DM551

Algorithms and Probability

September 17, 2018

(Poll) What do you think about the speed of the lectures?

(Poll) What do you think about the speed of the lectures?

(Poll) What do you think about what I mark on slides?

(Poll) What do you think about the speed of the lectures?

(Poll) What do you think about what I mark on slides?

(Poll) What do you think about this type of poll compared to Socrative?

(Poll) What do you think about the speed of the lectures?

(Poll) What do you think about what I mark on slides?

(Poll) What do you think about this type of poll compared to Socrative?

Do you have suggestions?

(Poll) What do you think about the speed of the lectures?

(Poll) What do you think about what I mark on slides?

(Poll) What do you think about this type of poll compared to Socrative?

Do you have suggestions?

I will be using Blackboard's Discussion Board to tell when there are corrections to the slides.

Discrete Probability

Fact. If all outcomes of a finite sample space S are equally likely, the probability of an event E is $p(E) = |E|/|S|$. This distribution of probabilities is called the *uniform distribution*.

Discrete Probability

Fact. If all outcomes of a finite sample space S are equally likely, the probability of an event E is $p(E) = |E|/|S|$. This distribution of probabilities is called the *uniform distribution*.

Def. Events $A_1, A_2, ...$ are *pairwise mutually exclusive* if $A_i \cap A_j = \emptyset$ for $i \neq j$.

Discrete Probability

Fact. If all outcomes of a finite sample space S are equally likely, the probability of an event E is $p(E) = |E|/|S|$. This distribution of probabilities is called the *uniform distribution*.

Def. Events $A_1, A_2, \ldots$ are *pairwise mutually exclusive* if $A_i \cap A_j = \emptyset$ for $i \neq j$.

Suppose sample space $S = \{x_1, x_2, \ldots, x_n\}$, and probability of x_i is $p(x_i)$. Must have

- $0 \leq p(x_i) \leq 1 \quad \forall i$
- $\sum_{i=1}^{n} p(x_i) \ = \ 1.$
- For any events $A_1, A_2, \ldots A_k$ that are pairwise mutually exclusive, $p(\bigcup_i A_i) = \sum_i p(A_i)$.

Then the function p is a *probability distribution*.

Discrete Probability

Suppose sample space $S = \{x_i \mid i \geq 1\}$ (countable). Then,

- $0 \leq p(x_i) \leq 1 \quad \forall i$
- $\sum_{i=1}^{\infty} p(x_i) \ = \ 1.$
- For any events $A_1, A_2, \ldots$ that are pairwise mutually exclusive, $p(\bigcup_i A_i) = \sum_i p(A_i)$.

Discrete Probability

Suppose sample space $S = \{x_i \mid i \geq 1\}$ (countable). Then,

- $0 \leq p(x_i) \leq 1 \quad \forall i$
- $\sum_{i=1}^{\infty} p(x_i) \;=\; 1.$
- For any events $A_1, A_2, ...$ that are pairwise mutually exclusive, $p(\bigcup_i A_i) = \sum_i p(A_i).$

Example:

experiment — a fair coin is flipped until "heads".

sample space — $S = \{H, TH, TTH, ..., T^n H, ...\}$.

event — 1 sequence of flips

probability — $p(T^n H) = (1/2)^{n+1}$ for $n \geq 0$.

Discrete Probability

Suppose sample space $S = \{x_i \mid i \geq 1\}$ (countable). Then,

- $0 \leq p(x_i) \leq 1 \quad \forall i$
- $\sum_{i=1}^{\infty} p(x_i) \;=\; 1.$
- For any events $A_1, A_2, \ldots$ that are pairwise mutually exclusive, $p(\bigcup_i A_i) = \sum_i p(A_i)$.

Example:

experiment — a fair coin is flipped until "heads".

sample space — $S = \{H, TH, TTH, \ldots, T^nH, \ldots\}$.

event — 1 sequence of flips

probability — $p(T^nH) = (1/2)^{n+1}$ for $n \geq 0$.

