Conditional Probability and Independence, chapter 3.1-3.3

Grethe Hystad

February 3, 2011

Conditional Probability, chapter 3.1

In excercise 3.2, the following table is given, which displays the results of a study, by educational level, of those who have smoked a cigarette within the past year for persons aged 26 and older.

Education	Smoked	Have not smoked	Total
$<$ High School Diploma	10,393	19,472	29,865
High School Graduate	17,798	39,247	57,045
Some College	13,463	30,969	44,432
College Graduate	8,320	43,357	51,677
Total	49,974	133,045	183,019

- The percentage of people who has smoked a cigarette within the past year is $\frac{49{,}974}{183{,}019} * 100 = 27.3$.

Conditional Probability, chapter 3.1

In excercise 3.2, the following table is given, which displays the results of a study, by educational level, of those who have smoked a cigarette within the past year for persons aged 26 and older.

Education	Smoked	Have not smoked	Total
$<$ High School Diploma	10,393	19,472	29,865
High School Graduate	17,798	39,247	57,045
Some College	13,463	30,969	44,432
College Graduate	8,320	43,357	51,677
Total	49,974	133,045	183,019

- The percentage of people who has smoked a cigarette within the past year is $\frac{49{,}974}{183{,}019} * 100 = 27.3$.
- The percentage of College Graduates who has smoked a cigarette within the past year is $\frac{8{,}320}{51{,}677} * 100 = 16.1$.

Conditional Probability

- Suppose a fair coin is tossed twice. The sample space is $S = \{HH, HT, TH, TT\}$. The probability that both the tosses result in a head is $\frac{1}{4}$.

Conditional Probability

- Suppose a fair coin is tossed twice. The sample space is $S = \{HH, HT, TH, TT\}$. The probability that both the tosses result in a head is $\frac{1}{4}$.
- Suppose that we are told at least one of the tosses is a head. What is the probability that both the tosses is a head? The sample space is now reduced to $S_R = \{HH, HT, TH\}$. Since the outcomes are equally likely, the probability that both tosses is a head is now $\frac{1}{3}$.

Conditional Probability

- Let A and B be the events
  - A: both tosses are head.
  - B: at least one of the tosses is a head.

$A = \{HH\}$
$B = \{HH, HT, TH\}$.
$A \cap B = \{HH\}$.

Conditional Probability

- Let A and B be the events
  - A: both tosses are head.
  - B: at least one of the tosses is a head.

  $A = \{HH\}$
  $B = \{HH, HT, TH\}$.
  $A \cap B = \{HH\}$.
- $P(A \, given \, B)$ written $P(A \mid B) = \frac{1}{3}$.

Conditional Probability

- Let A and B be the events
  - A: both tosses are head.
  - B: at least one of the tosses is a head.

$A = \{HH\}$
$B = \{HH, HT, TH\}$.
$A \cap B = \{HH\}$.

- $P(A \; given \; B)$ written $P(A \mid B) = \frac{1}{3}$.
- Notice

$$P(A \mid B) = \frac{1}{3} = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{3/4}.$$

Conditional Probability

Definition

If A and B are any two events, then the conditional probability of A given B, denoted $P(A \mid B)$, is

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

provided that $P(B) > 0$.

Notice:
If it is given that B occurs, this will affect the probability of A, and all the elementary events which do not occur should be excluded, and thus only the ones in B should be left.

Axioms of probability

Conditional probability satisfy the three axioms of probability:

- $0 \leq P(A \mid B) \leq 1$
- $P(S \mid B) = 1$
- If $A_1, A_2, \ldots,$ are mutually exclusive events, then so are $A_1B, A_2B, \ldots,$ and

$$P\left(\cup_{i=1}^{\infty} (A_i \mid B)\right) = \sum_{i=1}^{\infty} P(A_i \mid B).$$

Example, Exercise 3.2

Example

From the first table, what is the probability that a randomly selected person who has smoked a cigarette in the past year has completed college?

Example, Exercise 3.2

Example

From the first table, what is the probability that a randomly selected person who has smoked a cigarette in the past year has completed college?

Solution

Let A and B denote the events, A: completed College
B: smoked a cigarette in the past year.
Then $P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} = \frac{0.161 * 0.282}{0.273} = 0.166.$

Example, Exercise 3.19

Example

Suppose that the probability that a student passes a test the first time is 0.8. For those who fail the first time, the probability of passing the test the second time is 0.6.
(A) Find the probability that a randomly selected student passes the test.
(B) If the student passes the test, what is the probability that she or he did so on the first try?

