Random variables, probability distributions, expected value, variance, binomial distribution, Gaussian distribution (normal distribution)

Relevant Readings: Table 5.2, Section 5.3 up through 5.3.4 in Mitchell

CS495 - Machine Learning, Fall 2009

Prob and stats

- Probability, statistics, and sampling theory aren't only just useful in ML. These are generally useful in computer science (such is in the study of randomized algorithms)

Prob and stats

- Probability, statistics, and sampling theory aren't only just useful in ML. These are generally useful in computer science (such is in the study of randomized algorithms)

Some basics

- A *random variable* is a numerical outcome of some random "experiment"

Some basics

- A *random variable* is a numerical outcome of some random "experiment"
    - It can be thought of as an unknown value that changes every time it is observed

Some basics

- A *random variable* is a numerical outcome of some random “experiment”
  - It can be thought of as an unknown value that changes every time it is observed
    - Similar to Math.random()

Some basics

- A *random variable* is a numerical outcome of some random "experiment"
  - It can be thought of as an unknown value that changes every time it is observed
    - Similar to Math.random()
  - Example: The result of rolling a die

Some basics

- A *random variable* is a numerical outcome of some random "experiment"
  - It can be thought of as an unknown value that changes every time it is observed
    - Similar to Math.random()
  - Example: The result of rolling a die
- A random variable's behavior is governed its *probability distribution*

Some basics

- A *random variable* is a numerical outcome of some random "experiment"
  - It can be thought of as an unknown value that changes every time it is observed
    - Similar to Math.random()
  - Example: The result of rolling a die
- A random variable's behavior is governed its *probability distribution*
  - The probability distribution is a specification for the likelihood of the random variable taking on different values

Some basics

- A *random variable* is a numerical outcome of some random "experiment"
  - It can be thought of as an unknown value that changes every time it is observed
    - Similar to Math.random()
  - Example: The result of rolling a die
- A random variable's behavior is governed its *probability distribution*
  - The probability distribution is a specification for the likelihood of the random variable taking on different values
  - Example: For rolling a die, the probability distribution is *uniform*, meaning that each possibility (1, 2, 3, 4, 5, 6) has equal probability of coming up (1/6).

Some basics

- A *random variable* is a numerical outcome of some random "experiment"
  - It can be thought of as an unknown value that changes every time it is observed
    - Similar to Math.random()
  - Example: The result of rolling a die
- A random variable's behavior is governed its *probability distribution*
  - The probability distribution is a specification for the likelihood of the random variable taking on different values
  - Example: For rolling a die, the probability distribution is *uniform*, meaning that each possibility (1, 2, 3, 4, 5, 6) has equal probability of coming up (1/6).
  - Random variables can be discrete or continuous

Some basics

- A *random variable* is a numerical outcome of some random "experiment"
  - It can be thought of as an unknown value that changes every time it is observed
    - Similar to Math.random()
  - Example: The result of rolling a die
- A random variable's behavior is governed its *probability distribution*
  - The probability distribution is a specification for the likelihood of the random variable taking on different values
  - Example: For rolling a die, the probability distribution is *uniform*, meaning that each possibility (1, 2, 3, 4, 5, 6) has equal probability of coming up (1/6).
  - Random variables can be discrete or continuous
  - The probability that random variable X takes on the value x is denoted $\Pr(X = x)$

Some basics

- A *random variable* is a numerical outcome of some random "experiment"
  - It can be thought of as an unknown value that changes every time it is observed
    - Similar to Math.random()
  - Example: The result of rolling a die
- A random variable's behavior is governed its *probability distribution*
  - The probability distribution is a specification for the likelihood of the random variable taking on different values
  - Example: For rolling a die, the probability distribution is *uniform*, meaning that each possibility (1, 2, 3, 4, 5, 6) has equal probability of coming up (1/6).
  - Random variables can be discrete or continuous
  - The probability that random variable X takes on the value x is denoted $\Pr(X = x)$

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.
  - What is $\Pr(X = 1/4)$?

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.
  - What is $\Pr(X = 1/4)$?
    - Zero.

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.
  - What is $\Pr(X = 1/4)$?
    - Zero.
  - What is $\Pr(1/4 \leq X \leq 1/2)$?

