AE1MCS: Mathematics for Computer Scientists

Huan Jin and Tony Bellotti
University of Nottingham Ningbo China

Oct 2020

Reading

Kenneth H. Rosen, *Discrete Mathematics and Its Applications*, 7th Edition, 2013.

- Chapter 7, Section 7.2 Probability Theory
- Chapter 7, Section 7.3 Bayes' Theorem
- Chapter 7, Section 7.4 Expected Value and Variance

Probability Distribution

Let s be the sample space of an experiment with a finite or countable number of outcomes. We assign a probability $p(s)$ to each outcome. We require that two conditions be met:

1. $0 \leq p(s) \leq 1$ for each $s \in S$
2. $\sum_{s \in S} p(s) = 1$.

The function p from the set of all outcomes of the sample space S is called a **probability distribution.**

Conditional Probability

Given an event F occurs, the probability that event E occurs is the **conditional probability** of E given F.

Let E and F be events with $p(F) > 0$. The **conditional probability** of E given F, denoted by $p(E|F)$, is defined as:

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

.

Conditional Probability

In a best 2 out of 3 serious, the probability of wining first game is 1/2, the probability of winning a game following a win is 2/3, the probability of winning after a loss is 1/3. What is the probability of winning the series given a first win?

Conditional Probability

A bit string of length four is generated at random so that each of the 16 bit strings of length four is equally likely. What is the probability that it contains at least two consecutive 0s, given that its first bit is a 0?

Independence

When two events are independent, the occurrence of one of the events gives no information about the probability of that the other event occurs.
The events E and F are independent **if and only if**

$$p(E \cap F) = p(E)p(F)$$

.

Independence

Roll two fair independent coins, let A be the event that both coins match, let B be the event that the first coin is Head. Are A and B independent? In what circumstances, these two events are independent?

Independence

Rolling two fair regular dice (six sided with values 1, 2,. . . , 6 appearing equally likely), let D_1 and D_2 be the face that comes up for the first and second die respectively. Let S be the sum of the two dice.

Independence

Suppose E is the event that a randomly generated bit string of length four begin with a 1 and F is the event that this bit string contains an even number of 1s. Are E and F independent, if the 16 bit strings of length four are equally likely?

Bernoulli Trials and Binomial Distribution

Each performance of an experiment with two possible outcomes is called a **Bernoulli trial.**

In general, a possible outcome of a Bernoulli trial is called a **success** or a **failure.**

- Generate a bit, $\{\ 0, 1\}$.
- Flip a coin, $\{$Heads, tails$\}$.

If p is the probability of a success and q is the probability of a failure, it follows that $p + q = 1$.

Binomial Distribution

- The probability of exactly k successes in n independent Bernoulli trials, with probability of success p and probability of failure $q = 1 - p$, is $C(n, k)p^k q^{n-k}$.
- We denote by $b(k; n, p)$ the probability of k successes in n independent Bernoulli trials with probability of success p and probability of failure $q = 1 - p$.
- Considered as a function of k, we call this function the **binomial distribution.**

$$b(k; n, p) = C(n, k)p^k q^{n-k}.$$

Binomial Distribution

Example

A coin is biased so that the probability of heads is 2/3. What is the probability that exactly four heads come up when the coin is flipped seven times, assuming that the flips are independent?

Example

An airline on average assumes that just 95% of all ticket purchasers actually show up for a flight. If the airline sells 105 tickets for a 100 seat flight, what is the probability that a flight is overbooked?

Bayes' Theorem

Suppose we know p(F), the probability that an event F occurs, but we have knowledge that an event E occurs.
The conditional probability that F occurs given that E occurs, $p(F|E)$

Bayes' Theorem

Bayes' Theorem

Suppose that E and F are events from a sample space S such that $P(E) \neq 0$ and $P(F) \neq 0$. Then

$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\overline{F})p(\overline{F})}$$

Generalizing Bayes' Theorem

Suppose that E is an event from a sample space S and that $F_1, F_2, ..., F_n$ are mutually exclusive events such that $\cup_{i=1}^n F_i = S$. Assume that $p(E) \neq 0$ and $p(F_i) \neq 0$ for $i = 1, 2, ..., n$. Then

$$p(F_j|E) = \frac{p(E|F_j)p(F_j)}{\sum_{i=1}^n p(E|Fi)p(Fi).}$$

Bayesian Spam Filters

Suppose that we have found that the word ”Rolex” occurs in 250 of 2000 messages known to be spam and in 5 of 1000 messages known not to be spam. Estimate the probability that an incoming message containing the word “Rolex” is spam, assuming that it is equally likely that an incoming message is spam or not spam. If our threshold for rejecting a message as spam is 0.9, will we reject such messages?