Solution

Solution

(A) Let A_i and B be the events,
A_i: student passes the test the i^{th} time.
B: student passes the test.

$$\begin{aligned} P(B) = P(A_1) + P(A_2 \cap \bar{A}_1) &= P(A_1) + P(A_2 \mid \overline{A}_1)P(\overline{A}_1) \\ &= 0.80 + 0.60 * 0.20 \\ &= 0.92. \end{aligned}$$

Solution

Solution

(B)

$$P(A_1 \mid B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{P(A_1)}{P(B)} = \frac{0.80}{0.92} = 0.87$$

Problems

Problem

Suppose we have two Urns, Urn I and II, which contains balls. Urn I contains 7 *green balls and* 2 *red balls.*
Urn II contains 3 *green balls and* 5 *red balls.*
Select a urn at random and select a ball from its urn. What is the probability that the selected ball is red?

Problems

Problem

Suppose we have two Urns, Urn I and II, which contains balls. Urn I contains 7 green balls and 2 red balls.
Urn II contains 3 green balls and 5 red balls.
Select a urn at random and select a ball from its urn. What is the probability that the selected ball is red?

Problem

Now consider the previous problem, but select instead two balls without replacing the first.
(A) What is the probability of selecting two red balls from Urn II?
(B) What is the probability that the second ball selected is red?

Screening

	$+$	$-$	Sum
Test result $+$	a	b	$a+b$
Test result $-$	c	d	$c+d$
Sum	$a+c$	$b+d$	$a+b+c+d=n$

$+$ presence of the disease

$-$ absence of the disease

- n people are tested.
- a really have the disease.
- b do not have the disease (false positives).
- c actually have the disease (false negatives).

Screening

- *Sensitivity* $= \frac{a}{a+c}$ which is the conditional probability of having a positive test, given that the person has the disease.
- *Specificity* $= \frac{d}{b+d}$ which represents the conditional probability of having a negative test, given that the person does not have the disease.
- *Predictive value* $= \frac{a}{a+b}$: The conditional probability that a randomly selected person actually has the disease, given that he or she tested positive.

Screening Example 1

Test 1

	+	−	Sum
Test result +	90	10	100
Test result −	10	90	100
Sum	100	100	200

- Prevalence rate is 50%
- $\textit{Sensitivity} = \frac{90}{100} = 0.90 = \textit{Specificity} = \textit{Predictive value}$
- Good test.

Screening Example 2

		+	−	Sum
Test 2	Test result +	90	1000	1090
	Test result −	10	9000	9010
	Sum	100	10,000	10,100

- Prevalence rate is $\frac{100}{10100} = 0.099$
- *Sensitivity* $= 0.90 =$ *Specificity*
- *Predictive value* $= \frac{90}{1090} = 0.083$, So only 8% of those tested positive actually have the disease.

Independence, Chapter 3.2

If the knowledge that an event B has occurred, does not result in a change of the probability of A, that is $P(A \mid B) = P(A)$, the events A and B are said to be *independent*.

Definition

Two events A and B are said to be independent iff

$$P(A \cap B) = P(A)P(B).$$

If $P(B) > 0$, this is equivalent

$$P(A \mid B) = P(A) \quad \text{or}$$

if $P(A) > 0$

$$P(B \mid A) = P(B).$$

Multiplicative Rule

Theorem

Multiplicative Rule. If A and B are any two events, then

$$P(A \cap B) = P(A)P(B \mid A) = P(B)P(A \mid B).$$

If A and B are independent, then

$$P(A \cap B) = P(A)P(B).$$

Example

Two dice are rolled and the number on the upper face is observed. Is the event that the sum of the numbers on the upper face is 7 independent of the event that the number observed on the second die is a 5?

Solution

Solution

Let A and B be the events,

- *A: sum of the numbers on the upper face is* 7*,*
- *B: number on the upper face on second die is* 5*.*
- $A = \{\{(1,6)\},\{(6,1)\},\{(2,5)\},\{(5,2)\},\{3,4\},\{4,3\}\}$,
- $B = \{\{(1,5)\},\{(2,5)\},\{(3,5)\},\{(4,5)\},\{(5,5)\},\{(6,5)\}\}$,
- $A \cap B = \{\{(2,5)\}\}$.
- *Total outcome is* $6 \times 6 = 36$.
- $P(A) = P(B) = \frac{6}{36} = \frac{1}{6}$
- $P(A \cap B) = \frac{1}{36}$
- $P(A)P(B) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} = P(A \cap B)$ *so A and B are independent events.*

Example

Example

A card is drawn from a standard deck of 52 cards. Without replacing the first card, a second card is drawn. Is the event that the second card is a diamond independent of the event that the first card is a diamond?

