Introductory Statistics

2024 Lectures
Part 7 - Conditional Probability

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Conditional probability

- So far, we know how to compute the probability of events based on outcomes of an experiment(s). What if the event has already occurred or we have some extra information about the event...this, of course, may affect the probability of the event

Example 28: You roll a die, I see the result and tell you that it is an even number. What is the probability that number 6 shows up?

Example 29: You are to guess on the probability of rain today. Will your answer be the same if you sitting inside a windowless office or if you are outside watching the sky?

Conditional probability

Definition 20: Let $A, B \in \mathcal{A}$, $P(B) > 0$. The conditional probability of A given event B is defined as

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

- $P(A|B)$ can be approximated by a frequency of A among the cases where B has occurred

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

- Clearly $P(AB) = P(A|B)P(B) = P(B|A)P(A)$ regardless of the probabilities of events A and B.

Conditional sample space

- since it is given that B occurs, the sample space has been reduced to the subset B
- B can be understood as a new conditional sample space

Theorem 10: If $B \in \mathcal{A}, P(B) > 0$ then $P(\cdot|B)$ as a function on $\mathcal{A}$ is a probability satisfying the axioms.

- conditional probability is thus a standard (axiomatic) probability on the conditional sample space B

Example 30: A perplexed investor must choose an investment instrument from among 15 different stocks, 10 different bonds, and 5 different mutual funds. Allowing each instrument an equal probability of being chosen, the investor randomly chooses an instrument. Given that the chosen instrument was not a bond, what is the probability that a stock was chosen?

Chain rule

- As an immediate consequence of the definition of conditional probability, we have the following chain rule.

Theorem 11: For any events $A_1, \ldots, A_n \in \mathcal{A}$ we have

$$P(A_1 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)\ldots P(A_n|A_1 \ldots A_{n-1})$$

provided $P(A_1 \ldots A_{n-1}) > 0$.

Example 31: A fruit basket contains 25 apples and oranges of which 20 are apples. If three fruits are randomly picked in a sequence, what is the probability that all three are apples?

Example 32: (Problem about keys)
A man has N keys of which only one opens the door. For some reason he tries them at random (eliminating the keys that have already been tried). What is the probability that he opens the door on the kth attempt?

Example 33: (during Control Session 3) Let events A, B, and C be such that $P(A) > 0$, $P(B) > 0$, and $P(C) > 0$. Label the following statements as true or false:

(i) The conditional probability $P(A|B)$ can never exceed the unconditional probability $P(A)$.

(ii) If $P(A|B) = P(A|C)$ then $P(B) = P(C)$.

(iii) $P(A|B) + P(A^c|B) = 1$.

(iv) $P(A|B) + P(A|B^c) = 1$.

Total probability formula

Theorem 12: Let $P(\cup_n B_n) = 1$ where $\{B_n\}$ is finite or countable collection of mutually disjoint events. If $P(B_n) > 0$ for all n then for $A \in \mathcal{A}$

$$P(A) = \sum_n P(A|B_n)P(B_n).$$

Definition 21: The collection of sets $\{B_n\}$ in the above theorem is called the partition of S.

- One can think of B_n's as events describing subpopulations and of $P(A)$ as the weighted average, with weights $P(B_n), n = 1, 2, \ldots$.

Bayes formula

- Sometimes it is not possible to calculate the conditional probability directly; other probabilities related to the probability in question are available

Theorem 13: (Bayes)
Let $P(\cup_n B_n) = 1$ where $\{B_n\}$ is finite or countable collection of mutually disjoint events. If $P(B_n) > 0$ for all n then for $A \in \mathcal{A}$, $P(A) > 0$

$$P(B_m|A) = \frac{P(A|B_m)P(B_m)}{\sum_n P(A|B_n)P(B_n)}.$$

- when the partition $\{B_n\}$ represents all possible mutually exclusive conditions that are logically possible, $P(B_m)$ and $P(B_m|A)$ are often called prior and posterior probabilities of B_m.
- Bayes formula is thus a tool for inverting conditional probabilities

Bayes formula

Example 34: Assume that 1% of the population has a certain disease (prevalence). A laboratory test is 98% effective if the person has the disease (sensitivity). However, the test also yields a “false positive” result for 0.5% of the healthy persons tested. What is the probability of correct diagnosis given that the test result is

a) positive?

b) negative?

