Generating permutations

Suppose you want to test your sorting algorithm.

Try it on all permutations of $\{1, 2, \ldots, n\}$.

How large should n be?

Generating permutations

Suppose you want to test your sorting algorithm.

Try it on all permutations of $\{1, 2, ..., n\}$.

How large should n be?

If n is 10, this is 10!=3,628,800, so not very large n.

Generating permutations

Suppose you want to test your sorting algorithm.

Try it on all permutations of $\{1, 2, ..., n\}$.

How large should n be?

If n is 10, this is 10!=3,628,800, so not very large n.

Lexicographic order would be nice.

$(2, 4, 7, 3, 8, 1, 9, 6, 5, 10) < (2, 4, 7, 5, 6, 10, 9, 1, 3, 8)$ because $3 < 5$.

How to get the next larger permutation? (after $(a_1, a_2, ..., a_n)$)

(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1).

Generating permutations

(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1).

If $a_{n-1} < a_n$, switch them.

Otherwise, find largest j with $a_j < a_{j+1}$.

$a_{j+1} > a_{j+2} > ... > a_n$.

Find the least value $a_k > a_j$ from $\{a_{j+1}, a_{j+2}, ..., a_n\}$.

Put a_k where a_j was.

Put the remaining elements from $\{a_j, a_{j+1}, ..., a_n\}$ in increasing order.

$(1, 2, 5, 10, 7, \mathbf{4}, 9, 8, 6, 3) \rightarrow$

$(1, 2, 5, 10, 7, \mathbf{6}, 3, 4, 8, 9)$.

Generating permutations

procedure next_permutation($a_1, a_2, ..., a_n$)
{ a permutation of $(1, 2, ..., n)$, $\neq (n, n-1, ..., 1)$}
$j \leftarrow n - 1$
while $a_j > a_{j+1}$; $j \leftarrow j - 1$
{ j is largest subscript with $a_j < a_{j+1}$}

$k \leftarrow n$
while $a_j > a_k$; $k \leftarrow k - 1$
{ a_k is smallest value $> a_j$ to right of a_j}

switch a_j and a_k
$r \leftarrow n$
$s \leftarrow j + 1$
while $r > s$
    switch a_r and a_s
    $r \leftarrow r - 1$; $s \leftarrow s + 1$
{ this reverses the order after a_j}

Generating combinations

Combinations are subsets:

So use binary strings to represent them.

Lexicographic order of strings is increasing order of integers.

$(000), (001), (010), (011), (100), (101), (110), (111).$

To get next integer, find rightmost 0.

Change it to 1. Change all 1's to right to 0's.

Generating combinations

Combinations are subsets:

So use binary strings to represent them.

Lexicographic order of strings is increasing order of integers.

$(000), (001), (010), (011), (100), (101), (110), (111)$.

To get next integer, find rightmost 0.

Change it to 1. Change all 1's to right to 0's.

For an r-combination, have r ones in the string.

$(00111), (01011), (01101), (01110), (10011), (10101), (10110),$
$(11001), (11010), (11100)$.

Find the rightmost 01.

Change it to 10.

Move all the 1's to the right as far right as possible.

Generating combinations

To find the next r-combination of $(1, 2, ..., n)$:

$(11100) \leftrightarrow (1, 2, 3)$ — the smallest.

Lexicographic order is not lexicographic order of strings.

To get the next, find the rightmost 10.

Change it to 01.

Move all the 1's to the right as far left as possible.

Generating combinations

To do this directly:

Suppose have $(a_1, a_2, \ldots, a_r)$.

Generating combinations

To do this directly:

Suppose have $(a_1, a_2, ..., a_r)$.

If this looks like $(a_1, ..., a_k, n - r + k + 1, n - r + k + 2, ..., n - 1, n)$,

should change a_k to $a_k + 1$.

(in what follows, use the old value of a_k)

$a_{k+1} \leftarrow a_k + 2$.

$a_{k+2} \leftarrow a_k + 3$.

Generally, $a_j = a_k + j - k + 1$, for $k + 1 \leq j \leq r$.

Generating combinations

To do this directly:

Suppose have $(a_1, a_2, \ldots, a_r)$.

If this looks like $(a_1, \ldots, a_k, n-r+k+1, n-r+k+2, \ldots, n-1, n)$,

should change a_k to $a_k + 1$.

(in what follows, use the old value of a_k)

$a_{k+1} \leftarrow a_k + 2$.

$a_{k+2} \leftarrow a_k + 3$.

Generally, $a_j = a_k + j - k + 1$, for $k + 1 \leq j \leq r$.

Suppose $n = 6$. Consider $(2, 5, 6) \to (3, 4, 5)$.

Generating combinations

procedure next_combination($a_1, a_2, ..., a_r$)
{ a r-combination of $(1, 2, ..., n)$, $\neq (n - r + 1, ..., n)$}
$i \leftarrow r$
while $a_i = n - r + i$
    $i \leftarrow i - 1$
$a_i \leftarrow a_i + 1$
for $j \leftarrow i + 1$ **to** r
    $a_j \leftarrow a_i + j - i$

Discrete Probability

Example: 6-sided dice — 2

experiment — throwing the dice, getting 2 numbers

sample space — all pairs of numbers (i, j), where $1 \leq i \leq j \leq 6$.

event — pairs of numbers that sum to 7

probability that the sum is 7 — $\frac{6}{6 \cdot 6} = \frac{1}{6} = \frac{|E|}{|S|}$,

where E is an event and S is a finite sample space.

