parameters, denoted by α and β, that control the shape of the distribution.



Exponential distributions

A special case of Gamma is the exponential distribution ($\alpha = 1$)

$$f(x) = \beta e^{-\beta x}, \qquad x > 0$$