Bayesian Spam Filters

Suppose that we train a Bayesian spam filter on a set of 2000 spam messages and 1000 messages that are not spam. The word “stock” appears in 400 spam messages and 60 messages that are not spam, and the word “undervalued” appears in 200 spam messages and 25 messages that are not spam. Estimate the probability that an incoming message containing both the words “stock” and “undervalued” is spam, assuming that we have no prior knowledge about whether it is spam. Will we reject such messages as spam when we set the threshold at 0.9?

Homework

Expected Value and Variance

The **expected value**, also called the expectation or mean, of the random variable X on the sample space S:

$$E(X) = \sum_{s \in S} p(s)X(S)$$

If X is a random variable and $p(X = r)$ is the probability that $X = r$, so that $p(X = r) = \sum_{s \in S, X(s) = r} p(s)$, then

$$E(X) = \sum_{r \in X(S)} p(X = r)r$$

Expected Value

The expected number of successes when n mutually independent Bernoulli trials are performed, where p is the probability of success on each trial, is np.

Linearity of Expectations

If X_i, $i = 1, 2, ..., n$ with n a positive integer, are random variables on S, and if a and b are real numbers, then

1. $E(X_1 + X_2 + \cdots + Xn) = E(X_1) + E(X_2) + \cdots + E(X_n)$
2. $E(aX + b) = aE(X) + b.$

Example: Rolling a Die

Let X be the number that comes up when a fair die is rolled. What is the expected value and variance of X?

Geometric Distribution

A random variable X has a geometric distribution with parameter p if $p(X = k) = (1 - p)^{k-1}p$ for $k = 1, 2, 3, \ldots$, where p is a real number with $0 \leq p \leq 1$.

$$E(X) = 1/p$$

Geometric Distribution

Suppose that the probability that a coin comes up tails is p. This coin is flipped repeatedly until it comes up tails. What is the expected number of flips until this coin comes up tails?
The random variable X that equals the number of flips expected before a coin comes up tails is an example of a random variable with a geometric distribution.

Independent Random Variables

The random variables X and Y on a sample space S are independent if

$$p(X = r1 \text{ and } Y = r2) = p(X = r1) \cdot p(Y = r2)$$

or, if the probability that $X = r1$ and $Y = r2$ equals the product of the probabilities that $X = r1 and Y = r2$, for all real numbers $r1$ and $r2$.

Variance

Variance provides a measure of how widely X is distributed about its expected value.

Definition

Let X be a random variable on a sample space S. The variance of X, denoted by $V(X)$, is

$$V(x) = \sum_{s \in S} (X(s) - E(X))^2 p(s).$$

That is, V (X) is the weighted average of the square of the deviation of X. The standard deviation of X, denoted $\sigma(X)$, is defined to be $\sqrt{V(X)}$.

Variance

Theorem

If X is a random variable on a sample space S, then
$V(X) = E(X^2) - E(X)^2$

Corollary

If X is a random variable on a sample space S and $E(X) = \mu$, then
$V(X) = E((X - \mu)^2)$.

How to prove it?

Example

What is the variance of the random variable $X((i,j)) = 2i$, where i is the number appearing on the first die and j is the number appearing on the second die, when two fair dice are rolled?

Discrete Probability

- Uniform (e.g. rolling a die)
- Bernoulli (Success or Failure)
- Binomial (Number of successes in fixed number of trials)
- Geometric (Number of trials until success)
- Poisson (Number of arrivals in fixed time interval)
- Many others...

Reading

Kenneth H. Rosen, *Discrete Mathematics and Its Applications*, 7th Edition, 2013.

- Chapter 7, Section 7.2 Probability Theory
- Chapter 7, Section 7.3 Bayes' Theorem
- Chapter 7, Section 7.4 Expected Value and Variance