STAT 320: Principles of Probability
Unit 4: Conditional Probability & Statistical Independence

United Arab Emirates University

Department of Statistics

Outline

1. Conditional Probability
2. Partition of an Event
3. Law of Total Probability
4. Bayes' Theorem
5. The Notion of Statistical Independence

Conditional Probability

Example0

Example : Suppose that we toss 2 dice, and suppose that each of the 36 possible outcomes is equally likely to occur and hence has probability $\frac{1}{36}$. Suppose further that we observe that the first die is a 3. Then, given this information, what is the probability that the sum of the 2 dice equals 8?

Definition (Conditional Probability)

Let E, and F are two events such that $P(F) > 0$, then the conditional probability of E **given** F is defined to be,

$$P(E \mid F) := \frac{P(E \cap F)}{P(F)}.$$

Example

Question : : Suppose that a balanced die is tossed once. Find the conditional probability that 1 appears, given that an odd number was obtained.

A:= { Observe a 1.}, B:={Observe an odd number.}

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

Example

Example : A student is taking a one-hour-time-limit makeup examination. Suppose the probability that the student will finish the exam in less than x hours is $\frac{x}{2}$, for all $0 \leq x \leq 1$. Then, given that the student is still working after 0.75 hour, what is the conditional probability that the full hour is used?

Solution: Let L_x denote the event that the student finishes the exam in less than x hours, $0 \leq x \leq 1$, and let F be the event that the student uses the full hour. Because F is the event that the student is not finished in less than 1 hour,
$P(F) = P(\overline{L_1}) = 1 - P(L_1) = 0.5$.
Now, the event that the student is still working at time 0.75 is the complement of the event $L_{0.75}$, so the desired probability is obtained from

$$P(F \mid \overline{L_{0.75}}) = \frac{F \cap \overline{L_{0.75}}}{\overline{L_{0.75}}} = \frac{P(F)}{1 - P(F_{0.75})} = \frac{0.5}{0.625} = 0.8$$

Example

Example : A coin is flipped twice. Assuming that all four points in the sample space $\mathscr{S} = \{HH, HT, TH, TT\}$ are equally likely, what is the conditional probability that both flips land on heads, given that (a) the first flip lands on heads? (b) at least one flip lands on heads?

Solution:

Example

Example : A total of n balls are sequentially and randomly chosen, without replacement, from an urn containing r red and b blue balls ($n \leq r + b$). Given that k of the n balls are blue, what is the conditional probability that the first ball chosen is blue?

Solution: Show that this probability is $\frac{k}{n}$.

Definition (Multiplication Rule)

Let E and F are two events, then

$$\boxed{P(E \cap F)} := \boxed{P(E \mid F)} \times \boxed{P(F)}.$$

Example

Example : Celine is undecided as to whether to take a French course or a chemistry course. She estimates that her probability of receiving an A grade would be $\frac{1}{2}$ in a French course and $\frac{2}{3}$ in a chemistry course. If Celine decides to base her decision on the flip of a fair coin, what is the probability that she gets an A in chemistry?

Solution: .

Example

Example : Suppose that an urn contains 8 red balls and 4 white balls. We draw 2 balls from the urn without replacement. If we assume that at each draw each ball in the urn is equally likely to be chosen, what is the probability that both balls drawn are red?

Example

Example : Suppose that an urn contains 8 red balls and 4 white balls. We draw 2 balls from the urn without replacement. If we assume that at each draw each ball in the urn is equally likely to be chosen, what is the probability that both balls drawn are red?

