MT4531/MT5731: (Advanced) Bayesian Inference

Bayes' Theorem for events

Nicolò Margaritella

School of Mathematics and Statistics, University of St Andrews

–

Outline

Outline

Bayes' Theorem

- Let A and B denote possible events, such that $\mathbb{P}(B) > 0$.
- Then, Bayes' Theorem states that:

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

Bayes' Theorem

- Let A and B denote possible events, such that $\mathbb{P}(B) > 0$.
- Then, Bayes' Theorem states that:

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

Bayes' Theorem proof

- **Proof:**

By definition of conditional probability,

Bayes' Theorem proof

- **Proof:**

By definition of conditional probability,

$$\begin{aligned}\mathbb{P}(A|B) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \\ &= \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}.\end{aligned}$$

Alternative form of Bayes' Theorem (1)

- The denominator in the expression for Bayes' Theorem is most often expressed in an alternative way.
- Denoting by A^c the complement of A, by the law of total probability,

$$\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c) = \mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c)$$

Now,

$$\begin{aligned}\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} &= \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c)} \\ &= \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c)}.\end{aligned}$$

Alternative form of Bayes' Theorem (1)

- The denominator in the expression for Bayes' Theorem is most often expressed in an alternative way.
- Denoting by A^c the complement of A, by the law of total probability,

$$\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c) = \mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c)$$

Now,

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c)}.$$

Alternative form of Bayes' Theorem (2)

- More generally, suppose that we let A_i, denote a set of mutually exclusive and exhaustive events, for $i = 1, \ldots$,n.
- By the law of total probability we can express $\mathbb{P}(B)$ in the form,

$$\mathbb{P}(B) = \mathbb{P}(B \cap A_1) + \ldots + \mathbb{P}(B \cap A_n) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

- Thus, an alternative expression for Bayes theorem is given by,

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j)\mathbb{P}(A_j)}{\sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i)}.$$

Alternative form of Bayes' Theorem (2)

- More generally, suppose that we let A_i, denote a set of mutually exclusive and exhaustive events, for $i = 1, \ldots$,n.
- By the law of total probability we can express $\mathbb{P}(B)$ in the form,

$$\mathbb{P}(B) = \mathbb{P}(B \cap A_1) + \ldots + \mathbb{P}(B \cap A_n) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

- Thus, an alternative expression for Bayes theorem is given by,

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j)\mathbb{P}(A_j)}{\sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i)}.$$

Alternative form of Bayes' Theorem (2)

- More generally, suppose that we let A_i, denote a set of mutually exclusive and exhaustive events, for $i = 1, \dots,$n.
- By the law of total probability we can express $\mathbb{P}(B)$ in the form,

$$\mathbb{P}(B) = \mathbb{P}(B \cap A_1) + \ldots + \mathbb{P}(B \cap A_n) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

- Thus, an alternative expression for Bayes theorem is given by,

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j)\mathbb{P}(A_j)}{\sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i)}.$$

Outline

Example:

- A test for Covid-19 is given to a population.
- It is believed that only 0.1% have Covid-19 (from Riley et al., 2020, community prevalence in England during May 2020).
- The test is 70% accurate for people who have Covid-19 (sensitivity), and 95% accurate for those who do not have the virus (specificity). (Watson et al., 2020)
- **(i)** Given that a person has a negative test result, what is the probability they do have Covid-19 (*false negative*)?
- **(ii)** Given that a person has a positive test result, what is the probability they do not have Covid-19 (*false positive*)?
- **(iii)** what is the probability that a person has Covid-19 given a positive result (*positive predictive value*)?

Example:

- A test for Covid-19 is given to a population.
- It is believed that only 0.1% have Covid-19 (from Riley et al., 2020, community prevalence in England during May 2020).
- The test is 70% accurate for people who have Covid-19 (sensitivity), and 95% accurate for those who do not have the virus (specificity). (Watson et al., 2020)
- **(i)** Given that a person has a negative test result, what is the probability they do have Covid-19 (*false negative*)?
- **(ii)** Given that a person has a positive test result, what is the probability they do not have Covid-19 (*false positive*)?
- **(iii)** what is the probability that a person has Covid-19 given a positive result (*positive predictive value*)?

