CS 441 Discrete Mathematics for CS
Lecture 20

Probabilities

Milos Hauskrecht
milos@cs.pitt.edu
5329 Sennott Square

Probabilities

Three axioms of the probability theory:

(1) Probability of a discrete outcome is:

- 0 < = P(s) <= 1

(2) Sum of probabilities of all (disjoint) outcomes is = 1

(3) For any two events E1 and E2 holds:

$P(E1 \cup E2) = P(E1) + P(E2) - P(E1 \cap E2)$

Probability distribution

Definition: A function **p: S → [0,1]** satisfying the three conditions is called **a probability distribution**

Example: a biased coin

- Probability of head 0.6, probability of a tail 0.4
- **Probability distribution:**
  - Head → 0.6 The sum of the probabilities sums to 1
  - Tail → 0.4

Note: a uniform distribution is a special distribution that assigns equal probabilities to each outcome.

M. Hauskrecht

Probability of an Event

Definition: The probability of the event *E* is the sum of the probabilities of the outcomes in *E*.

$$p(E) = \sum_{s \in S} p(s)$$

- Note that now no assumption is being made about the distribution.

Complement:

- P(~E) = 1 – P(E)

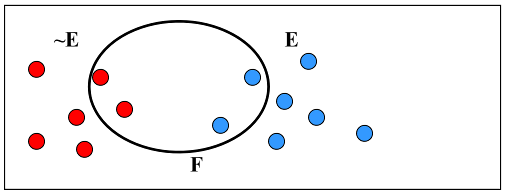
M. Hauskrecht

Complements

Let E and F are two events. Then:

- $P(F) = P(F \cap E) + P(F \cap \sim E)$

Sample space



Conditional probability

Definition: Let E and F be two events such that P(F) > 0. The **conditional probability** of E given F

- $P(E| F) = P(E \cap F) / P(F)$

Corrolary: Let E and F be two events such that P(F) > 0. Then:

- $P(E \cap F) = P(E| F) * P(F)$

This result is also referred to as a product rule.

Conditional probability

Product rule:

- $P(E \cap F) = P(E|F)\, P(F)$

Example:

- Assume the probability of getting a flu is 0.2
- Assume the probability of having a high fever given the flu: 0.9

What is the probability of getting a flu with fever?

$P(flu \wedge fever) = P(fever \mid flu)P(flu) = 0.9*0.2 = 0.18$

- When is this useful?

Sometimes conditional probabilities are easier to estimate.

M. Hauskrecht

Bayes theorem

Definition: Let E and F be two events such that $P(F) > 0$. Then:

- $P(E| F) = P(F| E)P(E) / P(F)$

  $= P(F|E)P(E) / (P(F|E)P(E) + P(F|\sim E)P(\sim E))$

Proof:

$P(E| F) = P(E \cap F) / P(F)$

$= P(F| E) P(E) / P(F)$

$P(F) = P(F \cap E) + P(F \cap \sim E)$

$= P(F| E) P(E) + P(F| \sim E) P(\sim E)$

Hence:

$P(E| F) = P(F|E)P(E) / (P(F|E)P(E) + P(F|\sim E)P(\sim E))$

Idea: Simply switch the conditioning events.

M. Hauskrecht

Bayes theorem

Definition: Let E and F be two events such that P(F) > 0. Then:

- $P(E|F) = P(F|E)P(E) / P(F)$

  $= P(F|E)P(E) / (P(F|E)P(E) + P(F|\sim E)P(\sim E))$

Example:

- Assume the probability of getting a flu is 0.2
- Assume the probability of getting a fever is 0.3
- Assume the probability of having a high fever given the flu: 0.9
- What is the probability of having a flu given the fever?

