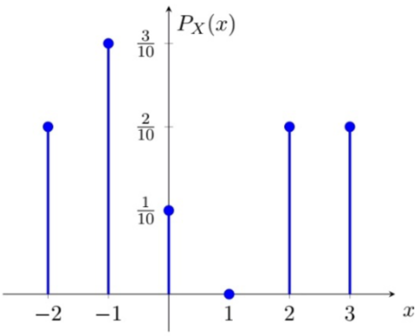
Jilin Zheng U49258796

EK381 Homework #3

Problem 3.1 (Video 2.3 - 2.6, Lecture Problem)



Let X be a discrete random variable with probability mass function (PMF) as above. Let event $A = \{-2, 1, 3\}$.

(a) Given that $\{X \in A\}$ occurs, what is the conditional probability that $X > 1$, that is $\mathbb{P}[X > 1 | X \in A]$?

(b) Determine $\mathbb{E}[X]$ and $\mathbb{E}[3X + 2]$.

(c) Determine $\mathsf{Var}[X]$ and $\mathsf{Var}[2X - 1]$.

(a) $\mathbb{P}[X>1 \mid X\in A] = \dfrac{\mathbb{P}[X>1 \cap X\in A]}{\mathbb{P}[X\in A]} = \dfrac{2/10}{2/10 + 0 + 2/10} = 2/10 \cdot 10/4 = 1/2$

(b) $\mathbb{E}[X] = \sum_{x\in R_X} x \cdot P_X(x)$

$= -2 \cdot 2/10 + (-1)\cdot 3/10 + 0 \cdot 1/10 + 2 \cdot 2/10 + 3 \cdot 2/10 = -4/10 - 3/10 + 4/10 + 6/10 = 3/10$

$\mathbb{E}[3X+2] = 3\mathbb{E}[X] + 2 = 9/10 + 2 = 29/10$

(c) $\mathrm{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{x\in R_X} (x-\mu_X)^2 \cdot P_X(x)$

$= (-2-3/10)^2 \cdot 2/10 + (-1-3/10)^2 \cdot 3/10 + (0-3/10)^2 \cdot 1/10 + (2-3/10)^2 \cdot 2/10 + (3-3/10)^2 \cdot 2/10$

$= 3.61$

$\mathrm{Var}[2X-1] = 2^2\,\mathrm{Var}[X] = 14.44$

Problem 3.2 (Video 2.3 - 2.6, Quick Calculations)

Calculate each of the requested quantities.

(a) Your favorite sports team wins a game with probability $\frac{3}{4}$, independently of other games. Let X be the number of games they win out of 20. What kind of random variable is X? (Don't forget the parameters.) Calculate $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$.

(b) Let X be Poisson(λ) and assume that $\mathbb{E}[X] = 2$. Calculate λ, and $\mathbb{P}[X \leq 3]$ and $\mathbb{P}[X \leq 3 | X > 0]$.

(c) Roll a six-sided die until the first 2 appears. Let X denote the number of rolls. What kind of a random variable is X? (Don't forget the parameters.) Calculate $\mathbb{E}[2X - 1]$ and $\mathsf{Var}[2X - 1]$.

(d) Let X be a random variable with $\mathbb{E}[X] = -1$ and $\mathsf{Var}[X] = 4$. Let $Y = -3X + 2$. Calculate $\mathbb{E}[Y]$ and $\mathsf{Var}[Y]$.

(e) Let X be a random variable with $\mathbb{E}[X] = 0$ and $\mathsf{Var}[X] = 2$. Calculate $\mathbb{E}[X^2]$ and $\mathbb{E}[(2X - 1)^2]$.

(a) X is Binomial$(20, 3/4)$

$\mathbb{E}[X] = np = 3/4 \cdot 20 = 15$

$\mathbb{E}[X^2] = (\mathbb{E}[X])^2 + \mathrm{Var}[X]$

$= 15^2 + 3/4(20)(1 - 3/4)$

$= 228.75$

(b) since $\mathbb{E}[X] = 2$, $\lambda = 2$

$\mathbb{P}[X \leq 3] = P_X(0) + P_X(1) + P_X(2) + P_X(3)$

$= \frac{2^0}{0!}e^{-2} + \frac{2^1}{1!}e^{-2} + \frac{2^2}{2!}e^{-2} + \frac{2^3}{3!}e^{-2}$

