8 Intro to Probability, Kolmogorov Complexity, and other Whale-y goodness

Computer Science Mentors 70

Mar 29 - Apr 2
Spring 2021

Amogh Gupta, Sylvia Jin, Aekus Bhathal, Abinav Routhu, Debayan Bandyopadhyay, and Rahul Shah

Roast us here: `https://tinyurl.com/csm70-feedback20`

Introduction to Kolmogorov Complexity

1. Recall from HW 7 that the *Kolmogorov complexity* of a string $K(x)$ is the length of an optimally-compressed copy of x; that is, $K(x)$ is the length of the shortest program that returns x. Let's explore this definition a little bit.

   (a) True or false? The Kolmogorov complexity of any string is always smaller than the length of the string itself.

   (b) True or false? When you use a program to "zip" files on your computer into a compressed archive, the resulting .zip file has a smaller Kolmogorov complexity than the original files.

   (c) In plain English, write the shortest possible representation of the following strings. There are multiple possible solutions because of the constraints of the English language.

      (a) UPEUPEUPEUPEUPE

      (b) `3.14159265358979323846264...`

      (c) hellooooooooo / it's me / I was wondering if after all these years you'd like to meet
      To go over everything! / they say time's supposed to heal yuh / but I ain't done much healing
      hello from the other (oh my god I'm crying) side / I must've called whale-thousand times ...

Introduction to a whale-ful of Probability

1. Suppose two integers a and b are drawn uniformly from $[-n \,..\, n]$, that is $a, b \in \mathbb{Z}$ and $-n \leq a, b \leq n$.

   (a) Define a probability space for (a, b). Does each sample point occur with uniform probability?

(b) Find the probability that $\max\{0, a\} = \min\{0, b\}$.

(c) Find the probability that $|a - b| \leq k$. You may assume $k < \frac{n}{2}$.

(d) Suppose we choose two closed intervals $u = [a \,..\, b], v = [c \,..\, d]$ uniformly at random from $[-n \,..\, n]$. What is the probability that u is enveloped by v, meaning that $u \subset v$ and $c < a \leq b < d$. What happens to this probability as n approaches ∞?

2. Prakash and Michelle are playing Yahtzee, a game involving rolling 5 dice.

(a) First, define a probability space representing the possible outcome of Prakash or Michelle's rolls of the 5 dice. Assume all dice are fair and labeled 1 through 6.

Prakash and Michelle each roll 1 die to see who goes first. The person with the higher roll goes first, and in case of a tie, they both roll their die again.

(b) What's the chance Michelle rolls a higher number on the first roll?

(c) What's the chance Michelle goes first?

(d) They finally begin playing. Partway through the game, Prakash is missing the "three of a kind" category while Michelle is missing the "four of a kind" category. What is the probability of rolling...

1. exactly 3 of a kind?

2. exactly 4 of a kind?

3. Which one is more likely? 3 of a kind or 4 of a kind?

Inclusion-Exclusion Principle, Bayes' Theorem

1. Berkeley is experiencing bad weather because of HKN's "Gloomy - inator". It is always at least rainy, cloudy or windy, but because the inator is random we don't exactly know what it would be like. It rains with 0.5 probability, gets windy with 0.65 probability and gets cloudy with 0.45 probability. We experience at least 2 of these together with probability 0.45. Help UPE find the probability that all 3 of these happen together.

2. UPE owns a pizzeria. They observe that one of your customers, Andy, buys a cheese pizza on Saturday with probability 0.3 and on Sunday with probability 0.6.

   (a) If Andy's pizza purchasing habits on Sunday is independent from his pizza purchasing habits on Saturday, what is the probability that he buys pizza on a given weekend?

   (b) If Andy buying a pizza on Saturday means that he will not buy pizza on Sunday, what is the probability that he buys pizza on a given weekend (i.e if he buys pizza on one day, he is guaranteed to not buy a pizza the next day)? Note that the probability that he buys a pizza on Sunday, 0.6, is *not* a conditional probability, i.e. it is not conditioned on whether he buys a pizza on Saturday.

   (c) Suppose we don't know how Andy's pizza purchasing habits on Sundays depends on whether he bought a pizza on the preceding Saturday. Given that Andy buys pizza on a given weekend with probability 0.65, what is the probability that he buys pizza both days?

3. (*Note: before students attempt this problem, it is highly recommended that they attempt problem 1d*)

   Suppose Brian places m rectangles on an $n \times n$ grid, each chosen uniformly at random and independent from all others. Lower bound the probability that no rectangle is enveloped by another rectangle (that is, without any edges touching, the area of one rectangle is located entirely inside the area of another rectangle). Assuming n is sufficiently large, express this bound as a function of m