M. Hauskrecht

Bayes theorem

Definition: Let E and F be two events such that P(F) > 0. Then:

- $P(E|F) = P(F|E)P(E) / P(F)$

  $= P(F|E)P(E) / (P(F|E)P(E) + P(F|\sim E)P(\sim E))$

Example:

- Assume the probability of getting a flu is 0.2
- Assume the probability of getting a fever is 0.3

- Assume the probability of having a fever given the flu: 0.9
- What is the probability of having a flu given the fever?
- P(flu | fever) = P(fever | flu) P(flu) / P(fever) =

  = 0.9 x 0.2/0.3 = 0.18/0.3 = 0.6

M. Hauskrecht

Bayes theorem

Definition: Let E and F be two events such that P(F) > 0. Then:

- P(E| F) = P(F| E)P(E) / P(F)

  = P(F|E)P(E) / (P(F|E)P(E) + P(F|~E)P(~E))

Example (same as above but different probabilities are given):

- Assume the probability of getting a flu is 0.2
- Assume the probability of having a fever given the flu: 0.9
- Assume the probability of having a fever given no flue: 0.15
- What is the probability of having a flu given the fever?

Bayes theorem

Definition: Let E and F be two events such that P(F) > 0. Then:

- P(E| F) = P(F| E)P(E) / P(F)

  = P(F|E)P(E) / (P(F|E)P(E) + P(F|~E)P(~E))

Example:

- Assume the probability of getting a flu is 0.2
- Assume the probability of having a fever given the flu: 0.9
- Assume the probability of having a fever given no flue: 0.15
- What is the probability of having a flu given the fever?
- P(flu | fever) = P(fever | flu) P(flu) / P(fever)
- P(fever) = P(fever | flu) P(flu) + P(fever | ~flu) P(~flu)

  = 0.9*0.2 + 0.15*0.8 = 0.3

P(flu | fever) **= 0.9 x 0.2/0.3 = 0.18/0.3 = 0.6**

Independence

Definition: The two events E and F are said to be **independent** if:

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

Independence

Definition: The events E and F are said to be **independent** if:

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

Example. Assume that E denotes the family has three children of both sexes and F the fact that the family has at most one boy. Are E and F independent?

- All combos = {BBB, BBG, BGB, GBB,BGG,GBG,GGB,GGG} the number of elements = 8
- Both sexes = {BBG BGB GBB BGG GBG GGB} # = 6
- At most one boy = {GGG GGB GBG BGG} # = 4
- $E \cap F$ = {GGB GBG BGG} # = 3
- $P(E \cap F) = 3/8$ and $P(E)*P(F)= 4/8\ 6/8 = 3/8$
- **The two probabilities are equal → E and F are independent**

Independence

Example:

- Assume the probability of getting a flu is 0.2
- Assume the probability of getting a fever is 0.3
- Assume the probability of having a fever given the flu: 0.9
- Are flu and fever independent ?
- P(**flu ∩ fever) = P(fever | flu) * P(flu) = 0.2 * 0.9 = 0.18**
- **P(flu) * P(fever) = 0.2 * 0.3 = 0.06**
- **Independent or not? Not independent**

M. Hauskrecht

Bernoulli trial

Assume:

- **p = 0.6 is a probability of seeing head**
- **0.4 is the probability of seeing tail**

Assume we see a sequence of independent coin flips:

- **HHHTTHTHHT**
- **The probability of seeing this sequence:**
  $0.6^{6} * 0.4^{4}$

- **What is the probability of seeing a sequence of with 6 Heads and 4 tails?**
- The probability of each such sequence is $0.6^{6} * 0.4^{4}$
- **How many such sequences are there:** C(10,4)
- P(6H and 4T) = C(10,4) $*0.6^{6} * 0.4^{4}$

CS 1571 Intro to AI

M. Hauskrecht

Random variables

- **Definition: A random variable** is a function from the **sample space of an experiment** to the set of real numbers f: S → R. A random variable assigns a number to each possible outcome.

- **The distribution of a random variable X on the sample space S** is a set of pairs (r p(X=r)) for all r in S where r is the number and p(X=r) is the probability that X takes a value r.