Solution: Let R_1 and R_2 denote, respectively, the events that the first and second balls drawn are red. Now, given that the first ball selected is red, there are 7 remaining red balls and 4 white balls, so $P(R_2 \mid R_1) = \frac{7}{11}$. . As $P(R_1) = \frac{8}{12}$, the desired probability is

$$P(R_1 \mid R_2) = P(R_2 \mid R_1) \times P(R1) = \frac{7}{11}\frac{8}{12} = \frac{14}{33}.$$

$$P(E_1 \cap E_2 \cap E_3) = P(E_1 \mid E_2 \cap E_3) \times P(E_2 \mid E_3) \times P(E_3).$$

$$P(E_1 \cap E_2 \cap E_3)\, P(E_2 \cap E_1 \cap E_3) = P(E_2 \mid E_1 \cap E_3) \times P(E_1 \mid E_3) \times P(E_3).$$

$$P(E_1 \cap E_2 \cap E_3) = P(E_3 \cap E_2 \cap E_1) = P(E_3 \mid E_2 \cap E_1) \times P(E_2 \mid E_1) \times P(E_1).$$

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

$$P(A \cap B \cap C \cap D) = P(A \mid B \cap C \cap D) \times P(B \mid C \cap D) \times P(C \mid D) \times P(D).$$

$$P(A \cap B \cap C \cap D \cap E) = P(D \cap A \cap C \cap E \cap B) = P(D \mid A \cap C \cap E \cap B) \times P(A \cap C \cap E \cap B).$$

Definition (The Generalized Multiplication Rule:)

$P(E_1 \cap E_2 \cap \cdots E_k) := P(E_1) \times P(E_2 \mid E_1)P(E_3 \mid E_2 \cap E_1) \times \ldots \times P(E_k \mid E_1 \cap E_2 \cap \ldots \cap E_{k-1})$.

$$\begin{aligned} & P(E_1 \cap E_2 \cap \cdots \cap E_k) \\ := \; & P(E_1 \mid E_2 \cap \cdots \cap E_k) \times P(E_2 \mid E_3 \cap \cdots \cap E_k) \times \ldots \times \\ & \times P(E_{k-1} \mid E_k) \times P(E_k). \end{aligned}$$

$$\begin{aligned} & P(E_1, E_2, \cdots, E_k) \\ := \; & P(E_1 \mid E_2, \cdots, E_k) \times P(E_2 \mid E_3, \cdots, E_k) \times \ldots \times \\ & \times P(E_{k-1} \mid E_k) \times P(E_k). \end{aligned}$$

Example

Example : An ordinary deck of 52 playing cards is randomly divided into 4 piles of 13 cards each. Compute the probability that each pile has exactly 1 ace.

Example

Example : An ordinary deck of 52 playing cards is randomly divided into 4 piles of 13 cards each. Compute the probability that each pile has exactly 1 ace.

Solution: Define events E_i, $i = 1, 2, 3, 4$, as follows: $E_1 = \{$the ace of spades is in any one of the piles$\}$,
$E_2 = \{$the ace of spades and the ace of hearts are in different piles$\}$,
$E_3 = \{$ the aces of spades, hearts, and diamonds are all in different piles$\}$,
$E_4 = \{$all 4 aces are in different piles$\}$,
The desired probability is $P(E_1 \cap E_2 \cap E_3 \cap E_4)$, and by the multiplication rule,
$P(E_1 \cap E_2 \cap E_3 \cap E_4) = P(E_1)P(E_2 \mid E_1)P(E_3 \mid E_1 \cap E_2)P(E_4 \mid E_1 \cap E_2 \cap E_3)$
Note that $P(E_1 1) = 1$, because E_1 is equal to the sample space. $P(E_2 \mid E_1) = \frac{39}{51}$, $P(E3 \mid E_1 \cap E_2) = -\frac{26}{50}$,
$P(E_4 \mid E_1 \cap E_2 \cap E_3) = \frac{13}{49}$.