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.
  - What is $\Pr(X = 1/4)$?
    - Zero.
  - What is $\Pr(1/4 \leq X \leq 1/2)$?
    - $1/8$

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.
  - What is $\Pr(X = 1/4)$?
    - Zero.
  - What is $\Pr(1/4 \leq X \leq 1/2)$?
    - $1/8$
  - So then for continuous random variables, we work in terms of ranges of values, rather than the probability of equaling some value exactly

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.
  - What is $\Pr(X = 1/4)$?
    - Zero.
  - What is $\Pr(1/4 \leq X \leq 1/2)$?
    - $1/8$
  - So then for continuous random variables, we work in terms of ranges of values, rather than the probability of equaling some value exactly
  - Think of it this way: what's the probability of throwing a dart and hitting *exactly* the right spot on the wall?

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.
  - What is $\Pr(X = 1/4)$?
    - Zero.
  - What is $\Pr(1/4 \leq X \leq 1/2)$?
    - $1/8$
  - So then for continuous random variables, we work in terms of ranges of values, rather than the probability of equaling some value exactly
  - Think of it this way: what's the probability of throwing a dart and hitting *exactly* the right spot on the wall?
    - Zero.

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.
  - What is $\Pr(X = 1/4)$?
    - Zero.
  - What is $\Pr(1/4 \leq X \leq 1/2)$?
    - $1/8$
  - So then for continuous random variables, we work in terms of ranges of values, rather than the probability of equaling some value exactly
  - Think of it this way: what's the probability of throwing a dart and hitting *exactly* the right spot on the wall?
    - Zero.
    - But after it's thrown, it actually hit *some* exact spot.

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.
  - What is $\Pr(X = 1/4)$?
    - Zero.
  - What is $\Pr(1/4 \leq X \leq 1/2)$?
    - $1/8$
  - So then for continuous random variables, we work in terms of ranges of values, rather than the probability of equaling some value exactly
  - Think of it this way: what's the probability of throwing a dart and hitting *exactly* the right spot on the wall?
    - Zero.
    - But after it's thrown, it actually hit *some* exact spot.
    - Therefore zero-probability doesn't necessarily mean "impossible"

Continuous probability distributions

- Consider a uniform continuous random variable X over the range [0,2]
  - What is $\Pr(X = 1/2)$?
    - Zero.
  - What is $\Pr(X = 1/4)$?
    - Zero.
  - What is $\Pr(1/4 \leq X \leq 1/2)$?
    - $1/8$
  - So then for continuous random variables, we work in terms of ranges of values, rather than the probability of equaling some value exactly
  - Think of it this way: what's the probability of throwing a dart and hitting *exactly* the right spot on the wall?
    - Zero.
    - But after it's thrown, it actually hit *some* exact spot.
    - Therefore zero-probability doesn't necessarily mean "impossible"

Some basic notions

- Summation notation: $\sum_{i=0}^{n} a_i$ means $a_0 + a_1 + \cdots + a_n$

Some basic notions

- Summation notation: $\sum_{i=0}^{n} a_i$ means $a_0 + a_1 + \cdots + a_n$
- The *expected value* (aka. *mean*) of a random variable Y is $E[Y] = \sum_i y_i \cdot \Pr(Y = y_i)$, where the sum is over all possible outcomes

Some basic notions

- Summation notation: $\sum_{i=0}^{n} a_i$ means $a_0 + a_1 + \cdots + a_n$
- The *expected value* (aka. *mean*) of a random variable Y is $E[Y] = \sum_i y_i \cdot \Pr(Y = y_i)$, where the sum is over all possible outcomes
  - Just a weighted average (the continuous case would be an integral).

Some basic notions

- Summation notation: $\sum_{i=0}^{n} a_i$ means $a_0 + a_1 + \cdots + a_n$
- The *expected value* (aka. *mean*) of a random variable Y is $E[Y] = \sum_i y_i \cdot \Pr(Y = y_i)$, where the sum is over all possible outcomes
  - Just a weighted average (the continuous case would be an integral).
- The *variance* of a random variable Y is $Var(Y) = E[(Y - E[Y])^2]$

Some basic notions

- Summation notation: $\sum_{i=0}^{n} a_i$ means $a_0 + a_1 + \cdots + a_n$
- The *expected value* (aka. *mean*) of a random variable Y is $E[Y] = \sum_i y_i \cdot \Pr(Y = y_i)$, where the sum is over all possible outcomes
  - Just a weighted average (the continuous case would be an integral).
- The *variance* of a random variable Y is $Var(Y) = E[(Y - E[Y])^2]$
  - This is a measure of how much Y deviates from the mean