Random variables

Example:

Let S be the outcomes of a two-dice roll

Let random variable X denotes the sum of outcomes

(1,1) → 2

(1,2) and (2,1) → 3

(1,3), (3,1) and (2,2) → 4

…

Distribution of X:

- 2 → 1/36,
- 3 → 2/36,
- 4 → 3/36 …
- 12 → 1/36

Probabilities

- **Assume a repeated coin flip**
- P(head) =0.6 and the probability of a tail is 0.4. Each coin flip is independent of the previous.
- What is the probability of seeing:
  - HHHHH - 5 heads in a row
- **P(HHHHH) = 0.6^5 =**
  - Assume the outcome is HHTTT
- **P(HHTTT)= $0.6*0.6*0.4^{3}$= $0.6^2*0.4^3$**
  - Assume the outcome is TTHHT
- **P(TTHHT)=$0.4^{2*}0.6^2*0.4$ = $0.6^2*0.4^3$**
- What is the probability of seeing three tails and two heads?
- The number of two-head-three tail combinations = **C(5,2)**
- P(two-heads-three tails) = **C(5,2) *0.6^2 *0.4^3**

M. Hauskrecht

Probabilities

- **Assume a variant of a repeated coin flip problem**
- The space of possible outcomes is the count of occurrences of heads in 5 coin flips. For example:
- TTTTT yields outcome 0
- HTTTT or TTHTT yields 1
- HTHHT yields 3 …

- What is the probability of an outcome 0?
- P(outcome=0) = **0.6^0 *0.4^5**
- P(outcome=1) = **C(5,1) 0.6^1 *0.4^4**
- P(outcome =2) = **C(5,2) 0.6^2 *0.4^3**
- P(outcome =3) = **C(5,3) 0.6^3 *0.4^2**
- …

M. Hauskrecht

Expected value and variance

Definition: **The expected value** of the random variable X(s) on the sample space is equal to:

$$E(X) = \sum_{s \in S} p(s) X(s)$$

Example: roll of a dice

- Outcomes: 1 2 3 4 5 6
- Expected value:

E(X) = ?

M. Hauskrecht

Expected value and variance

Definition: **The expected value** of the random variable X(s) on the sample space is equal to:

$$E(X) = \sum_{s \in S} p(s) X(s)$$

Example: roll of a dice

- Outcomes: 1 2 3 4 5 6
- Expected value:

E(X) = 1*1/6 + 2*1/6+3*1/6 + 4*1/6 + 5*1/6 +6* 1/6 = 7/2

M. Hauskrecht

Expected value

Example:

Flip a fair coin 3 times. The outcome of the trial X is the number of heads. What is the expected value of the trial?

Answer:

Possible outcomes:

= {HHH HHT HTH THH HTT THT TTH TTT}

3 2 2 2 1 1 1 0

E(X) = ?

M. Hauskrecht

Expected value

Example:

Flip a fair coin 3 times. The outcome of the trial X is the number of heads. What is the expected value of the trial?

Answer:

Possible outcomes:

= {HHH HHT HTH THH HTT THT TTH TTT}

3 2 2 2 1 1 1 0

E(X) = 1/8 (3 + 3*2 +3*1 +0) = 12/8= 3/2

M. Hauskrecht

Expected value

- **Theorem:** If Xi i=1,2,3, n with n being a positive integer, are random variables on S, and a and b are real numbers then:
  - E(X1+X2+ ... Xn) = E(X1)+E(X2) + ... E(Xn)
  - E(aX+b) = aE(X) +b

M. Hauskrecht

Expected value

Example:

- Roll a pair of dices. What is the expected value of the sum of outcomes?
- **Approach 1:**
- Outcomes: (1,1) (1,2) (1,3) .... (6,1)... (6,6)

  2 3 4 7 12

Expected value: 1/36 (2*1 + ...) = 7

- Approach 2 (theorem):
- E(X1+X2) = E(X1) +E(X2)
- E(X1) =7/2 E(X2) = 7/2
- E(X1+X2) = 7

M. Hauskrecht