Math 318 – homework 1 solutions

Problem 1. A coin is tossed until either two Tails appear successively, or until the fifth toss, whichever comes first. The outcome is the resulting sequence of coins.
(a) Write down the sample space, and determine the probability of each outcome in the sample space.
(b) For each $i \in \{2, 3, 4, 5\}$, let E_i be the event that the coin is tossed exactly i times. Determine $P(E_i)$ for each i.

Solution.
(a) The sample space is $S = \{$TT, HTT, HHTT, THTT, HHHHH, HHHHT, HHHTH, HHHTT, HHTHH, HHTHT, HTHHH, HTHHT, HTHTH, HTHTT, THHHH, THHHT, THHTH, THHTT, THTHH, THTHT$\}$.
We have $P(TT) = 1/4$, $P(HTT) = 1/8$, $P(HHTT) = P(THTT) = 1/16$ and the other 16 elements have probability 1/32.
(b) $E_2 = \{TT\}$ so $P(E_2) = 1/4$. Similarly, $P(E_3) = P(\{HTT\}) = 1/8$ and $P(E_4) = P(\{HHTT, THTT\}) = 2/16$. The remaining points are E_5 with $P(E_5) = 16/32 = 1/2$.

Problem 2. An herpetology graduate student is sent to estimate the number of frogs in a pond. She captures 40 frogs, marks each with a dot of paint, and then releases them. The next day, she goes back and captures another sample of 50. She finds that 14 of the frogs were previously marked, and 36 unmarked.
(a) Assuming that the frog population has size n, and that every frog is equally likely to be captured, determine the probability $L(n)$ that a sample of 50 frogs will contain exactly 14 marked ones.
(b) Show that the function $L(n)$ is increasing up to some n_*, and decreasing afterwards. Hint: when does the inequality $L(n)/L(n-1) \le 1$ hold?
(c) Find the maximum likelihood estimate for n; that is the value n_* which maximizes $L(n)$.

Solution.
(a) Since there are 40 marked and $n - 40$ unmarked frogs, the number of ways to pick 14 marked and 36 unmarked frogs is $\binom{40}{14}\binom{n-40}{36}$. The sample space has size $\binom{n}{50}$. Therefore

$$L(n) = \frac{\binom{40}{14}\binom{n-40}{36}}{\binom{n}{50}}.$$

(b) Note that $L(n)$ is positive for all $n \ge 76$. We find after cancellations that

$$\frac{L(n)}{L(n-1)} = \frac{(n-40)(n-50)}{n(n-76)} = 1 - \frac{14n - 2000}{n(n-76)}.$$

It follows that $L(n) > L(n-1)$ for $n < \frac{2000}{14} = 142.8\ldots$ and $L(n) < L(n-1)$ for larger n. Therefore L is maximized at $n_* = 142$ where $L(n_*) = 0.154\ldots$.

Remark. If the 50 frogs are sampled with replacement, the likelyhood that 14 of them are marked is $\binom{50}{14}\left(\frac{40}{n}\right)^{14}\left(1 - \frac{40}{n}\right)^{36}$. This is not equal to $L(n)$, but has is also maximized at the same n_*. See plot in python notebook.

Problem 3. (a) Compute the probability that a poker hand contains:
(i) one pair ($aabcd$ with a, b, c, d) distinct face values; answer: 0.42)
(ii) two pairs ($aabbc$ with a, b, c distinct face values; answer: 0.047)
(b) Poker dice is played by simultaneously rolling 5 dice. Compute the probabilities of the following outcomes:
(i) one pair ($aabcd$ with a, b, c, d) distinct face values; answer: 0.46)
(ii) two pairs ($aabbc$ with a, b, c distinct face values; answer: 0.23)

Solution.

(a) The sample space is 5 unordered cards, with $|S| = \binom{52}{5}$.

(i) $|E| = 13 \cdot \binom{4}{2} \cdot \binom{12}{3} \cdot 4^3$, so
$$P(E) = \frac{13 \cdot \binom{4}{2} \cdot \binom{12}{3} \cdot 4^3}{\binom{52}{5}}.$$

(ii) $|E| = \binom{13}{2} \cdot \binom{4}{2}^2 \cdot 11 \cdot 4$, so
$$P(E) = \frac{\binom{13}{2} \cdot \binom{4}{2}^2 \cdot 11 \cdot 4}{\binom{52}{5}}.$$

(b) The sample space is 5 dice in order, with $|S| = 6^5$.