Solution

Solution

- *Let A_i be the event that the i^{th} card is a diamond.*

Solution

Solution

- *Let* A_i *be the event that the* i^{th} *card is a diamond.*
- *Since there are a total of* 13 *diamonds and* 12 *left if the first card drawn is a diamond, we have* $P(A_2 \mid A_1) = \frac{12}{51} = \frac{4}{17}$.

Solution

Solution

- *Let A_i be the event that the i^{th} card is a diamond.*
- *Since there are a total of 13 diamonds and 12 left if the first card drawn is a diamond, we have* $P(A_2 \mid A_1) = \frac{12}{51} = \frac{4}{17}$.
- *Now*

$$
\begin{aligned}
P(A_2) &= P(A_1)P(A_2 \mid A_1) + P(\overline{A}_1)P(A_2 \mid \overline{A}_1) \\
&= \frac{13}{52}\frac{12}{51} + \frac{39}{52}\frac{13}{51} \\
&= \frac{1}{4}.
\end{aligned}
$$

Solution

Solution

- *Let* A_i *be the event that the* i^{th} *card is a diamond.*
- *Since there are a total of* 13 *diamonds and* 12 *left if the first card drawn is a diamond, we have* $P(A_2 \mid A_1) = \frac{12}{51} = \frac{4}{17}$.
- *Now*

$$\begin{aligned} P(A_2) &= P(A_1)P(A_2 \mid A_1) + P(\overline{A}_1)P(A_2 \mid \overline{A}_1) \\ &= \frac{13}{52}\frac{12}{51} + \frac{39}{52}\frac{13}{51} \\ &= \frac{1}{4}. \end{aligned}$$

- *Hence* $P(A_2 \mid A_1) \neq P(A_2)$, *so* A_2 *and* A_1 *are not independent events.*

Independence

Theorem

If A and B are independent events each with positive probability, then they cannot be mutually exclusive.

Theorem

A and B are independent events iff $\bar{A}$ and $\bar{B}$ are independent events.

Inclusion-Exclusion Principle

Theorem

$$P(\cup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i) - \sum_{i_1<i_2} P(E_{i_1}E_{i_2}) + \cdots$$
$$+ (-1)^{k+1} \sum_{i_1<i_2<\cdots<i_k} P(E_{i_1}E_{i_2}\cdots E_{i_k}) + \cdots$$
$$+ (-1)^{n+1} P(E_1E_2\cdots E_n).$$

Exercise 3.35 from the text book

Example

A portion of an electric circuit is displayed next. The switches operate independently of each other and the probability that each relay closes when the switch is thrown is displayed by the switch. What is the probability that current will flow from s to t when the switch is thrown?

Solution

Let C_i denote the event that the i^{th} switch is closed for $i = 1, 2, 3, 4$. The events C_i are independent so for any k-tuple of distinct indicies $(i_1, i_2, ..., i_k)$, where $2 \leq k \leq 4$, we have

$$P(C_{i_1} \cap C_{i_2} \cap C_{i_3} \cap C_{i_4}) = P(C_{i_1})P(C_{i_2}) \cdots P(C_{i_k}).$$

Solution cont. Method 1

Solution

- *The event that the circuit is closed and the switch is thrown is* $C = (C_1 \cap C_2) \cup (C_3 \cup C_4)$.

Solution cont. Method 1

Solution

- *The event that the circuit is closed and the switch is thrown is* $C = (C_1 \cap C_2) \cup (C_3 \cup C_4)$.
- *Then by Inclusion-Exclusion principle,*

$$\begin{aligned}
P(C) &= P[(C_1 \cap C_2) \cup (C_3 \cup C_4)] \\
&= P(C_1 \cap C_2) + P(C_3) + P(C_4) - P(C_1 \cap C_2 \cap C_3) - \\
&- P(C_1 \cap C_2 \cap C_4) - P(C_3 \cap C_4) + P(C_1 \cap C_2 \cap C_3 \cap C_4) \\
&= P(C_1)P(C_2) + P(C_3) + P(C_4) - P(C_1)P(C_2)P(C_3) - \\
&- P(C_1)P(C_2)P(C_4) - P(C_3)P(C_4) + P(C_1)P(C_2)P(C_3)P(C_4 \\
&= 0.8 * 0.7 + 0.9 + 0.8 - 0.8 * 0.7 * 0.9 - 0.8 * 0.7 * 0.8 - \\
&- 0.9 * 0.8 + 0.8 * 0.7 * 0.9 * 0.8 = 0.9912.
\end{aligned}$$

Theorem of Total Probability and Bayes' Rule, Chapter 3.3

Definition

Events B_1, B_2,..., B_k are said to partition a sample space S if the following two conditions are satisfied:

- $B_i \cap B_j = \emptyset$ for any pair i and j with $i \neq j$.
- $B_1 \cup B_2 \cup \cdots \cup B_k = S$.