Therefore, $P(E_1 \cap E_2 \cap E_3 \cap E_4) = P(E_1)P(E_2 \mid E_1)P(E_3 \mid E_1 \cap E_2)P(E_4 \mid E_1 \cap E_2 \cap E_3 = 1 \times \frac{39}{51} \times \frac{26}{50} \times \frac{13}{49} = 0.105$

Outline

Partition of an Event Using Partion of Sample Space

Reminder from Unit1 Slides: Disjoint Sets

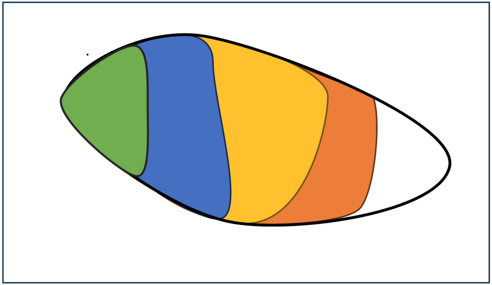
Disjoint Sets: Two sets A, and B are said to be **Disjoint Sets** or **mutually exclusive sets** if A and B does not have any elements in common.

A and B are Disjoint $\Leftrightarrow$ $A \cap B = \emptyset$.

Reminder from Unit1 Slides: Partition

Partition: A collection of sets $\{A_1, A_2, \cdots, A_k\}$ is called a **partition for a set** C if

1. $A_i \cap A_j = \emptyset$ for $1 \leq i \neq j \leq k$ (i.e. ,A_i, and A_j are Disjoint if $i \neq j$), and
2. $A_1 \cup A_2 \cup \cdots \cup A_k = C$.



Partition of Sample Space

Partition of Sample Space: A collection of sets $\{A_1, A_2, \cdots, A_n\}$ is called a **partition** for a set $\mathscr{S}$ if

1. the collection of events $\{A_i\}_{i=1}^{n}$ are **pairwise disjoint**, and
2. $A_1 \cup A_2 \cup \cdots \cup A_n = \mathscr{S}$.

Comment : In the above definition, we may replace n by ∞ and the definition extends naturally.

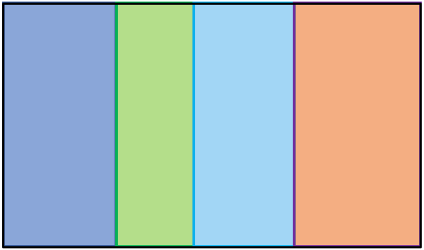
Partition of Sample Space

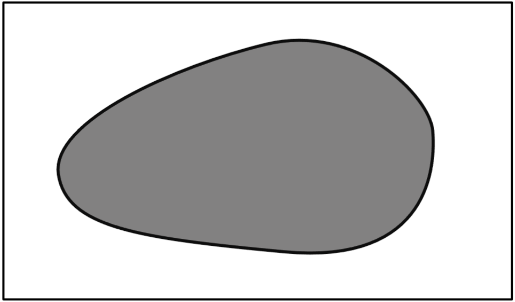
Partition of Sample Space: A collection of sets $\{A_1, A_2, \cdots, A_n\}$ is called a **partition** for a set $\mathscr{S}$ if

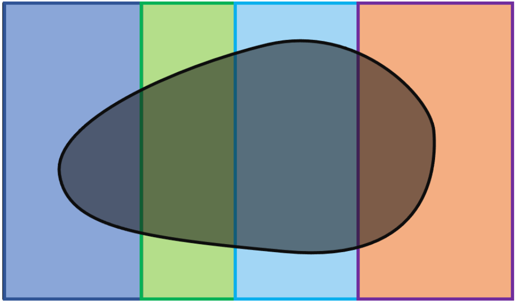
1. the collection of events $\{A_i\}_{i=1}^n$ are **pairwise disjoint**, and
2. $A_1 \cup A_2 \cup \cdots \cup A_n = \mathscr{S}$.

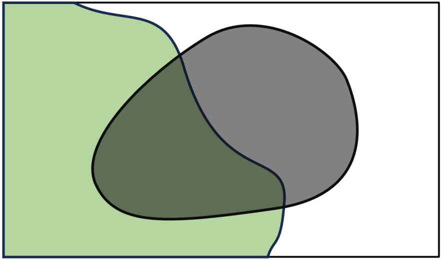
Comment : In the above definition, we may replace n by ∞ and the definition extends naturally.