Some basic notions

- Summation notation: $\sum_{i=0}^{n} a_i$ means $a_0 + a_1 + \cdots + a_n$
- The *expected value* (aka. *mean*) of a random variable Y is $E[Y] = \sum_i y_i \cdot \Pr(Y = y_i)$, where the sum is over all possible outcomes
  - Just a weighted average (the continuous case would be an integral).
- The *variance* of a random variable Y is $Var(Y) = E[(Y - E[Y])^2]$
  - This is a measure of how much Y deviates from the mean
- The *standard deviation* of a random variable Y is simply $\sqrt{Var(Y)}$

Some basic notions

- Summation notation: $\sum_{i=0}^{n} a_i$ means $a_0 + a_1 + \cdots + a_n$
- The *expected value* (aka. *mean*) of a random variable Y is $E[Y] = \sum_i y_i \cdot \Pr(Y = y_i)$, where the sum is over all possible outcomes
  - Just a weighted average (the continuous case would be an integral).
- The *variance* of a random variable Y is $Var(Y) = E[(Y - E[Y])^2]$
  - This is a measure of how much Y deviates from the mean
- The *standard deviation* of a random variable Y is simply $\sqrt{Var(Y)}$

Discrete uniform distribution

- Suppose X is a discrete random variable with uniform probability distribution over the set $\{1, 2, 3, \ldots, m\}$

Discrete uniform distribution

- Suppose X is a discrete random variable with uniform probability distribution over the set $\{1, 2, 3, \ldots, m\}$
- What is $E[X]$?

Discrete uniform distribution

- Suppose X is a discrete random variable with uniform probability distribution over the set $\{1, 2, 3, \ldots, m\}$
- What is $E[X]$?
  - $(m + 1)/2$

Discrete uniform distribution

- Suppose X is a discrete random variable with uniform probability distribution over the set $\{1, 2, 3, \ldots, m\}$
- What is $E[X]$?
  - $(m+1)/2$
- What is $Var[X]$?

Discrete uniform distribution

- Suppose X is a discrete random variable with uniform probability distribution over the set $\{1, 2, 3, \ldots, m\}$
- What is $E[X]$?
  - $(m + 1)/2$
- What is $Var[X]$?
  - $(m^2 - 1)/12$

Discrete uniform distribution

- Suppose X is a discrete random variable with uniform probability distribution over the set $\{1, 2, 3, \dots, m\}$
- What is $E[X]$?
  - $(m+1)/2$
- What is $Var[X]$?
  - $(m^2-1)/12$

Continuous uniform distribution

- Suppose X is a continuous random variable with uniform probability distribution over the interval $[a, b]$

Continuous uniform distribution

- Suppose X is a continuous random variable with uniform probability distribution over the interval $[a, b]$
- What is $E[X]$?

Continuous uniform distribution

- Suppose X is a continuous random variable with uniform probability distribution over the interval $[a, b]$
- What is $E[X]$?
  - $(a + b)/2$

Continuous uniform distribution

- Suppose X is a continuous random variable with uniform probability distribution over the interval $[a, b]$
- What is $E[X]$?
  - $(a + b)/2$
- What is $Var[X]$?

Continuous uniform distribution

- Suppose X is a continuous random variable with uniform probability distribution over the interval $[a, b]$
- What is $E[X]$?
  - $(a + b)/2$
- What is $Var[X]$?
  - $(b - a)^2/12$

Continuous uniform distribution

- Suppose X is a continuous random variable with uniform probability distribution over the interval $[a, b]$
- What is $E[X]$?
  - $(a + b)/2$
- What is $Var[X]$?
  - $(b - a)^2/12$

Binomial distribution

- Imagine an unfair coin n times that has probability p of landing on heads

Binomial distribution

- Imagine an unfair coin n times that has probability p of landing on heads
- A binomial distribution is: "how many times did we get heads?"

Binomial distribution

- Imagine an unfair coin n times that has probability p of landing on heads
- A binomial distribution is: "how many times did we get heads?"
- Is this a discrete or continuous distribution?