For example if the events B_1 and B_2 partition the sample space S, then for any event A in the sample space S,

$$A = AB_1 \cup AB_2,$$

where AB_1 and AB_2 are mutually exclusive events. Thus

$$P(A) = P(AB_1) + P(AB_2) = P(B_1)P(A \mid B_1) + P(B_2)P(A \mid B_2).$$

Bayes' Rule

Theorem

Theorem of Total Probability: If $B_1, B_2, ..., B_k$ is a collection of mutually exclusive and exhaustive events, then for any event A,

$$P(A) = \sum_{i=1}^{k} P(B_i)P(A \mid B_i).$$

Theorem

Bayes' Rule: If the events $B_1, B_2, ..., B_k$ form a partition of the sample space S, and A is any event in S, then

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{\sum_{j=1}^{k} P(A \mid B_j)P(B_j)}.$$

Example from R. Schraffer's book

Example

A diagnostic test for a certain disease is said to be 90% accurate; that is, if a person has the disease, the test will detect it with probability 0.9. Moreover, if a person does not have the disease, the test will report that he or she doesn't have it with probability 0.9. Only 1% of the popuIation has the disease in question. If the diagnostic test reports that a person chosen at random from the population has the disease, what is the conditional probability that the person does in fact, have the disease? Are you surprised by the size of the answer? Do you consider this diagnostic test reliable?

Solution

Solution

Here the sensitivity and specificity is 0.90. *The prevalence is* 0.01.
Let S, N, TS, *and* TN *be the events,*
S: *has the disease*
TS: *positive test result*
Then by Bayes' Rule,

$$P(S \mid TS) = \frac{P(S \cap TS)}{P(TS)} = \frac{P(TS \mid S)P(S)}{P(TS \mid S)P(S) + P(TS \mid \bar{S})P(\bar{S})}$$
$$= \frac{0.90 * 0.01}{0.90 * 0.01 + 0.10 * 0.99} = 0.0833.$$

Hence the predictive value is low.

Excersie 3.35, Method 2

Solution

By Theorem of Total Probability,

$$P(C) = P(C \mid C_1 \cap C_2)P(C_1 \cap C_2) + P(C \mid \overline{C_1 \cap C_2})P(\overline{C_1 \cap C_2})$$
$$= P(C \mid C_1 \cap C_2)P(C_1)P(C_2) + P(C \mid \overline{C_1 \cap C_2})(1 - P(C_1 \cap C_2)).$$

If C_1 and C_2 are closed, the current will flow, so $P(C \mid C_1 \cap C_2) = 1$. If C_1 and C_2 are both closed, C_3 and C_4 must be open in order for the current to flow, i.e.,

$$P(C \mid \overline{C_1 \cap C_2}) = P(C_3 \cup C_4) = P(C_3) + P(C_4) - P(C_3 \cap C_4).$$

Excersie 3.35, Solution Cont.

Solution

Hence

$$
\begin{aligned}
P(C) &= P(C_1)P(C_2) + (P(C_3) + P(C_4) \\
&\quad - P(C_3)P(C_4)) * (1 - P(C_1)P(C_2)) \\
&= 0.8 * 0.7 + (0.9 + 0.8 - 0.9 * 0.8) * (1 - 0.8 * 0.7) \\
&= 0.9912.
\end{aligned}
$$

Problem

Lightbulbs are manufactored by three factories, I, II, and III.

- 10% *of the bulbs were made in factory I*
- 40% *of the bulbs were made in factory II*
- 50% *of the bulbs were made in factory III.*

Suppose that

- 2% *of the lightbulbs produced by I are defective*
- 5% *of the lightbulbs produced by II are defective*
- 3% *of the lightbulbs produced by III are defective.*

Assume that an lightbulb selected at random from the stockpile is observed to be defective.
What is the probability that the lightbulb came from factory I?

Problem

Suppose that a box contains 4 red balls and 2 green balls. Suppose that 3 balls are removed at random from the box without replacement. Let R_i and G_i be the events that the ball drawn on the i^{th} draw is red and green, respectively.
Find
(i) $P(R_3)$
(ii) $P(R_1 G_3)$

In general (exercise 3.39), suppose that a box contains n balls, of which k are white. Suppose that m balls is removed at random from the box without replacement. Let A_j, where $j = 1, 2, ..., m$ denote the event that the ball drawn on the j^{th} draw is white. Then $P(A_j) = \frac{k}{n}$.