Randomness, Probability Basics, PMFs

Submit a PDF of your answers to Canvas

1. ***Laws of Probability.*** Prove or disprove if the following statements are true in general:

a) $\mathbb{P}(A^c, B^c) = \mathbb{P}(A^c) - \mathbb{P}(B)$

b) $\mathbb{P}(A^c \cap B^c) \geq \mathbb{P}(A^c) - \mathbb{P}(B)$

c) $\mathbb{P}(A) - \mathbb{P}(B) \leq \mathbb{P}(B^c) - \mathbb{P}(A^c)$

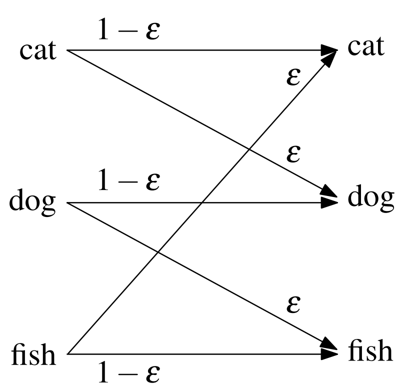
2. ***The Monty Hall Paradox (Adapted from Vos Savant, 1990).***

You're on a gameshow, and given the choice of three doors, denoted A, B, and C. Behind one is gold, and behind the other two is nothing. Imagine you a pick door. Monty Hall, the host, knows what is behind each door, and opens one of the remaining two doors and shows you there is nothing behind it. He then says to you - 'Do you want pick another door?'

a) Before you make your decision to move, what is the sample space for this random experiment? *Hint: write out all possible outcomes of the triple (door with gold, door selected by player, door that Monty opens).*

b) Assume that the gold was placed at random, and that your initial choice of door is independent of the gold placement, and that Monty's choice is random among possible empty doors. Specify the probability measure for the outcomes, and use it to compute the probability of winning if the contestant stays or moves.

3. ***Tertiary (3-class) Classification problem and crossover probabilities***. You build a classifier that aims to classify pictures of cats, dogs, and fish. The classifier succeeds at classifying cats as cats with probability $1-\varepsilon$, but incorrectly labels cat pictures as dogs with probability ε, and so on, as specified in the transition diagram below:



Pictures of cats occur with probability 1/2, pictures of dogs with probability 1/4, and pictures of fish with probability 1/4.

a) Find, as a function of ε, the probabilities of the labels.

b) Suppose the classifier predicts dog. What is the probability that the input was cat? Fish? Dog? Express your answer a function of ε.

4. ***Multinomial Probabilities by Enumeration***. The multinomial distribution is an extension of the binomial distribution. For n independent trials of an experiment that leads to one of k possible categories (i.e, the roll of a $k = 6$ sided die, n times), the multinomial distribution models the probability of the counts of each category.

Imagine that you toss a fair, 6-sided die 7 times.

a) ***Categorical outcomes.*** What is Ω in this experiment? Write a script to enumerate the possible outcomes (note that you can do this with nested for loops, or a recursive function). Your script should produce a list of tuples, where each tuple has 7 elements indicating the sequence of results from tossing the dice seven times.

b) How many total outcomes are there? What is the probability of any given outcome?

c) ***Multinomial outcomes***. Suppose you are only interested in the number of times each of the categories (sides of the die, 1-6) appeared in the 7 throws. This *count or frequency* of each categories is a multinomial random variable. What is Ω for this experiment?

d) Write a script to convert the outcomes from (a) into this format, keeping track of how many outcomes in (a) map to a single multinomial outcome. It may be useful to use a dictionary data structure.

e) How many possible multinomial outcomes are there?

f) What are the most probable multinomial outcomes, and what is their probability?

g) (Optional) Derive a general expression for the total number of multinomial outcomes given k categories and n trials.