Comment : Any set A and it's complement, $\bar{A}$, creates a partition of $\mathscr{S}$.





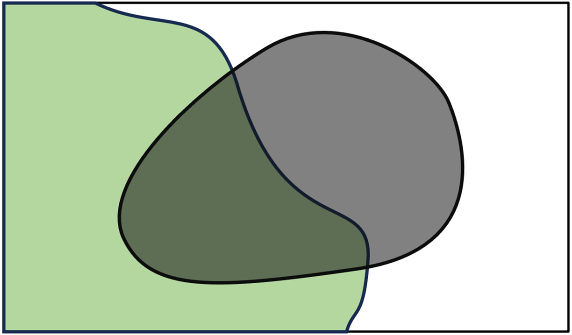




Outline

Law of Total Probability

Law of Total Probability



$$E = (E \cap F) \cup (E \cap \overline{F})$$

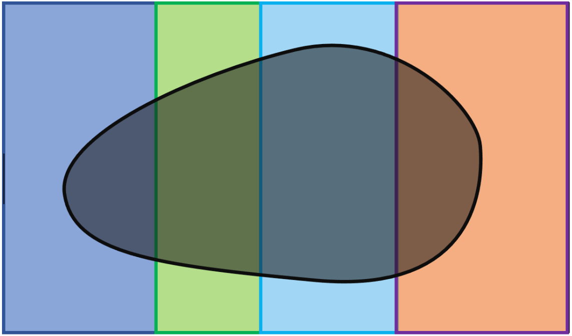
$$P(E) = P(E \cap F) + P(E \cap \overline{F})$$

Law of Total Probability

Let E and F be two events, then

$$P(E) = P(E \mid F)P(F) + P(E \mid \overline{F}))(\overline{F})$$

Law of Total Probability (General)



Law of Total Probability (General)

Law of Total Probability (General): Let E be an event. Assuming that the collection of sets $\{F_1, F_2, \ldots, F_k\}$ forms a partition of $\mathscr{S}$, we have

$$P(E) = \boxed{\sum_{j=1}^{k} P(E \mid F_j)P(F_j)}$$

Example

Example : An insurance company believes that people can be divided into two classes: those who are accident prone and those who are not. The companys statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability 0.4, whereas this probability is 0.2 for a person who is not accident prone. If we assume that 30% of the population is accident prone, what is the probability that a new policyholder will have an accident within a year of purchasing a policy?

Example

Example : An insurance company believes that people can be divided into two classes: those who are accident prone and those who are not. The companys statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability 0.4, whereas this probability is 0.2 for a person who is not accident prone. If we assume that 30% of the population is accident prone, what is the probability that a new policyholder will have an accident within a year of purchasing a policy?

Solution: Let A_1 denote the event that the policyholder will have an accident within a year of purchasing the policy, and let A denote the event that the policyholder is accident prone. Hence, the desired probability is given by

$P(A_1) = P(A_1 \mid A)P(A) + P(A1 \mid \overline{A})P(\overline{A}) = (0.4)(0.3) + (0.2)(0.7) = 0.26$

Outline

Bayes' Theorem

Theorem

Let $F_1, F_2, \ldots, F_K$ *be a set of mutually exclusive and exhaustive events (meaning that exactly one of these events must occur). Suppose now that* E *has occurred and we are interested in determining which one of the* F_j *also occurred. Then, we have the following theorem*

$$\boxed{P(F_i \mid E) = \frac{P(E \mid F_i)P(F_i)}{\sum_{j=1}^{K} P(E \mid F_j)P(F_j)}}$$

- This theorem is well known as Bayes's theorem, after the English philosopher **Thomas Bayes**.