Is this a probability distribution?

Discrete Probability

Suppose the coin is biased — $p(\text{heads}) = p$; $p(\text{tails}) = q = 1 - p$.

$p(T^n H) = q^n p$ — the *geometric distribution*

Discrete Probability

Suppose the coin is biased — $p(\text{heads}) = p$; $p(\text{tails}) = q = 1 - p$.

$p(T^n H) = q^n p$ — the *geometric distribution*

Is $0 \leq q^n p \leq 1$? (Poll)

Discrete Probability

Suppose the coin is biased — $p(\text{heads}) = p$; $p(\text{tails}) = q = 1 - p$.

$p(T^nH) = q^n p$ — the *geometric distribution*

Is $0 \leq q^n p \leq 1$? (Poll)

Is this a probability distribution?

Discrete Probability

Suppose the coin is biased — $p(\text{heads}) = p$; $p(\text{tails}) = q = 1 - p$.

$p(T^n H) = q^n p$ — the *geometric distribution*

Is $0 \leq q^n p \leq 1$? (Poll)

Is this a probability distribution?

The probability of event E is

$$p(E) = \sum_{x_i \in E} p(x_i)$$

.

Discrete Probability

Example: For fair 6-sided dice: $p(\{5,6\}) = 1/3$.

Suppose a die is loaded so

$$\begin{array}{lll} p(1) = 1/3 & p(2) = 2/15 & p(3) = 2/15 \\ p(4) = 2/15 & p(5) = 2/15 & p(6) = 2/15 \end{array}$$

Then $\sum_{i=1}^{6} p(i) = 1/3 + 5(2/15) = 1$.

For this die, $p(\{5,6\}) = 4/15 < 1/3$.

Discrete Probability

Example: For fair 6-sided dice: $p(\{5, 6\}) = 1/3$.

Suppose a die is loaded so

$$\begin{array}{lll} p(1) = 1/3 & p(2) = 2/15 & p(3) = 2/15 \\ p(4) = 2/15 & p(5) = 2/15 & p(6) = 2/15 \end{array}$$

Then $\sum_{i=1}^{6} p(i) = 1/3 + 5(2/15) = 1$.

For this die, $p(\{5, 6\}) = 4/15 < 1/3$.

With 2 of these dice $p(\text{sum} = 7)$ is

$$2 \cdot \frac{1}{3} \cdot \frac{2}{15} + 4 \cdot \frac{2}{15} \cdot \frac{2}{15} = 4/25 < 1/6$$

Discrete Probability

Thm. $p(\overline{E}) = 1 - p(E)$.

Pf. $\sum_{i=1}^{n} p(x_i) \ = \ 1 \ = \ p(E) + p(\overline{E})$.

Thus, $P(\overline{E}) = 1 - p(E)$. $\square$

Discrete Probability

Thm. $p(\overline{E}) = 1 - p(E)$.

Pf. $\sum_{i=1}^{n} p(x_i) \; = \; 1 \; = \; p(E) + p(\overline{E})$.

Thus, $P(\overline{E}) = 1 - p(E)$. $\square$

Substituting ∞ for n changes nothing.

Discrete Probability

Thm. $p(\overline{E}) = 1 - p(E)$.

Pf. $\sum_{i=1}^{n} p(x_i) \ = \ 1 \ = \ p(E) + p(\overline{E})$.

Thus, $P(\overline{E}) = 1 - p(E)$. $\square$

Substituting ∞ for n changes nothing.

Thm. $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$.

Pf.

$$
\begin{aligned}
p(E_1 \cup E_2) &= \sum_{x_i \in E_1 \cup E_2} p(x_i) \\
&= \sum_{x_i \in E_1} p(x_i) + \sum_{x_i \in E_2} p(x_i) - \sum_{x_i \in E_1 \cap E_2} p(x_i) \\
&= p(E_1) + p(E_2) - p(E_1 \cap E_2). \qquad \square
\end{aligned}
$$

Conditional Probability

Def. The **conditional probability** of E given F is

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

Conditional Probability

Def. The **conditional probability** of E given F is

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

Example: With 2 fair dice, what is the probability that the sum is 7, given that both dice are ≥ 3?