(i) $|E| = \binom{5}{2} \cdot 6 \cdot 5 \cdot 4 \cdot 3 = 3600$, since we have to pick which two dice are equal, and then 4 different values for the dice. Therefore $P(E) = 3600/|S|$.

(ii) There are $\binom{5}{2,2,1} = 30$ ways to splie 5 dice into two pairs and a single, however swapping the pairs gives the same partition, so there are only 15 different ways. For each of these there are $6 \cdot 5 \cdot 4 = 120$ ways to pick the values, so $|E| = 15 \cdot 120 = 1800$ and $P(E) = 1800/|S|$.

Problem 4. A coin is tossed $2n$ times. Let p_n be the probability that exactly half the outcomes are heads.

(a) Find a formula for p_n.

(b) Calculate p_{n+1}/p_n, and show that p_n is decreasing in n (i.e., $p_{n+1} < p_n$).

(c) We are interested in the asymptotics of p_n. We use the notation $a_n \sim b_n$ if $\lim a_n/b_n = 1$. Using Stirling's formula
$$n! \sim \sqrt{2\pi n}(n/e)^n,$$
prove that there is some α so that $p_n \sim \alpha/\sqrt{n}$, and find the value of α.

Solution.

(a) All sequences of coins have the same probability 2^{-2n}. We have $p_n = \binom{2n}{n} 2^{-2n}$.

(b) Simplifying the factorials gives
$$\frac{p_{n+1}}{p_n} = \frac{2n+1}{2n+2} < 1.$$

(c) We have
$$p_n = \frac{(2n)!}{4^n n!^2} \sim \frac{\sqrt{2\pi(2n)}(2n/e)^{2n}}{2^{2n}(\sqrt{2\pi n}(n/e)^n)^2} = \frac{1}{\sqrt{\pi n}}.$$
Thus the claim holds with $\alpha = 1/\sqrt{\pi}$.

Problem 5. The number of ways to place n *distinguishable* balls in m urns is m^n, since each ball can be placed in any one of the m urns. The multinomial coefficient $\binom{n}{n_1,\ldots,n_m} = \frac{n!}{n_1!\ldots n_m!}$ counts the number of ways that n_i balls are in urn i for each $i = 1, 2, \ldots, m$, so when each ball is randomly assigned to an urn, the probability that n_i balls are in urn i, for each i, is equal to $\binom{n}{n_1,\ldots,n_m} m^{-n}$. Systems described by these probabilities are said to obey Maxwell–Boltzmann statistics.

(a) Suppose instead that the balls are *indistinguishable*; now we speak of Bose–Einstein statistics. When there are $m = 2$ urns, the number of ways of putting the n balls in the 2 urns is $n + 1$, because an outcome is specified by saying how many balls are in urn 1 and the possibilities are $\{0, 1, 2, \ldots, n\}$. For the case of general $m \geq 1$, how many ways are there to place n indistinguishable balls in m urns?
Hint: This is the number of ways to arrange $m - 1$ barriers among a row of n balls, e.g., for $n = 7$ and $m = 3$ the configuration with $n_1 = 0, n_2 = 2, n_3 = 5$ is graphically described by $| \bullet\bullet | \bullet\bullet\bullet\bullet\bullet$.

(b) Indistinguishable particles are said to obey Fermi–Dirac statistics if all arrangements that have at most one ball per urn have the same probability, and those are the only arrangements. How many ways can n of these particles be put into m urns (assuming $m \geq n$

Problem 6. write a program in python (A Jupyter notebook may be convenient) that will do the following.

(a) Write a function `birthday(n)` that:

(i) Generates a list containing n numbers uniformly distributed on $\{1, 2, \ldots, 365\}$ (think of this as the list of birthdays of n people, excluding leap year).

(ii) Returns 1 (or `True`) if there is at least one pair of people with coinciding birthdays (a "match") and 0 (or `False`) otherwise.

(b) For each n, from 2 to 60, run the function `birthday(n)` 1000 times, and compute the proportion $X(n)$ of the 1000 times in which there was a match.

(c) Let $Y(n)$ be the actual probability of a match:

$$Y(n) = 1 - \frac{365 \cdot 364 \cdots (365 - n + 1)}{365^n}.$$

In a single graph, plot of $X(n)$ and $Y(n)$ for $n \in [2, 60]$.

(d) Repeat the steps above for Martians. (Hint: The Martian year has 669 Martian days.)