- We can use the following Applet to find the conditional probability using Bayes's theorem:

- http://www.thomsonedu.com/statistics/book_content/0495110817_wackerly/applets/seeingstats/Chpt2/bayesTree.html

Example

Example : An insurance company believes that people can be divided into two classes: those who are accident prone and those who are not. The companys statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability 0.4, whereas this probability is 0.2 for a person who is not accident prone. If we assume that 30% of the population is accident prone.
Suppose that a new policyholder has an accident within a year of purchasing a policy. What is the probability that he or she is accident prone?

Example

Example : An insurance company believes that people can be divided into two classes: those who are accident prone and those who are not. The companys statistics show that an accident-prone person will have an accident at some time within a fixed 1-year period with probability 0.4, whereas this probability is 0.2 for a person who is not accident prone. If we assume that 30% of the population is accident prone.
Suppose that a new policyholder has an accident within a year of purchasing a policy. What is the probability that he or she is accident prone?

Solution: The desired probability is

$$P(A \mid A_1) = \frac{P(A \cap A_1)}{P(A_1)} = \frac{(0.4)(0.3)}{0.26} = \frac{6}{13}$$

Bayes' Theorem: Example

Question : : A diagnostic test for a disease is such that it (correctly) detects the disease in 90% of the individuals who actually have the disease. Also, if a person does not have the disease, the test will report that he or she does not have it with probability 0.9. Only 1% of the population has the disease in question. If a person is chosen at random from the population and the diagnostic test indicates that she has the disease, what is the conditional probability that she does, in fact, have the disease? Are you surprised by the answer? Would you call this diagnostic test reliable?

Bayes' Theorem: Example

Question : : A diagnostic test for a disease is such that it (correctly) detects the disease in 90% of the individuals who actually have the disease. Also, if a person does not have the disease, the test will report that he or she does not have it with probability 0.9. Only 1% of the population has the disease in question. If a person is chosen at random from the population and the diagnostic test indicates that she has the disease, what is the conditional probability that she does, in fact, have the disease? Are you surprised by the answer? Would you call this diagnostic test reliable?

$A = \{$The individial has the disease$\}$, $\overline{A} = \{$ does not have the disease$\}$, $B = \{$The test shows a POSITIVE result$\}$
$\overline{B} = \{$The test shows a NEGATIVE result$\}$

$P(B \mid A) = 0.9$, $P(\overline{B} \mid A^c) = 0.9$ and $P(A) = 0.01$

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid \overline{A})P(\overline{A})} = \frac{(0.9)(0.01)}{(0.9)(0.01) + (0.1)(0.99)} = 91\%$$

Example

Example : Suppose that we have 3 cards that are identical in form, except that both sides of the first card are colored red, both sides of the second card are colored black, and one side of the third card is colored red and the other side black. The 3 cards are mixed up in a hat, and 1 card is randomly selected and put down on the ground. If the upper side of the chosen card is colored red, what is the probability that the other side is colored black?

Example

Example : Suppose that we have 3 cards that are identical in form, except that both sides of the first card are colored red, both sides of the second card are colored black, and one side of the third card is colored red and the other side black. The 3 cards are mixed up in a hat, and 1 card is randomly selected and put down on the ground. If the upper side of the chosen card is colored red, what is the probability that the other side is colored black?

Solution: Let RR, BB, and RB denote, respectively, the events that the chosen card is all red, all black, or the redblack card. Also, let R be the event that the upturned side of the chosen card is red. Then the desired probability is obtained

$$P(RB \mid R) = \frac{P(R \mid RB)P(RB)}{P(R \mid RR)P(RR) + P(R \mid RB)P(RB) + P(R \mid BB)P(BB)} = \frac{\frac{1}{2} \times \frac{1}{3}}{(1) \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3} + (0) \times \frac{1}{3}} = \frac{1}{3}.$$

Example

Example : A new couple, known to have two children, has just moved into town. Suppose that the mother is encountered walking with one of her children. If this child is a girl, what is the probability that both children are girls?

Example

Example : A new couple, known to have two children, has just moved into town. Suppose that the mother is encountered walking with one of her children. If this child is a girl, what is the probability that both children are girls?