Example:

- A test for Covid-19 is given to a population.
- It is believed that only 0.1% have Covid-19 (from Riley et al., 2020, community prevalence in England during May 2020).
- The test is 70% accurate for people who have Covid-19 (sensitivity), and 95% accurate for those who do not have the virus (specificity). (Watson et al., 2020)
- **(i)** Given that a person has a negative test result, what is the probability they do have Covid-19 (*false negative*)?
- **(ii)** Given that a person has a positive test result, what is the probability they do not have Covid-19 (*false positive*)?
- **(iii)** what is the probability that a person has Covid-19 given a positive result (*positive predictive value*)?

Example:

- A test for Covid-19 is given to a population.
- It is believed that only 0.1% have Covid-19 (from Riley et al., 2020, community prevalence in England during May 2020).
- The test is 70% accurate for people who have Covid-19 (sensitivity), and 95% accurate for those who do not have the virus (specificity). (Watson et al., 2020)
- **(i)** Given that a person has a negative test result, what is the probability they do have Covid-19 (*false negative*)?
- **(ii)** Given that a person has a positive test result, what is the probability they do not have Covid-19 (*false positive*)?
- **(iii)** what is the probability that a person has Covid-19 given a positive result (*positive predictive value*)?

Example:

- A test for Covid-19 is given to a population.
- It is believed that only 0.1% have Covid-19 (from Riley et al., 2020, community prevalence in England during May 2020).
- The test is 70% accurate for people who have Covid-19 (sensitivity), and 95% accurate for those who do not have the virus (specificity). (Watson et al., 2020)
- **(i)** Given that a person has a negative test result, what is the probability they do have Covid-19 (*false negative*)?
- **(ii)** Given that a person has a positive test result, what is the probability they do not have Covid-19 (*false positive*)?
- **(iii)** what is the probability that a person has Covid-19 given a positive result (*positive predictive value*)?

Example:

- A test for Covid-19 is given to a population.
- It is believed that only 0.1% have Covid-19 (from Riley et al., 2020, community prevalence in England during May 2020).
- The test is 70% accurate for people who have Covid-19 (sensitivity), and 95% accurate for those who do not have the virus (specificity). (Watson et al., 2020)
- **(i)** Given that a person has a negative test result, what is the probability they do have Covid-19 (*false negative*)?
- **(ii)** Given that a person has a positive test result, what is the probability they do not have Covid-19 (*false positive*)?
- **(iii)** what is the probability that a person has Covid-19 given a positive result (*positive predictive value*)?

Example - Answers

- Let the event that an individual has Covid-19 be denoted by C, the event that their test result is positive be denoted by Po, and their complements be denoted C^c and Po^c respectively.
- Then, we have that,

$$
\begin{aligned}
\mathbb{P}(C) &= 0.001\\
\mathbb{P}(Po|C) &= 0.70\\
\mathbb{P}(Po^c|C^c) &= 0.95\\
\mathbb{P}(C^c) &= 0.999\\
\mathbb{P}(Po^c|C) &= 0.30\\
\mathbb{P}(Po|C^c) &= 0.05
\end{aligned}
$$

Example - Answers

- Let the event that an individual has Covid-19 be denoted by C, the event that their test result is positive be denoted by Po, and their complements be denoted C^c and Po^c respectively.
- Then, we have that,

$$\begin{aligned}
\mathbb{P}(C) &= 0.001 \\
\mathbb{P}(Po|C) &= 0.70 \\
\mathbb{P}(Po^c|C^c) &= 0.95 \\
\mathbb{P}(C^c) &= 0.999 \\
\mathbb{P}(Po^c|C) &= 0.30 \\
\mathbb{P}(Po|C^c) &= 0.05
\end{aligned}$$

Example - Answers

- (i) For the probability of a false negative,

$$\mathbb{P}(\text{Covid-19 given negative test}) \quad = \quad \mathbb{P}(C|Po^c)$$