Independence of two events

- If $P(A|B) = P(A)$ when $P(B) > 0$ then we say that A is independent of B.

Definition 22: Two events A and B with $P(A) > 0, P(B) > 0$ are said to be independent if $P(A|B) = P(A)$ or $P(B|A) = P(B)$. Otherwise, A and B are dependent.

- Alternatively, A and B are said to be independent if $P(AB) = P(A)P(B)$.

Theorem 14: If A and B are independent, so are events A and B^c, A^c and B, and A^c and B^c.

Mutual independence vs. pairwise independence

Definition 23: We say that $A_1, \ldots, A_n$ are (mutually) independent if for any set of indices $i_1, \ldots, i_k$ with $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ we have

$$P(A_{i_1} \ldots A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k}).$$

We say that $A_1, A_2, \ldots$ are independent if $A_1, \ldots, A_n$ are independent for any $n \in \mathbb{N}$.

Example 35: Set $S = \{s_1, s_2, s_3, s_4\}$ with $P(s_i) = 1/4, i = 1, \ldots, 4$.
Set $A = \{s_1, s_2\}, B = \{s_1, s_3\}, C = \{s_1, s_4\}$. Are A, B, C independent? Are they pairwise independent?

Example 36: Set $S = \{s_1, s_2, s_3, s_4, s_5\}$ with $P(s_1) = P(s_2) = P(s_3) = 8/27$, $P(s_4) = 1/27$, $P(s_5) = 2/27$.
Set $A = \{s_1, s_4\}, B = \{s_2, s_4\}, C = \{s_3, s_4\}$. Are A, B, C independent? Are they pairwise independent?

Theorem 15: If $A_1, \ldots, A_n$ are independent then also $A_1^c, \ldots, A_m^c, A_{m+1}, \ldots, A_n$ are independent for any $0 < m \leq n$.

Example 37: (during Control Session 3)
Label the statements true or false.

(i) The target is to be hit at least once. In three independent shots at the target (instead of one shot) you triple the chances of attaining the goal (assume each shot has the same positive chance of hitting the target).

(ii) If A and B are independent, then $P(A^c|B^c) = 1 - P(A)$.

(iii) If A and B are independent, then they must be disjoint.

Bayes formula and updates in evidence

- Let $P(\cup_n B_n) = 1$ where $\{B_n\}$ is finite or countable collection of mutually disjoint events. If $P(B_n) > 0$ for all n and $P(A) > 0$ then
$$P(B_m|A) = \frac{P(A|B_m)P(B_m)}{\sum_n P(A|B_n)P(B_n)}$$
- Suppose you get another evidence A'. How does the second evidence change the assessment the probabilities of the "state of the world"? Should one modify the prior probabilities given $A \cap A'$ or should one first obtain posterior probabilities given A and then use these posteriors as new priors?

Theorem 16: (Bayes formula for updated evidence)
Let $\{B_n\}$ be the partition with $P(B_n) > 0$ for all n. For $A, A' \in \mathcal{A}$ such that $P(AA') > 0$ we have

$$P(B_m|AA') = \frac{P(AA'|B_m)P(B_m)}{\sum_n P(AA'|B_n)P(B_n)} = \frac{P(A'|AB_m)P(B_m|A)}{\sum_n P(A'|AB_n)P(B_n|A)}.$$

Conditional independence

Definition 24:
Let A, A' and B be three events and let $P(B) > 0$. We say that events A and A' are conditional independent given B, if
$$P(AA'|B) = P(A|B)P(A'|B).$$

Theorem 17: (Bayes formula for conditionally independent updated evidence)
Let $\{B_n\}$ be the partition with $P(B_n) > 0$ for all n and A, A' be conditionally independent with respect to $\{B_n\}$. If $P(AA') > 0$ then
$$P(B_m|AA') = \frac{P(A|B_m)P(A'|B_m)P(B_m)}{\sum_n P(A|B_n)P(A'|B_n)P(B_n)} = \frac{P(A'|B_m)P(B_m|A)}{\sum_n P(A'|B_n)P(B_n|A)}.$$

Example 34 cont. Assume that 1% of the population has a certain disease (prevalence). A laboratory test is 98% effective if the person has the disease (sensitivity). However, the test also yields a “false positive” result for 0.5% of the healthy persons tested. What is the probability of correct diagnosis given that two independent test results are positive?