Tutorial on Basic Probability

Doron ZEILBERGER

IMPORTANT THEOREM (Principle of Inclusion-Exclusion for Two Events)

If A and B are events then

$$P(A^c \cap B^c) = 1 - P(A) - P(B) + P(A \cap B)$$

Proof: To figure out the probability that A did not happen and B did not happen, the first approximation is $1 - P(A) - P(B)$, but in doing this we subtracted too much, all those possibilities that both A and B happened, namely $A \cap B$ got removed twice, so we have to rectify it by adding $P(A \cap B)$.

REMARK A second version of the Principle of Inclusion-Exclusion for Two Sets is:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad .$$

Example A: In a certain class %30 smoke, %60 drink, and %20 smoke and drink, how many people neither smoke nor drink?

Solution to Example A: Let A be the event that someone smokes, and B be the event that someone drinks. Then

$$P(A) = 0.3 \quad , \quad P(B) = 0.6 \quad , P(A \cap B) = 0.2 \quad , hence$$

$$P(A^c \cap B^c) = 1 - 0.3 - 0.6 + 0.2 = 1 - 0.9 + 0.2 = 0.3 \quad .$$

Answer to Example A: %30 of the students neither smoke nor drink.

Do right now

Exercise 1: In a certain class %30 play basketball, %40 play soccer, and %10 play soccer and basketball, what percentage of the class plays neither soccer nor basketball?

Exercise 2: In a Labor Union in a small town, %25 have made at least one campaign contribution, %40 have volunteered for a campaign, and %10 did both. What percentage did neither?

In the above problems, you were given $P(A)$, $P(B)$ and $P(A \cap B)$, and had to figure out $P(A^c \cap B^c)$, but in the next type of problem, you are given $P(A), P(B)$ and $P(A^c \cap B^c)$ and have to figure out $P(A \cap B)$.

Example B: In a certain class %30 smoke, %60 drink, and %30 neither smoke nor drink, how many people smoke and drink?

Solution to Example B: Let A be the event that someone smokes, and B be the event that someone drinks. Then by (PIE)

$$P(A^c \cap B^c) = 1 - P(A) - P(B) + P(A \cap B)$$

Let's plug-in the data

$$P(A) = 0.3 \quad , \quad P(B) = 0.6 \quad , P(A^c \cap B^c) = 0.3 \qquad .$$

$$0.3 = 1 - 0.3 - 0.6 + P(A \cap B)$$

Solving for $P(A \cap B)$ we get

$$P(A \cap B) = 0.3 - 1 + 0.3 + 0.6 = 0.2 \quad .$$

Answer to Example B: %20 of the students smoke and drink.

Do right now

Exercise 3: In a certain class %30 play basketball, %40 play soccer, and %40 play neither soccer nor basketball, what percentage of the class plays both soccer and basketball?

Exercise 4: In a Labor Union in a small town, %25 have made at least one campaign contribution, %40 have volunteered for a campaign, and %45 did neither. What percentage did both?

Exercise 5: In an environmental Club in a small town, %20 have made at least one campaign contribution, %10 have volunteered for a campaign, and %75 did neither. What percentage did both?

Exercise 6: In the Democratic Club in a small town, %60 have made at least one campaign contribution, %50 have volunteered for a campaign, and %10 did neither. What percentage did both?

Example C: In a certain class %30 smoke, %60 drink, and %30 neither smoke nor drink, what is the probability that a student either smokes or drinks (or both).

Solution to Example C (first way): $P(A \cup B) = 1 - P(A^c \cap B^c) = 1 - 0.3 = 0.7$

Solution to Example C (second way): From Example B, we know that $P(A \cap B) = 0.2$, so $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.3 + 0.6 - 0.2 = 0.7$

Ans. to Example C: The probability that a student either smokes or drinks (or both) is 0.7.

Exercise 7: In a Labor Union in a small town, %25 have made at least one campaign contribution, %40 have volunteered for a campaign, and %45 did neither. What percentage did both?

Exercise 8: In an environmental Club in a small town, %20 have made at least one campaign contribution, %10 have volunteered for a campaign, and %75 did neither. What percentage did either contribute or participate (or both?)

Exercise 9: In the Democratic Club in a small town, %60 have made at least one campaign contribution, %50 have volunteered for a campaign, and %10 did neither. What percentage did either contribute or participate (or both?)

Important Formula (CONDITIONAL PROBABILITY)

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

In words: the probability that A happened given that we know that B happened is the probability that both happened divided by the probability that B happened.

Example D: In a certain class %30 smoke, %60 drink, and %20 smoke and drink.

(a) what is the probability that a student smokes given that he or she drinks?

(b) what is the probability that a student drinks given that he or she smokes?

Solution to Example D: Let A be the event that a student smokes, and B be the event that he drinks.