Solution:

Example

Example : A bin contains 3 different types of disposable ash-lights. The probability that a type 1 ashlight will give over 100 hours of use is 0.7, with the corresponding probabilities for type 2 and type 3 ashlights being 0.4 and 0.3, respectively. Suppose that 20 percent of the ashlights in the bin are type 1, 30 percent are type 2, and 50 percent are type 3.

1. What is the probability that a randomly chosen ashlight will give more than 100 hours of use?
2. Given that a ashlight lasted over 100 hours, what is the conditional probability that it was a type j ashlight, j = 1, 2, 3?

Solution:

Outline

The Notion of Statistical Independence

Statistical Independence

Definition (Statistically Independent Event)

Two events E and F are said to be statistically independent if

$$\boxed{P(E \cap F) = P(E) \times P(F)}$$

Statistical Independence

Definition (Statistically Independent Event)

Two events E and F are said to be statistically independent if

$$\boxed{P(E \cap F) = P(E) \times P(F)}$$

Corollary: Two events E and F are independent if and only if $P(E \mid F) = P(E)$ and $P(F \mid E) = P(F)$.

Example

Example : A card is selected at random from an ordinary deck of 52 playing cards. If E is the event that the selected card is an ace and F is the event that it is a spade, then E and F are independent.

Example

Example : A card is selected at random from an ordinary deck of 52 playing cards. If E is the event that the selected card is an ace and F is the event that it is a spade, then E and F are independent.

Solution: This follows because $P(E \cap F) = \frac{1}{52}$ **whereas** $P(E) = \frac{4}{52}$ **and** $P(F) = \frac{13}{52}$.

Example

Example : Two coins are flipped, and all 4 outcomes are assumed to be equally likely. If E is the event that the first coin lands on heads and F the event that the second lands on tails, then E and F are independent, since $P(E \cap F) = P(\{HT\}) = \frac{1}{4}$. Where as $P(E) = P(\{HH, HT\}) = \frac{1}{2}$ and $P(F) = P(\{HT, TT\}) = \frac{1}{2}$

Example

Example : Suppose that we toss 2 fair dice. Let E_1 denote the event that the sum of the dice is 6 and F denote the event that the first die equals 4. Is E_1 statistically independent of F?

Answer: $P(E_1 \cap F) = P(\{(4,2)\}) = \frac{1}{36}$ where as $P(E_1) = \frac{5}{36}$ and $P(F) = \frac{1}{6}$. Therefore, E_1 and F are not statistically independet because, $P(E_1 \cap F) \neq P(E_1) \times P(F)$.

Example : Now, suppose that we let E_2 be the event that the sum of the dice equals 7. Is E_2 statistically independent of F?

Answer: $P(E_2 \cap F) = P(\{(4,3)\}) = \frac{1}{36}$ where as $P(E_2) = \frac{1}{6}$ and $P(F) = \frac{1}{6}$. Therefore, E_2 and F are statistically independet because, $P(E_2 \cap F) = P(E_2) \times P(F)$.

Proposition: If E and F are independent, then so are E and $\overline{F}$.

Solution:

Example

Example : A card is selected at random from an ordinary deck of 52 playing cards. If E is the event that the selected card is an ace and F is the event that it is a spade, then E and F are independent.

Example

Example : A card is selected at random from an ordinary deck of 52 playing cards. If E is the event that the selected card is an ace and F is the event that it is a spade, then E and F are independent.

Solution: This follows because $P(E \cap F) = \frac{1}{52}$ **whereas** $P(E) = \frac{4}{52}$ **and** $P(F) = \frac{13}{52}$.

Generalized Definition of Statistical Independence

Definition (Statistically Independent Events)

Three events E, F, and G are said to be statistically independent if

$$P(E \cap F \cap G) = P(E) \times P(F) \times P(G)$$

$$P(E \cap F) = P(E) \times P(F)$$

$$P(E \cap G) = P(E) \times P(G)$$

$$P(F \cap G) = P(F) \times P(G)$$

Questions?