Example - Answers

- (i) For the probability of a false negative,

$$\mathbb{P}(\text{Covid-19 given negative test}) = \mathbb{P}(C|Po^c)$$

$$\begin{aligned}
&= \frac{\mathbb{P}(Po^c \cap C)}{\mathbb{P}(Po^c)} \\
&= \frac{\mathbb{P}(Po^c \cap C)}{\mathbb{P}(Po^c \cap C) + \mathbb{P}(Po^c \cap C^c)} \\
&= \frac{\mathbb{P}(Po^c|C)\mathbb{P}(C)}{\mathbb{P}(Po^c|C)\mathbb{P}(C) + \mathbb{P}(Po^c|C^c)\mathbb{P}(C^c)} \\
&= \frac{0.30 \times 0.001}{(0.30 \times 0.001) + (0.95 \times 0.999)} \\
&= 0.00031 \text{ (Very good!)}
\end{aligned}$$

Example - Answers

- (ii) For a false positive result,

$$\mathbb{P}(\text{No Covid-19 given positive test}) = \mathbb{P}(C^c|Po)$$

Example - Answers

- (ii) For a false positive result,

$$\mathbb{P}(\text{No Covid-19 given positive test}) = \mathbb{P}(C^c|Po)$$

$$= \frac{\mathbb{P}(Po|C^c)\mathbb{P}(C^c)}{\mathbb{P}(Po|C^c)\mathbb{P}(C^c) + \mathbb{P}(Po|C)\mathbb{P}(C)}$$

$$= \frac{0.05 \times 0.999}{(0.05 \times 0.999) + (0.70 \times 0.001)}$$

$$= 0.986 \text{ (Very bad!!)}$$

- Note the high false positive probability! Why the counter intuitive result?
- The percentage of people without Covid is 99.9% which is a high percentage, in relation to the specificity ($\mathbb{P}(Po^c|C^c) = 0.95$) of the test.
- For a higher specificity, e.g. $\mathbb{P}(Po^c|C^c) = 0.99999$, we would obtain, $\mathbb{P}(\text{No Covid-19 given positive test}) = 0.014$.

- (iii) The probability that a person has the disease given a positive test is $1 - \mathbb{P}(C^c|Po) = 0.014$.

Example - Answers

- (ii) For a false positive result,

$$\mathbb{P}(\text{No Covid-19 given positive test}) = \mathbb{P}(C^c|Po)$$

$$\begin{aligned} &= \frac{\mathbb{P}(Po|C^c)\mathbb{P}(C^c)}{\mathbb{P}(Po|C^c)\mathbb{P}(C^c) + \mathbb{P}(Po|C)\mathbb{P}(C)} \\ &= \frac{0.05 \times 0.999}{(0.05 \times 0.999) + (0.70 \times 0.001)} \\ &= 0.986 \text{ (Very bad!!)} \end{aligned}$$

- Note the high false positive probability! Why the counter intuitive result?
- The percentage of people without Covid is 99.9% which is a high percentage, in relation to the specificity ($\mathbb{P}(Po^c|C^c) = 0.95$) of the test.
- For a higher specificity, e.g. $\mathbb{P}(Po^c|C^c) = 0.99999$, we would obtain, $\mathbb{P}(\text{No Covid-19 given positive test}) = 0.014$.

- (iii) The probability that a person has the disease given a positive test is $1 - \mathbb{P}(C^c|Po) = 0.014$.

Example - Answers

- (ii) For a false positive result,

$$
\begin{aligned}
\mathbb{P}(\text{No Covid-19 given positive test}) &= \mathbb{P}(C^c|Po) \\
&= \frac{\mathbb{P}(Po|C^c)\mathbb{P}(C^c)}{\mathbb{P}(Po|C^c)\mathbb{P}(C^c) + \mathbb{P}(Po|C)\mathbb{P}(C)} \\
&= \frac{0.05 \times 0.999}{(0.05 \times 0.999) + (0.70 \times 0.001)} \\
&= 0.986 \text{ (Very bad!!)}
\end{aligned}
$$

- Note the high false positive probability! Why the counter intuitive result?
- The percentage of people without Covid is 99.9% which is a high percentage, in relation to the specificity ($\mathbb{P}(Po^c|C^c) = 0.95$) of the test.
- For a higher specificity, e.g. $\mathbb{P}(Po^c|C^c) = 0.99999$, we would obtain, $\mathbb{P}(\text{No Covid-19 given positive test}) = 0.014$.
- (iii) The probability that a person has the disease given a positive test is $1 - \mathbb{P}(C^c|Po) = 0.014$.