Solution to 4(a):

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{.2}{.6} = \frac{1}{3}$$

Solution to 4(b):

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{.2}{.3} = \frac{2}{3} \quad .$$

Ans. to Example D: The probability that a student also smokes given that he or she drinks is $\frac{1}{3}$. The probability that a student also drinks given that he or she smokes is $\frac{2}{3}$.

Exercise 10: In a Labor Union in a small town, %25 have made at least one campaign contribution, %40 have volunteered for a campaign, and %45 did neither.

(a) What is the probability that she volunteered given that she made a contribution?

(b) What is the probability that she made a contribution given that she volunteered?

(Hint: You must first find the probability that she or he both made a contribution and volunteered (done in a previous exercise))

Exercise 11: In an environmental Club in a small town, %20 have made at least one campaign contribution, %10 have volunteered for a campaign, and %75 did neither.

(a) What is the probability that she volunteered given that she made a contribution?

(b) What is the probability that she made a contribution given that she volunteered?

(Hint: You must first find the probability that she or he both made a contribution and volunteered (done in a previous exercise))

Exercise 12: In the Democratic Club in a small town, %60 have made at least one campaign contribution, %50 have volunteered for a campaign, and %10 did neither.

(a) What is the probability that she volunteered given that she made a contribution?

(b) What is the probability that she made a contribution given that she volunteered?

(Hint: You must first find the probability that she or he both made a contribution and volunteered (done in a previous exercise))

Important Principle

There are two phases. In phase one there are n possibilities, $A_1, A_2, \ldots, A_n$, whose probabilities are $P(A_1), \ldots, P(A_n)$ respectively. In the second phase, a certain event, B, may or may not happen. We are given the individual conditional probabilities

$$P(B|A_1) \quad , \quad P(B|A_2) \quad , \quad P(B|A_n)$$

Then

$$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \ldots + P(A_n) \cdot P(B|A_n) \quad .$$

The above spelled-out for three possibilities in the first phase is

$$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + P(A_3) \cdot P(B|A_3) \quad .$$

Example E: In a certain town, %20 of the population is 18 or under, %40 are between 19 and 50, and the remaining one are 51 or over. The probability that a person 18 or under smokes is %10. The probability that a person between 19 and 50, smokes is %30. The probability that a person 51 or over , smokes is %25.

If a person is chosen totally at random, what is the probability that he (or she) smokes?

Solution to Example E:

Let the event that the picked person is 18 or under be A_1 .

Let the event that the picked person between 19 and 50 be A_2 .

Let the event that the picked person between 51 or over be A_3 .

Let B be the event that he or she smokes.

First, we must figure out $P(A_3) = 1 - P(A_1) - P(A_2) = 1 - 0.2 - 0.4 = 0.4$ By the data

$$P(A_1) = 0.2 \quad , \quad P(B|A_1) = 0.1$$

$$P(A_2) = 0.4 \quad , \quad P(B|A_2) = 0.3$$

$$P(A_3) = 0.4 \quad , \quad P(B|A_3) = 0.25$$

Hence

$$P(B) = P(A_1)\cdot P(B|A_1)+P(A_2)\cdot P(B|A_2)+P(A_3)\cdot P(B|A_3) = 0.2\cdot 0.1+0.4\cdot 0.3+0.4\cdot 0.25 = 0.02+0.12+0.1 = 0.24$$

Ans. to Example E: The probability that a random person smokes is %24.

Do Right Now

Long Exercise 13

In a certain town every person belongs to EXACTLY one club, as follows

• %50 are Union members

• %20 are from the environmental group

• %30 are from the Democratic Club

We have the following information on each of these

• In the Union: %25 have made at least one campaign contribution, %40 have volunteered for a campaign, and %45 did neither.

• In the environmental group: %20 have made at least one campaign contribution, %10 have volunteered for a campaign, and %75 did neither.

• In the Democratic Club: %60 have made at least one campaign contribution, %50 have volunteered for a campaign, and %10 did neither.

Find out the following

What is the probability that a person, chosen at random from the town

(a) Made at least one campaign contribution

(b) Volunteered

(c) Make BOTH a campaign contribution and volunteered

(d) Made neither campaign contribution nor volunteered?