Discrete Probability

Example: 6-sided dice — 2

experiment — throwing the dice, getting 2 numbers

sample space — all pairs of numbers (i, j), where $1 \leq i \leq j \leq 6$.

event — pairs of numbers that sum to 7

probability that the sum is 7 — $\frac{6}{6 \cdot 6} = \frac{1}{6} = \frac{|E|}{|S|}$,

where E is an event and S is a finite sample space.

Example: Deck of 52 playing cards:

experiment — taking 13 different cards at random

sample space — all hands containing 13 cards

event — all hands with no face cards or aces

probability — $\frac{\frac{36!}{13!23!}}{\frac{52!}{13!39!}} = \frac{36!39!}{23!52!} = .00363896\ldots$

Discrete Probability

Example: 36 numbered balls

experiment — randomly choosing 7 balls and then 4 more

sample space — all pairs of sets of numbers (A, B), where all numbers are distinct, $|A| = 7$, $|B| = 4$.

event — a specific set of pairs $(A_1, B_1), \ldots, (A_k, B_k)$, such that each A_i contains 6 of the 7 distinct numbers $x_1, x_2, x_3, x_4, x_5, x_6, x_7$ and each B_i contains the number missing from A_i.

probability is $\frac{k}{\frac{36!}{7!29!} \cdot \frac{29!}{4! \cdot 25!}} = \frac{k}{\frac{36!}{4!7!25!}}$.

Discrete Probability

Example: 36 numbered balls

experiment — randomly choosing 7 balls and then 4 more

sample space — all pairs of sets of numbers (A, B), where all numbers are distinct, $|A| = 7$, $|B| = 4$.

event — a specific set of pairs $(A_1, B_1), ..., (A_k, B_k)$, such that each A_i contains 6 of the 7 distinct numbers $x_1, x_2, x_3, x_4, x_5, x_6, x_7$ and each B_i contains the number missing from A_i.

probability is $\frac{k}{\frac{36!}{7!29!} \cdot \frac{29!}{4! \cdot 25!}} = \frac{k}{\frac{36!}{4!7!25!}}$.

So what is k? There are 7 possibilities for which x_j is not in A_i, 29 possibilities for that last number in A_i, and $28!/(3!25!)$ possibilities for the extra elements in B_i.

So $k = 7 \cdot 29 \cdot 28!/(3!25!)$,

and the entire probability is $\frac{7 \cdot 29!/(3!25!)}{\frac{36!}{4!7!25!}} = \frac{4 \cdot 7 \cdot 29! 7!}{36!} < 3.3542 * 10^{-6}$.

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 $i \neq j$.

Suppose sample space $S = \{x_1, x_2, ..., x_n\}$, 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 $A_1, A_2, ...A_k$ 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 $S = \{x_i \mid i \geq 1\}$. 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 $S = \{x_i \mid i \geq 1\}$. 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 $n \geq 0$.

Discrete Probability

Suppose sample space $S = \{x_i \mid i \geq 1\}$. 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 $n \geq 0$.

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

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

Discrete Probability

The probability of event E is

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

.

Discrete Probability

The probability of event E is

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

.

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

The probability of event E is

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

.

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.

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

Pf.

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

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

Example: 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?

Conditional Probability

Example: 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?

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

Example: 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?

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(prize behind A and host chose B) $= (1/3)(1/2) = 1/6$.

p(host chose B) $= (1/3)(1/2)+(1/3) = 1/2$.

Thus, p(prize behind A | the host chose B) $= (1/6)/(1/2) = 1/3$.

So p(prize behind A | 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 is 7 given that both dice are ≥ 3 is independent of the probabilities of the dice being 1 or 2. (Assuming that all probabilities are nonzero.)

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

Random Variables

Def. For a sample space S, *random variable* is a function $f : S \to \mathbb{R}$.

Example

Suppose a coin is flipped until the result is "heads". Let X be the random variable that equals the number of coins flipped.

Random Variables

Def. For a sample space S, *random variable* is a function $f : S \to \mathbb{R}$.

Example

Suppose a coin is flipped until the result is "heads". Let X be the random variable that equals the number of coins flipped.

$$X(H) = 1$$

$$X(TTH) = 3$$

$$X(TTTH) = 4$$

Random Variables

Def. For a sample space S, *random variable* is a function $f : S \to \mathbb{R}$.

Example

Suppose a coin is flipped until the result is "heads". Let X be the random variable that equals the number of coins flipped.

$X(H) = 1$

$X(TTH) = 3$

$X(TTTH) = 4$

Let $p(X = r)$ denote the probability the X takes the value r.

Def. The *distribution* of the random variable X on a sample space S is the set

$$\{(r, p(X = r)) \mid r \in X(S)\},$$

Random Variables

Def. For a sample space S, *random variable* is a function $f : S \to \mathbb{R}$.

Example

Suppose a coin is flipped until the result is "heads". Let X be the random variable that equals the number of coins flipped.

$X(H) = 1$

$X(TTH) = 3$

$X(TTTH) = 4$

Let $p(X = r)$ denote the probability the X takes the value r.

Def. The *distribution* of the random variable X on a sample space S is the set

$$\{(r, p(X = r)) \mid r \in X(S)\},$$

For a fair coin, the distribution of X is $\{(i, 1/2^i) \mid i \geq 1\}$.