Conditional Probability

Def. The **conditional probability** of E given F is

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

Example: With 2 fair dice, what is the probability that the sum is 7, given that both dice are ≥ 3?

Answer: $p(\text{sum is 7 and both} \geq 3)/p(\text{both} \geq 3) = 2 \cdot \frac{1}{6} \cdot \frac{1}{6}/(\frac{2}{3} \cdot \frac{2}{3})$ $= 1/8$.

Conditional Probability

Monte Hall 3-Door Puzzle: Suppose you are on a game show.

The host asks you to choose 1 of 3 doors for a prize.

You choose door A.

The host opens another door B. No big prize there.

You are told you can switch your choice.

Should you switch? (Poll)

Conditional Probability

Monte Hall 3-Door Puzzle: Suppose you are on a game show.

The host asks you to choose 1 of 3 doors for a prize.

You choose door A.

The host opens another door B. No big prize there.

You are told you can switch your choice.

Should you switch? (Poll)

Answer: Yes.

$p(\text{prize behind } A \mid \text{not behind } B) = p(A \wedge \neg B)/p(\neg B) = 1/2$ is not the answer. B was chosen after A.

Conditional Probability

Monte Hall 3-Door Puzzle: Suppose you are on a game show.

The host asks you to choose 1 of 3 doors for a prize.

You choose door A.

The host opens another door B. No big prize there.

You are told you can switch your choice.

Should you switch? (Poll)

Answer: Yes.

$p(\text{prize behind } A \mid \text{not behind } B) = p(A \wedge \neg B)/p(\neg B) = 1/2$ is not the answer. B was chosen after A.

$p(\text{prize behind } A \mid \text{the host chose } B)$ is the correct probability.

You could have chosen 3 doors.

If you chose the prize, the host has 2 choices; otherwise only 1.

Conditional Probability

Your choice	Location of Prize	Prob of B
A	A	(1/3)(1/2)
A	B	(1/3)(0)
A	C	(1/3)(1)

Conditional Probability

Your choice	Location of Prize	Prob of B
A	A	$(1/3)(1/2)$
A	B	$(1/3)(0)$
A	C	$(1/3)(1)$

So $p(\text{prize behind } A \text{ and host chose } B) = (1/3)(1/2) = 1/6$.

$p(\text{host chose } B) = (1/3)(1/2) + (1/3) = 1/2$.

Thus, $p(\text{prize behind } A \mid \text{the host chose } B) = (1/6)/(1/2) = 1/3$.

But $p(\text{prize behind } C \mid \text{the host chose } B) = 2/3$.

Conditional Probability

Another tricky example: Suppose you know that A has two children and you are told that one is a girl.

What is the probability that the other is also a girl?

Conditional Probability

Another tricky example: Suppose you know that A has two children and you are told that one is a girl.

What is the probability that the other is also a girl?

Wrong answer: 1/2 since there is always a 50-50 chance for a boy or a girl. This is wrong even assuming that the probabilities are exactly 50-50 and that the events are independent.

This is only correct if you said the oldest or the youngest, etc.

Conditional Probability

Another tricky example: Suppose you know that A has two children and you are told that one is a girl.

What is the probability that the other is also a girl?

Wrong answer: 1/2 since there is always a 50-50 chance for a boy or a girl. This is wrong even assuming that the probabilities are exactly 50-50 and that the events are independent.

This is only correct if you said the oldest or the youngest, etc.

Correct answer: 1/3.

There are 4 possibilities: (G,G), (G,B), (B,G), (B,B).

"One is a girl" only rules out the last!.

Conditional Probability

Def. E and F are **independent** iff $p(E \cap F) = p(E)p(F)$.

Fact: If E and F are independent, then
$p(E|F) = p(E \cap F)/p(F) = p(E)$.