$= 1/e^2 + 2/e^2 + 2/e^2 + 4/3e^2$

$= 0.857$

$\mathbb{P}[X \leq 3 | X > 0] = P_X(1) + P_X(2) + P_X(3) / \mathbb{P}[\{X \in B\}]$

$= [2/e^2 + 2/e^2 + 4/3e^2] / \mathbb{P}[X > 0]$

$= 0.835$

$\underbrace{\sum_{x \in C} P_{X|B}(x)}$

$\frac{P_X(x)}{\mathbb{P}[\{X \in B\}]}, \quad x \in B$

$\frac{2/e^2 + 2/e^2 + 4/3e^2}{1 - 1/e^2} =$

(c) X is a Geometric$(1/6)$ random variable

$\mathbb{E}[2X-1] = 2\mathbb{E}[X] - 1 = 2 \cdot 1/p - 1 = 2 \cdot 6 - 1 = 11$

$\mathrm{Var}[2X-1] = 2^2\,\mathrm{Var}[X] = 4 \cdot \frac{1-p}{p^2} = 4\frac{5/6}{(1/6)^2} = 4 \cdot 5/6 \cdot 36 = 120$

(d) $\mathbb{E}[X] = -1$, $\mathrm{Var}[X] = 4$; $Y = -3X + 2$

$\mathbb{E}[Y] = \mathbb{E}[-3X+2] = -3\mathbb{E}[X] + 2 = 3 + 2 = 5$

$\mathrm{Var}[Y] = \mathrm{Var}[-3X+2] = 9\,\mathrm{Var}[X] = 36$

(e) $\mathbb{E}[X] = 0$, $\mathrm{Var}[X] = 2$

$\mathbb{E}[X^2] = \mathrm{Var}[X] + (\mathbb{E}[X])^2 = 2$

$\mathbb{E}[(2X-1)^2] = \mathbb{E}[4X^2 - 4X + 1] = 4\mathbb{E}[X^2] - 4\mathbb{E}[X] + 1$

$= 8 - 0 + 1 = 9$

Problem 3.3 (Video 2.3 - 2.6, **Fall 2020 Exam 1 Problem**)

You are practicing your free throws for an upcoming basketball game. Every throw is successful with probability 2/3, independently of the others. Let X denote the number of successful throws out of 5.

(a) What kind of random variable is X? (Don't forget the parameters.)

(b) What is the probability that you successfully make at least 3 out of the 5 free throws?

(c) Given that you successfully make at least 3 out of 5 free throws, what is the probability that you successfully make exactly 3 out of 5?

(d) What is the probability of scoring exactly 3 **consecutive** free throws within the set of 5?

(e) You keep practicing with sets of 5 free throws. What is the average number of sets until your first set where you miss every free throw?

(a) X is a Binomial$(5, 2/3)$ random variable

(b) $\mathbb{P}[X \geq 3] = P_X(3) + P_X(4) + P_X(5)$

$$= \binom{5}{3}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^2 + \binom{5}{4}\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^1 + \binom{5}{5}\left(\frac{2}{3}\right)^5\left(\frac{1}{3}\right)^0$$

$$= 0.790$$

(c) $\mathbb{P}[X=3 \mid X \geq 3] = \sum_{x \in C} P_{X|B}(x)$

$$= \frac{P_X(3)}{0.790} = \frac{\binom{5}{3}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^2}{0.790} = 0.417$$

(d) $3\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^2 = 0.0988$

(e) $\left(\frac{1}{3}\right)^5 = \frac{1}{243}$ (probability of all misses)

Let Y be a Geometric$\left(\frac{1}{243}\right)$ random variable

$\mathbb{E}[Y] = \frac{1}{1/243} = 243$ sets

Problem 3.4 (**Video 2.4**, **Lecture Problem**) You are interested in calculating the probability that your favorite[1] Game of Thrones character is eliminated in episode X. You have decided to model X as a Geometric(1/4) random variable.

(a) Unfortunately, you have learned a spoiler: your favorite character does not appear in episode 4 or beyond. What is the conditional PMF $P_{X|B}(x)$ of X given the event $B = \{X < 4\}$?

(b) Given this spoiler, what is the probability that your favorite character is eliminated in one of the first two episodes?

(c) Given this spoiler, what is the expected value of X conditioned on the event B?

(d) Let's consider yet another scenario: After watching the show for 2 episodes, you are happy to see that your favorite character *has not* been eliminated yet. What is the conditional PMF $P_{X|C}(x)$ of X given the event $C = \{X > 2\}$?

[1]Somehow, you have already managed to decide on a favorite character before watching any episodes.

