Lecture 7
Probability III
Text: Chapter 4

STAT 8010 Statistical Methods I
September 10, 2020

Whitney Huang
Clemson University

Notes

Agenda

1. Law of Total Probability
2. Bayes' Rule
3. Random Variables

Notes

Review: Independence and Conditional Probability

Conditional Probability

Let A and B be events. The probability that event B occurs given (knowing) that event A occurs is called a conditional probability and is denoted by $P(B|A)$. The formula of conditional probability is

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Independent events

Suppose $P(A) > 0,\ P(B) > 0$. We say that event B is independent of event A if the occurrence of event A does not affect the probability that event B occurs.

$$\boxed{P(B|A) = P(B) \Rightarrow P(B \cap A) = P(B)P(A)}$$

Notes

Law of Partitions & Multiplication Rule

Law of partitions

Let $A_1, A_2, \cdots, A_k$ form a partition of Ω. Then, for all events B,

$$\mathbb{P}(B) = \sum_{i=1}^{k} \mathbb{P}(A_i \cap B)$$

Multiplication rule

- 2 events:

$$\mathbb{P}(B \cap A) = \mathbb{P}(A) \times \mathbb{P}(B|A) = \mathbb{P}(B) \times \mathbb{P}(A|B)$$

- More than 2 events:

$$\mathbb{P}(\cap_{i=1}^{n} A_i) = \mathbb{P}(A_1) \times \mathbb{P}(A_2|A_1) \times \mathbb{P}(A_3|A_1 \cap A_2) \times \cdots \times \mathbb{P}(A_n|A_{n-1} \cap \cdots \cap A_1)$$

Notes

Law of Total Probability

Let $A_1, A_2, \cdots, A_k$ form a partition of Ω. Then, for all events B,

$$\mathbb{P}(B) = \underbrace{\sum_{i=1}^{k} \mathbb{P}(A_i \cap B)}_{\text{Law of partitions}} = \underbrace{\sum_{i=1}^{k} \mathbb{P}(B|A_i) \times \mathbb{P}(A_i)}_{\text{Multiplication rule}}$$

Notes

Example

Suppose that two factories supply light bulbs to the market. Factory X's bulbs work for over 5000 hours in 99% of cases, whereas factory Y's bulbs work for over 5000 hours in 95% of cases. It is known that factory X supplies 60% of the total bulbs available and Y supplies 40% of the total bulbs available. What is the chance that a purchased bulb will work for longer than 5000 hours?

Notes

Bayes' Rule: Motivating example

The Monty Hall Problem

There was an old television show called Let's Make a Deal, whose original host was named Monty Hall. The set–up is as follows. You are on a game show and you are given the choice of three doors. Behind one door is a car, behind the others are goats. You pick a door, and the host, who knows what is behind the doors, opens another door (not your pick) which has a goat behind it. Then he asks you if you want to change your original pick. The question we ask you is, "Is it to your advantage to switch your choice?"

Notes

The Monty Hall Problem

Notes

The Monty Hall Problem Solution

Notes

Bayes' Rule

General form

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

Let $A_1, A_2, \cdots, A_k$ form a partition of the sample space. Then for every event B in the sample space,

$$\boxed{\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j) \times \mathbb{P}(A_j)}{\sum_{i=1}^{k} \mathbb{P}(B|A_i) \times \mathbb{P}(A_i)}}, \; j = 1, 2, \cdots, k$$

Notes

Example

Let us assume that a specific disease is only present in 5 out of every 1,000 people. Suppose that the test for the disease is accurate 99% of the time a person has the disease and 95% of the time that a person lacks the disease. What is the probability that the person has the disease given that they tested positive?

Solution.

$$\mathbb{P}(D|+) = \frac{\mathbb{P}(D\cap+)}{\mathbb{P}(+)} = \frac{.005 \times .99}{.005 \times .99 + .995 \times .05} = \frac{.00495}{.0547} = .0905$$

The reason we get such a surprising result is because the disease is so rare that the number of false positives greatly outnumbers the people who truly have the disease.

Notes

Review: Probability Rules

- $0 \le \mathbb{P}(A) \le 1$ for any event A, $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$
- Complement rule: $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$
- General addition rule: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- Multiplication rule: $\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \times \mathbb{P}(B) = \mathbb{P}(B|A) \times \mathbb{P}(A)$
- Conditional probability: $\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$
- Law of total probability: $\mathbb{P}(B) = \sum_{i=1}^{k} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{k} \mathbb{P}(B|A_i) \times \mathbb{P}(A_i)$
- Independence: if A and B are independent, then $\mathbb{P}(A|B) = \mathbb{P}(A)$, $\mathbb{P}(B|A) = \mathbb{P}(B)$, and $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

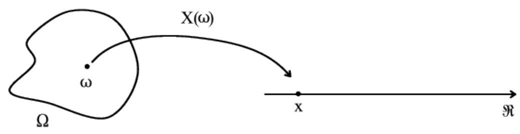
Notes

Random Variables

A random variable is a real–valued function whose domain is the sample space of a random experiment. In other words, a random variable is a function

$$X : \Omega \mapsto \mathbb{R}$$

where Ω is the sample space of the random experiment under consideration and $\mathbb{R}$ represents the set of all real numbers.