Example - Answers

- (ii) For a false positive result,

$$\mathbb{P}(\text{No Covid-19 given positive test}) = \mathbb{P}(C^c|Po)$$

$$= \frac{\mathbb{P}(Po|C^c)\mathbb{P}(C^c)}{\mathbb{P}(Po|C^c)\mathbb{P}(C^c) + \mathbb{P}(Po|C)\mathbb{P}(C)}$$

$$= \frac{0.05 \times 0.999}{(0.05 \times 0.999) + (0.70 \times 0.001)}$$

$$= 0.986 \text{ (Very bad!!)}$$

- Note the high false positive probability! Why the counter intuitive result?
- The percentage of people without Covid is 99.9% which is a high percentage, in relation to the specificity ($\mathbb{P}(Po^c|C^c) = 0.95$) of the test.
- For a higher specificity, e.g. $\mathbb{P}(Po^c|C^c) = 0.99999$, we would obtain, $\mathbb{P}(\text{No Covid-19 given positive test}) = 0.014$.

- (iii) The probability that a person has the disease given a positive test is $1 - \mathbb{P}(C^c|Po) = 0.014$.

Example - Answers

- (ii) For a false positive result,

$$\begin{aligned}\mathbb{P}(\text{No Covid-19 given positive test}) &= \mathbb{P}(C^c|Po)\\ &= \frac{\mathbb{P}(Po|C^c)\mathbb{P}(C^c)}{\mathbb{P}(Po|C^c)\mathbb{P}(C^c)+\mathbb{P}(Po|C)\mathbb{P}(C)}\\ &= \frac{0.05\times 0.999}{(0.05\times 0.999)+(0.70\times 0.001)}\\ &= 0.986 \text{ (Very bad!!)}\end{aligned}$$

- Note the high false positive probability! Why the counter intuitive result?
- The percentage of people without Covid is 99.9% which is a high percentage, in relation to the specificity ($\mathbb{P}(Po^c|C^c) = 0.95$) of the test.
- For a higher specificity, e.g. $\mathbb{P}(Po^c|C^c) = 0.99999$, we would obtain, $\mathbb{P}(\text{No Covid-19 given positive test}) = 0.014$.

- (iii) The probability that a person has the disease given a positive test is $1 - \mathbb{P}(C^c|Po) = 0.014$.

Example - Answers

- Let $Po2$ be the event that a second test is also positive.
- For some testing procedures, it may be justifiable to assume that the two tests are independent, given the disease status and the problem specifications. We will make this assumption for this problem.
- **Note**: It may not be sensible to assume the two tests are independent for unknown disease status, as learning about the outcome of the first test could affect our beliefs on the probabilities for the outcome of the second test.

Example - Answers

- Let $Po2$ be the event that a second test is also positive.
- For some testing procedures, it may be justifiable to assume that the two tests are independent, given the disease status and the problem specifications. We will make this assumption for this problem.
- **Note**: It may not be sensible to assume the two tests are independent for unknown disease status, as learning about the outcome of the first test could affect our beliefs on the probabilities for the outcome of the second test.

Example - Answers

- Let $Po2$ be the event that a second test is also positive.
- For some testing procedures, it may be justifiable to assume that the two tests are independent, given the disease status and the problem specifications. We will make this assumption for this problem.
- **Note**: It may not be sensible to assume the two tests are independent for unknown disease status, as learning about the outcome of the first test could affect our beliefs on the probabilities for the outcome of the second test.

Example - Answers

- After a second positive test,

$$\mathbb{P}(\text{Covid-19 given 2 positive tests}) = \mathbb{P}(C|Po, Po2)$$

Example - Answers

- After a second positive test,

$$\mathbb{P}(\text{Covid-19 given 2 positive tests}) = \mathbb{P}(C|Po, Po2)$$

$$\begin{aligned}
&= \frac{\mathbb{P}(Po, Po2|C)\mathbb{P}(C)}{\mathbb{P}(Po, Po2|C)\mathbb{P}(C) + \mathbb{P}(Po, Po2|C^c)\mathbb{P}(C^c)} \\
&= \frac{\mathbb{P}(Po|C)\mathbb{P}(Po2|C)\mathbb{P}(C)}{\mathbb{P}(Po|C)\mathbb{P}(Po2|C)\mathbb{P}(C) + \mathbb{P}(Po|C^c)\mathbb{P}(Po2|C^c)\mathbb{P}(C^c)} \\
&= \frac{0.70 \times 0.70 \times 0.001}{(0.70 \times 0.70 \times 0.001) + (0.05 \times 0.05 \times 0.999)} \\
&= 0.164
\end{aligned}$$

- One can use the posterior given Po as the current prior, and update beliefs in a sequential manner when $Po2$ is given.
- See Tutorial 1, for a different derivation of $\mathbb{P}(C|Po, Po2)$.