(e) Let $Y = X - 2$ be the number of additional episodes after the 2nd that it takes for your favorite character to be eliminated. Using part (d), quickly determine the conditional PMF $P_{Y|C}(y)$ of Y given the event $C = \{X > 2\}$. Determine the family of random variables this conditional PMF belongs to, along with the associated parameter(s).

(f) Using what you learned in part (e), determine the conditional mean $\mathbb{E}[X|C]$.

(a)

$$P_{X|B}(x) = \frac{P_X(x)}{\mathbb{P}[B]} = \begin{cases} \frac{1}{4}\cdot\frac{64}{37} = \frac{16}{37}, & x=1 \\ \frac{3}{16}\cdot\frac{64}{37} = \frac{12}{37}, & x=2 \\ \frac{9}{64}\cdot\frac{64}{37} = \frac{9}{37}, & x=3 \end{cases} \qquad 0, \text{ otherwise}$$

$$\mathbb{P}[B] = P_X(1)+P_X(2)+P_X(3) = \tfrac{1}{4}\left(\tfrac{3}{4}\right)^0 + \tfrac{1}{4}\left(\tfrac{3}{4}\right)^1 + \tfrac{1}{4}\left(\tfrac{3}{4}\right)^2 = \tfrac{1}{4} + \tfrac{3}{16} + \tfrac{9}{64} = \tfrac{37}{64}$$

(b) $P_{X|B}(1) + P_{X|B}(2) = \frac{16}{37} + \frac{12}{37} = \frac{28}{37}$

(c) $\mathbb{E}[X|B] = \sum_{x\in B} x\cdot P_{X|B}(x) = 1\cdot\frac{16}{37} + 2\cdot\frac{12}{37} + 3\cdot\frac{9}{37} = \frac{16}{37} + \frac{24}{37} + \frac{27}{37} = \frac{67}{37}$

(d) $C = \{x>2\}, C' = \{X \le 2\}$

$$P_{X|C}(x) = \frac{P_X(x)}{\mathbb{P}[C]} = \begin{cases} \frac{p(1-p)^{x-1}}{9/16}, & x>2 \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbb{P}[C'] = P_X(1)+P_X(2) = \tfrac{1}{4}\left(\tfrac{3}{4}\right)^0 + \tfrac{1}{4}\left(\tfrac{3}{4}\right)^1 = \tfrac{1}{4} + \tfrac{3}{16} = \tfrac{7}{16}$$

$$\mathbb{P}[C] = 1 - \mathbb{P}[C'] = \tfrac{9}{16}$$

(e) $Y = X-2$, $C = \{X>2\} = \{Y+2>2\} = \{Y>0\}$

$$P_{Y|C}(y) = \frac{P_Y(y)}{\mathbb{P}[\{y\in C\}]} = \begin{cases} \frac{p(1-p)^{y-1}}{1}, & y>0 \\ 0, & \text{otherwise} \end{cases}$$

Y is a Geometric$\left(\frac{1}{4}\right)$ random variable

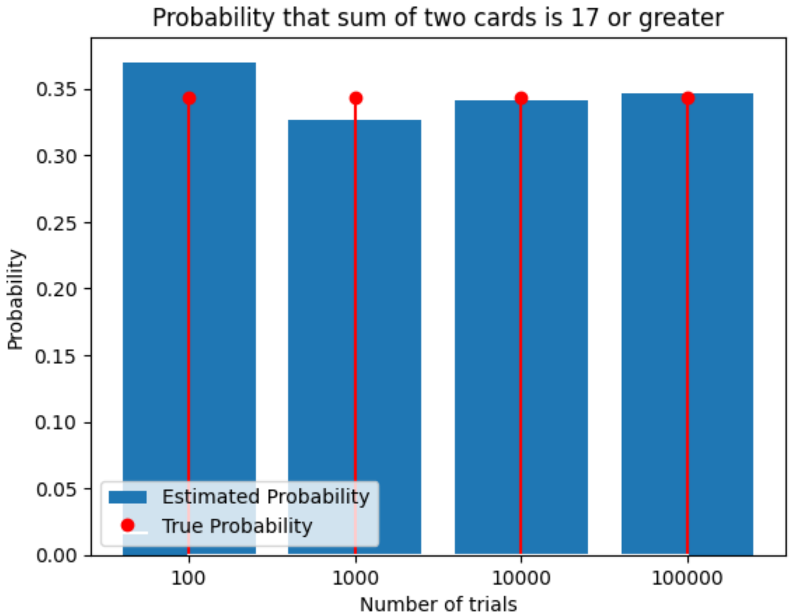
(f) $\mathbb{E}[X|C] = \mathbb{E}[Y+2|C] = \mathbb{E}[Y|C] + 2 = \frac{1}{1/4} + 2 = 6$