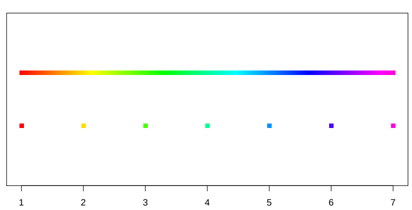
Notes

Discrete and Continuous Random Variables

There are two main types of quantitative random variables (r.v.s): discrete and continuous. A discrete r.v. often involves a count of something.

Discrete random variable

A random variable X is called a discrete random variable if the outcome of the random variable is limited to a countable set of real numbers (usually integers).



Notes

Example

The following is a chart describing the number of siblings each student in a particular class has.

Siblings	Frequency	Relative Frequency
0	8	.200
1	17	.425
2	11	.275
3	3	.075
4	1	.025
Total	40	1

Let's define the event A as the event that a randomly chosen student has 2 or more siblings. What is $\mathbb{P}(A)$?

Solution.

$$\begin{aligned}\mathbb{P}(A) = \mathbb{P}(X \geq 2) &= \mathbb{P}(X = 2) + \mathbb{P}(X = 3) + \mathbb{P}(X = 4) \\ &= .275 + .075 + .025 = .375\end{aligned}$$

Notes

Probability Mass Function

Let X be a discrete random variable. Then the probability mass function (pmf) of X is the real–valued function defined on $\mathbb{R}$ by

$$p_X(x) = \mathbb{P}(X = x)$$

The capital letter, X, is used to denote random variable. Lowercase letter, x, is used to denote possible values of the random variable.

> $p_X(x)$: The probability that the discrete random variable X is exactly equal to x.

Notes

Probability Mass Function Example

Flip a fair coin 3 times. Let X denote the number of heads tossed in the 3 flips. Create a pmf for X

Solution.

The random variable X maps any outcome to an integer (e.g. X ((T, T, T)) $= 0, X((H, H, T)) = 2$)

x	0	1	2	3
$p_X(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Notes

Properties of a PMF

- $0 \le p_X(x) \le 1, \ x \in \{0, 1, 2, \cdots\}$
- $\sum_x p_X(x) = 1$

Notes

Example

Let X be a random variable with pmf defined as follows:

$$p_X(x) = \begin{cases} k(5-x) & \text{if } x = 0,1,2,3,4 \\ 0 & \text{otherwise} \end{cases}$$

1. Find the value of k that makes $p_X(x)$ a legitimate pmf.
2. What is the probability that X is between 1 and 3 inclusive?
3. If X is not 0, what is the probability that X is less than 3?

Notes

Mean of Discrete Random Variables

The mean of a discrete r.v. X, denoted by $\mathbb{E}[X]$, is defined by

$$\mathbb{E}[X] = \sum_x x \times p_X(x)$$

Remark:

The mean of a discrete r.v. is a weighted average of its possible values, and the weight used is its probability. Sometimes we refer to the expected value as the expectation (expected value), or the first moment.

For any function, say $g(X)$, we can also find an expectation of that function. It is

$$\mathbb{E}[g(X)] = \sum_x g(x) \times p_X(x)$$

Example

$$\mathbb{E}[X^2] = \sum_x x^2 \times p_X(x)$$

Notes

Properties of Mean

Let X and Y be discrete r.v.s defined on the same sample space and having finite expectation (i.e. $\mathbb{E}[X], \mathbb{E}[Y] < \infty$). Let a and b be constants. Then the following hold:

- $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
- $\mathbb{E}[aX+b] = a \times \mathbb{E}[X] + b$

Notes

Number of Siblings Example Revisited

Siblings (X)	Frequency	Relative Frequency
0	8	.200
1	17	.425
2	11	.275
3	3	.075
4	1	.025
Total	40	1

Find the expected value of the number of siblings

Solution.

$\mathbb{E}[X] = \sum_x x p_X(x) =$
$0 \times .200 + 1 \times .425 + 2 \times .275 + 3 \times .075 + 4 \times .025 = 1.3$

Notes

Variance/Standard Deviation of Discrete r.v.'s

The **variance** of a (discrete) r.v., denoted by $Var(X)$, is a measure of the spread, or variability, in the r.v. $Var(X)$ is defined by

$$\boxed{Var(X) = \mathbb{E}[(X - \mathbb{E}[x])^2]}$$

or

$$\boxed{Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2}$$

The **standard deviation**, denoted by $sd(X)$, is the square root of its variance

Notes

Properties of Variance

Let c be a constant. Then the following hold:

- $Var(cX) = c^2 \times Var(X)$
- $Var(X + c) = Var(X)$

Notes

Example

Suppose X and Y are random variables with $\mathbb{E}[X] = 3$, $\mathbb{E}[Y] = 4$ and $Var(X) = 4$. Find:

1. $\mathbb{E}[2X + 1]$
2. $\mathbb{E}[X - Y]$
3. $\mathbb{E}[X^2]$
4. $\mathbb{E}[X^2 - 4]$
5. $\mathbb{E}[(X - 4)^2]$
6. $Var(2X - 4)$

Notes

Notes

Notes