Binomial distribution

- Imagine an unfair coin n times that has probability p of landing on heads
- A binomial distribution is: "how many times did we get heads?"
- Is this a discrete or continuous distribution?
  - Discrete

Binomial distribution

- Imagine an unfair coin n times that has probability p of landing on heads
- A binomial distribution is: "how many times did we get heads?"
- Is this a discrete or continuous distribution?
  - Discrete
- Suppose X is a random variable with binomial probability distribution

Binomial distribution

- Imagine an unfair coin n times that has probability p of landing on heads
- A binomial distribution is: "how many times did we get heads?"
- Is this a discrete or continuous distribution?
  - Discrete
- Suppose X is a random variable with binomial probability distribution
- What is $E[X]$?

Binomial distribution

- Imagine an unfair coin n times that has probability p of landing on heads
- A binomial distribution is: "how many times did we get heads?"
- Is this a discrete or continuous distribution?
  - Discrete
- Suppose X is a random variable with binomial probability distribution
- What is $E[X]$?
  - np

Binomial distribution

- Imagine an unfair coin n times that has probability p of landing on heads
- A binomial distribution is: "how many times did we get heads?"
- Is this a discrete or continuous distribution?
  - Discrete
- Suppose X is a random variable with binomial probability distribution
- What is $E[X]$?
  - np
- What is $Var[X]$?

Binomial distribution

- Imagine an unfair coin n times that has probability p of landing on heads
- A binomial distribution is: "how many times did we get heads?"
- Is this a discrete or continuous distribution?
  - Discrete
- Suppose X is a random variable with binomial probability distribution
- What is $E[X]$?
  - np
- What is $Var[X]$?
  - $np(1-p)$

Binomial distribution

- Imagine an unfair coin n times that has probability p of landing on heads
- A binomial distribution is: "how many times did we get heads?"
- Is this a discrete or continuous distribution?
  - Discrete
- Suppose X is a random variable with binomial probability distribution
- What is $E[X]$?
  - np
- What is $Var[X]$?
  - $np(1-p)$

Gaussian (normal) distribution

- If you add lots of random variables together, the sum is in itself a random variable whose probability distribution tends to look like a Gaussian.

Gaussian (normal) distribution

- If you add lots of random variables together, the sum is in itself a random variable whose probability distribution tends to look like a Gaussian.
- Thus, Gaussian distributions show up a lot in nature

Gaussian (normal) distribution

- If you add lots of random variables together, the sum is in itself a random variable whose probability distribution tends to look like a Gaussian.
- Thus, Gaussian distributions show up a lot in nature
- It's a "bell shaped curve"

Gaussian (normal) distribution

- If you add lots of random variables together, the sum is in itself a random variable whose probability distribution tends to look like a Gaussian.
- Thus, Gaussian distributions show up a lot in nature
- It's a "bell shaped curve"
- The probability distribution for a Gaussian with mean μ and variance σ^2 is defined as:

Gaussian (normal) distribution

- If you add lots of random variables together, the sum is in itself a random variable whose probability distribution tends to look like a Gaussian.
- Thus, Gaussian distributions show up a lot in nature
- It's a "bell shaped curve"
- The probability distribution for a Gaussian with mean μ and variance σ^2 is defined as:
  - $e^{-(x-\mu)^2/(2\sigma^2)}/(\sigma\sqrt{2\pi})$

Gaussian (normal) distribution

- If you add lots of random variables together, the sum is in itself a random variable whose probability distribution tends to look like a Gaussian.
- Thus, Gaussian distributions show up a lot in nature
- It's a "bell shaped curve"
- The probability distribution for a Gaussian with mean μ and variance σ^2 is defined as:
  - $e^{-(x-\mu)^2/(2\sigma^2)}/(\sigma\sqrt{2\pi})$

Application

- Suppose we are programming Naive Bayes and want to handle a continuous variable

Application

- Suppose we are programming Naive Bayes and want to handle a continuous variable
- One way we could do this: measure the mean and variance of the relevant examples, model them with a probability distribution (such as a Gaussian), and use that distribution to help determine whether “yes” or “no” is more likely for the given instance

Application

- Suppose we are programming Naive Bayes and want to handle a continuous variable
- One way we could do this: measure the mean and variance of the relevant examples, model them with a probability distribution (such as a Gaussian), and use that distribution to help determine whether "yes" or "no" is more likely for the given instance