Example - Answers

- After a second positive test,

$$\mathbb{P}(\text{Covid-19 given 2 positive tests}) = \mathbb{P}(C|Po, Po2)$$

$$\begin{aligned}
&= \frac{\mathbb{P}(Po, Po2|C)\mathbb{P}(C)}{\mathbb{P}(Po, Po2|C)\mathbb{P}(C) + \mathbb{P}(Po, Po2|C^c)\mathbb{P}(C^c)} \\
&= \frac{\mathbb{P}(Po|C)\mathbb{P}(Po2|C)\mathbb{P}(C)}{\mathbb{P}(Po|C)\mathbb{P}(Po2|C)\mathbb{P}(C) + \mathbb{P}(Po|C^c)\mathbb{P}(Po2|C^c)\mathbb{P}(C^c)} \\
&= \frac{0.70 \times 0.70 \times 0.001}{(0.70 \times 0.70 \times 0.001) + (0.05 \times 0.05 \times 0.999)} \\
&= 0.164
\end{aligned}$$

- One can use the posterior given Po as the current prior, and update beliefs in a sequential manner when $Po2$ is given.
- See Tutorial 1, for a different derivation of $\mathbb{P}(C|Po, Po2)$.

Example - Answers

- After a second positive test,

$$\mathbb{P}(\text{Covid-19 given 2 positive tests}) = \mathbb{P}(C|Po, Po2)$$

$$\begin{aligned}
&= \frac{\mathbb{P}(Po, Po2|C)\mathbb{P}(C)}{\mathbb{P}(Po, Po2|C)\mathbb{P}(C) + \mathbb{P}(Po, Po2|C^c)\mathbb{P}(C^c)} \\
&= \frac{\mathbb{P}(Po|C)\mathbb{P}(Po2|C)\mathbb{P}(C)}{\mathbb{P}(Po|C)\mathbb{P}(Po2|C)\mathbb{P}(C) + \mathbb{P}(Po|C^c)\mathbb{P}(Po2|C^c)\mathbb{P}(C^c)} \\
&= \frac{0.70 \times 0.70 \times 0.001}{(0.70 \times 0.70 \times 0.001) + (0.05 \times 0.05 \times 0.999)} \\
&= 0.164
\end{aligned}$$

- One can use the posterior given Po as the current prior, and update beliefs in a sequential manner when $Po2$ is given.
- See Tutorial 1, for a different derivation of $\mathbb{P}(C|Po, Po2)$.

Bayesian inversion (Bayesian flip)

- We saw in (ii) how, when probabilities are 'inverted', the results may be counter intuitive.
- **Task**: Read lecture notes up to section 1.2.1 (included) and complete the example re a new HIV test on page 4 (solutions will be uploaded during the weekend).
- Finally, there are some (optional) materials on Bayesian inversion and the prosecutor's fallacy uploaded to Moodle.

Bayesian inversion (Bayesian flip)

- We saw in (ii) how, when probabilities are 'inverted', the results may be counter intuitive.
- **Task**: Read lecture notes up to section 1.2.1 (included) and complete the example re a new HIV test on page 4 (solutions will be uploaded during the weekend).
- Finally, there are some (optional) materials on Bayesian inversion and the prosecutor's fallacy uploaded to Moodle.

Bayesian inversion (Bayesian flip)

- We saw in (ii) how, when probabilities are 'inverted', the results may be counter intuitive.
- **Task**: Read lecture notes up to section 1.2.1 (included) and complete the example re a new HIV test on page 4 (solutions will be uploaded during the weekend).
- Finally, there are some (optional) materials on Bayesian inversion and the prosecutor's fallacy uploaded to Moodle.