(e) Made a campaign contribution but did not volunteer? (Hint: $P(A \cap B^c) = P(A) - P(A \cap B)$)

Important Theorem (Bayes' Law

There are two phases. In phase one there are n possibilities, $A_1, A_2, \ldots, A_n$, whose probabilities are $P(A_1), \ldots, P(A_n)$ respectively. In the second phase, a certain event, B, may or may not happen. We are given the individual conditional probabilities

$$P(B|A_1) \quad , \quad P(B|A_2) \quad , \quad P(B|A_n)$$

Then we know, from the above Important Principle that

$$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \ldots + P(A_n) \cdot P(B|A_n) \quad .$$

Now , suppose that you know that B happened, then, for $i = 1, 2, \ldots, n$

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{P(B)} \quad .$$

The above spelled-out for three possibilities in the first phase is

$$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + P(A_3) \cdot P(B|A_3) \quad .$$

and the *posterior probabilities* are

$$P(A_1|B) = \frac{P(A_1) \cdot P(B|A_1)}{P(B)} \quad , \quad P(A_2|B) = \frac{P(A_2) \cdot P(B|A_2)}{P(B)} \quad , \quad P(A_3|B) = \frac{P(A_3) \cdot P(B|A_3)}{P(B)} \quad .$$

Example F: In a certain class %20 only smoke, %40 only drink, and %40 smoke and drink (no one neither smokes nor drinks)

- If you only smoke your chances of dying before the age 80 is %50
- If you only drink your chances of dying before the age 80 is %40
- If you smoke and drink your chances of dying before the age 80 is %70

If you know about someone that he died before the age of 80, what are the chances that

- He only smoked
- He only drank
- He smoke and drank

Solutions to Example F

Important First Step: Let B the event "died younger than 80", the IMPORTANT first step is to use the Fundamental Fact above to find $P(B)$.

Let A_1, A_2, A_3 be the events "only smokes", "only drinks", and "smokes and drinks". Then

$$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + P(A_3) \cdot P(B|A_3) =$$

$$0.2 \cdot 0.5 + 0.4 \cdot 0.4 + 04. \cdot 0.7 = 0.54$$

(Note: this is an important INTERMEDIATE step).

Note that the 0.54 has three contributions, coming from A_1 (only smoke), A_2 (only drink), A_3 (smoke and drink), the respective answers to the questions are the *relative contributions*

$$P(A_1|B) = \frac{0.2 \cdot 0.5}{0.2 \cdot 0.5 + 0.4 \cdot 0.4 + 04. \cdot 0.7} = \frac{0.1}{0.54} = 0.1852$$

$$P(A_2|B) = \frac{0.4 \cdot 0.4}{0.2 \cdot 0.5 + 0.4 \cdot 0.4 + 04. \cdot 0.7} = \frac{0.16}{0.54} = 0.2963$$

$$P(A_1|B) = \frac{0.4 \cdot 0.7}{0.2 \cdot 0.5 + 0.4 \cdot 0.4 + 04. \cdot 0.7} = \frac{0.28}{0.54} = 0.5185$$

Answer to Example F: If it is known that someone died before the age of 80, the probability that he only smoked is %18.52, that he only drank is %29.62, that he drank and smoke is %51.85.

(Note: the probability that he neither drank nor smoke is 0, since in this artificial problem the assumption was that everyone either smokes, drinks, or both).

Do Right Now

Very Long Exercise 14

(You may use your answers to Exercise 13, the assumptions are the same, and are repeated her for the sake of clarity).

In a certain town every person belongs to EXACTLY one club, as follows

- %50 are Union members
- %20 are from the environmental group
- %30 are from the Democratic Club

We have the following information on each of these

- In the Union: %25 have made at least one campaign contribution, %40 have volunteered for a campaign, and %45 did neither.

• In the environmental group: %20 have made at least one campaign contribution, %10 have volunteered for a campaign, and %75 did neither.

• In the Democratic Club: %60 have made at least one campaign contribution, %50 have volunteered for a campaign, and %10 did neither.

Find out the following

(a) Given that you know that he made at least one campaign contribution what are

(a)[i]: The probability that he is a Union member

(a)[ii]: The probability that he is a member of the environmental group

(a)[iii]:The probability that he is a member of the Democratic Club

(b) Given that you know that he volunteered whare are

(b)[i]: The probability that he is a Union member

(b)[ii]: The probability that he is a member of the environmental group

(b)[iii]:The probability that he is a member of the Democratic Club

(c) Given that you know that he made BOTH a campaign contribution and volunteered

(c)[i]: The probability that he is a Union member

(c)[ii]: The probability that he is a member of the environmental group

(c)[iii]:The probability that he is a member of the Democratic Club

(d) Given that you know for sure that he made neither campaign contribution nor volunteered?

(d)[i]: The probability that he is a Union member

(d)[ii]: The probability that he is a member of the environmental group

(d)[iii]:The probability that he is a member of the Democratic Club

(e) Given that you know for sure that he made a campaign contribution but did not volunteer?

(e)[i]: The probability that he is a Union member

(e)[ii]: The probability that he is a member of the environmental group

(e)[iii]:The probability that he is a member of the Democratic Club

---

Doron Zeilberger, Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA.
Email: `zeilberg at math dot rutgers dot edu` .