Department of Computer Science
UNIVERSITY OF COLORADO **BOULDER**

Probability Distributions: Discrete

Introduction to Data Science Algorithms

Jordan Boyd-Graber and Michael Paul

SEPTEMBER 27, 2016

Poisson distribution

- We showed that the Bernoulli/binomial/categorical/multinomial are all related to each other
- Lastly, we will show something a little different
- The **Poisson** distribution gives the probability that an event will occur a certain number of times within a time interval
- Examples:
  - The number of goals in a soccer match
  - The amount of mail received in a day
  - The number of times a river will flood in a decade

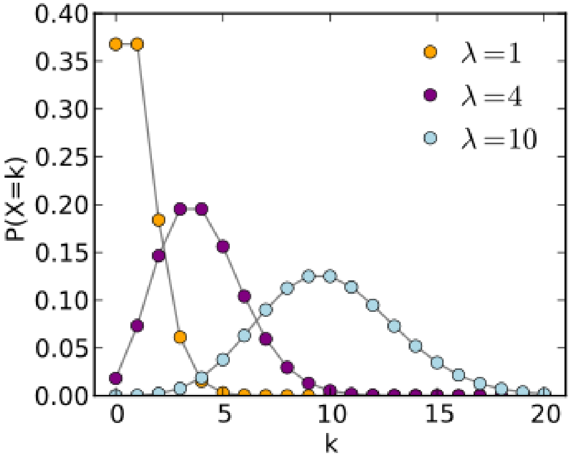
Poisson distribution

- Let the random variable X refer to the count of the number of events over whatever interval we are modeling.
  - X can be any positive integer or zero: $\{0, 1, 2, \ldots\}$
- The probability mass function for the Poisson distribution is:

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

- The Poisson distribution has one parameter λ, which is the average number of events in an interval.
  - $\mathbb{E}[X] = \lambda$

Poisson distribution



Poisson distribution

- Example: Poisson is good model of World Cup match having a certain number of goals
- A World Cup match has an average of 2.5 goals scored: $\lambda = 2.5$
-
  - $P(X=0) = \frac{2.5^0 e^{-2.5}}{0!} = \frac{e^{-2.5}}{1} = 0.082$
  - $P(X=1) = \frac{2.5^1 e^{-2.5}}{1!} = \frac{2.5e^{-2.5}}{1} = 0.205$
  - $P(X=2) = \frac{2.5^2 e^{-2.5}}{2!} = \frac{6.25e^{-2.5}}{2} = 0.257$
  - $P(X=3) = \frac{2.5^3 e^{-2.5}}{3!} = \frac{15.625e^{-2.5}}{6} = 0.213$
  - $P(X=4) = \frac{2.5^4 e^{-2.5}}{4!} = \frac{39.0625e^{-2.5}}{24} = 0.133$

  $\dots$
  - $P(X=10) = \frac{2.5^{10} e^{-2.5}}{10!} = \frac{9536.7432e^{-2.5}}{3628800} = 0.00022$

  $\dots$

Wrap up

- Next time: practice with discrete distributions
- Next week: continuous distributions
- Homework 2
  - Building language model for Republican and Democratic presidents
  - Building distributions for Republican and Democractic states' districts