Conditional Probability

Def. E and F are **independent** iff $p(E \cap F) = p(E)p(F)$.

Fact: If E and F are independent, then
$p(E|F) = p(E \cap F)/p(F) = p(E)$.

Example: With 2 fair dice, the probability that the sum is 7 is *not* independent of both dice being ≥ 3.

$p(\text{sum is 7 and both} \geq 3)/p(\text{both} \geq 3) = 2 \cdot \frac{1}{6} \cdot \frac{1}{6}/(\frac{2}{3} \cdot \frac{2}{3}) =$
$1/8 \neq 1/6 = p(\text{sum is 7})$.

Conditional Probability

Def. E and F are **independent** iff $p(E \cap F) = p(E)p(F)$.

Fact: If E and F are independent, then
$p(E|F) = p(E \cap F)/p(F) = p(E)$.

Example: With 2 fair dice, the probability that the sum is 7 is *not* independent of both dice being ≥ 3.

$p(\text{sum is 7 and both} \geq 3)/p(\text{both} \geq 3) = 2 \cdot \frac{1}{6} \cdot \frac{1}{6}/(\frac{2}{3} \cdot \frac{2}{3}) =$
$1/8 \neq 1/6 = p(\text{sum is 7})$.

Example: The probability that the sum of the first two dice is 7 given that both dice are ≥ 3 is independent of the probability of a third die being 1 or 2.

Bernouli Trials

Bernoulli trial: – an experiment with 2 outcomes; success with probability p, failure with probability $q = 1 - p$.

Bernoulli trials: — k independent repetitions of a Bernoulli trial.

Bernouli Trials

Bernoulli trial: – an experiment with 2 outcomes; success with probability p, failure with probability $q = 1 - p$.

Bernoulli trials: — k independent repetitions of a Bernoulli trial.

Example: Consider 2 fair dice. Throw k times. What is the probability that the first time you get a sum of 7 is on throw j, where $j \leq k$.

Bernouli Trials

Bernoulli trial: – an experiment with 2 outcomes; success with probability p, failure with probability $q = 1 - p$.

Bernoulli trials: — k independent repetitions of a Bernoulli trial.

Example: Consider 2 fair dice. Throw k times. What is the probability that the first time you get a sum of 7 is on throw j, where $j \leq k$.

Answer: $(\frac{5}{6})^{j-1}(\frac{1}{6})$.

This is the same as the biased coin with probabilities p of "heads" and $q = 1 - p$ of "tails". The answer is from the geometric distribution: $q^{j-1}p$.

Try tree diagrams.

Bernouli Trials

Thm. The probability of k successes in n independent Bernoulli trials is $\binom{n}{k} p^k q^{n-k}$.

Pf. The outcome is $(x_1, x_2, ..., x_n)$, where

$$x_i = \begin{cases} S, & \text{if the } i\text{th trial is success} \\ F, & \text{if the } i\text{th trial is failure} \end{cases}$$

$p((x_1, x_2, ..., x_n))$, with k successes and $n - k$ failures is $p^k q^{n-k}$.

There are $\binom{n}{k}$ outcomes with k successes and $n - k$ failures.

The probability is $\binom{n}{k} p^k q^{n-k}$. $\square$

Bernouli Trials

Thm. The probability of k successes in n independent Bernoulli trials is $\binom{n}{k} p^k q^{n-k}$.

Pf. The outcome is $(x_1, x_2, ..., x_n)$, where

$$x_i = \begin{cases} S, & \text{if the } i\text{th trial is success} \\ F, & \text{if the } i\text{th trial is failure} \end{cases}$$

$p((x_1, x_2, ..., x_n))$, with k successes and $n-k$ failures is $p^k q^{n-k}$.

There are $\binom{n}{k}$ outcomes with k successes and $n-k$ failures.

The probability is $\binom{n}{k} p^k q^{n-k}$. $\square$

The **binomial distribution** is $b(k; n, p) = \binom{n}{k} p^k q^{n-k}$.

Note $\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$.