The use of simulation to generate probabilties for counting events. We know that, by counting the number of times that certain combinations of outcomes occur, we can compute the probability that those outcomes occur analytically. This is typically done when analyzing card games such as Blackjack or Poker, where one computes the probability of getting a set of cards that satisfies some criterion. An alternative is to simulate the experiment, and observe the fraction of times that the cards satisfy the criteria. We implement the approach below for three problems, and compare the simulation results with the analytical computations of probability, for different lengths of experiments.

In [ ]:
```
import matplotlib.pyplot as plt
import numpy as np
import math
```

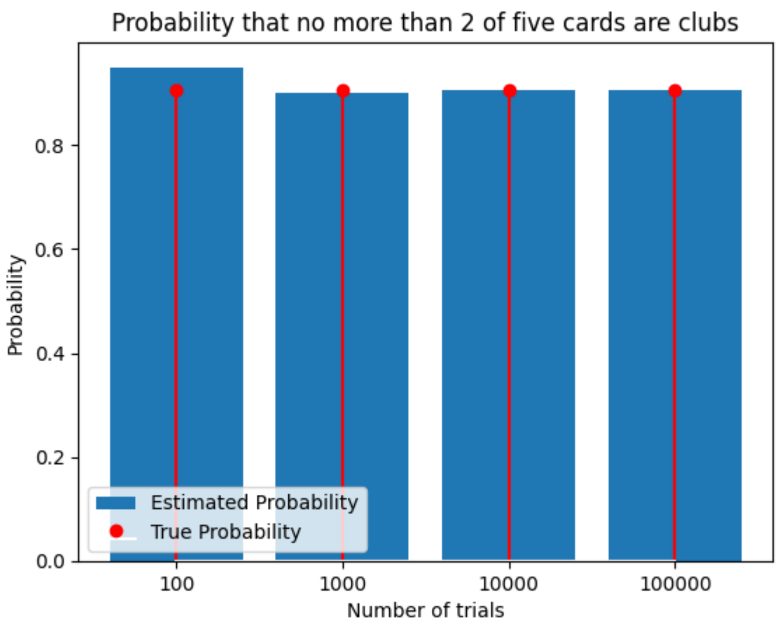
Part a. Run the program below to estimate the probability that the first five cards of a randomly shuffled deck have no more than two clubs. Remember that clubs are cards numbered 1-13. Do it for numtrials = 100, 1000, 10000, 100000. Plot the bar chart of probability estimate versus number of trials. Plot as a stem plot the true probability, which is (nchoosek(39,5) + 13*nchoosek(39,4) + nchoosek(13,2)*nchoosek(39,3))/nchoosek(52,5).

In [ ]:
```
N = 52 # number of cards
trials = [100,1000,10000,100000]  # number of cases
probability = np.zeros(4,);
cards = list(range(1,53)) # a list with elements 1, 2, ..., 52

for i in range(4):
    numtrials = trials[i]
    successes = 0

    for k in range(numtrials):

        p = np.random.permutation(cards) # generates random permutation of numbers from 1 to 52: a shuffle

        total = 0
        first = (p[0]-1) % 13     # modular arithmetic
        if (first == 0):
            total = total + 11
        elif (first >= 9):
            total = total + 10
        else:
            total = total + first + 1

        second = (p[1]-1) % 13
        if (second == 0):
            total = total + 11
        elif (second >= 9):
            total = total + 10
        else:
            total = total + second + 1

        if (total >= 17) and (total <= 21):
            successes =  successes + 1

    probability[i] = successes/numtrials

trueProb = (4*32 + 4*16 + 4*20 + 4* 16 + math.comb(16,2))/math.comb(52,2)



fig = plt.figure()
plt.bar(np.arange(4),probability)
plt.xticks(np.arange(4), trials)

plt.xlabel('Number of trials')
plt.ylabel('Probability')
plt.stem(np.arange(4),trueProb*np.ones(4,),linefmt='r',markerfmt='ro',basefmt='w')
plt.title(f"Probability that sum of two cards is 17 or greater")
plt.legend(['Estimated Probability','True Probability'])
plt.show()
fig.savefig("P3_5a.png",bbox_inches='tight')
```



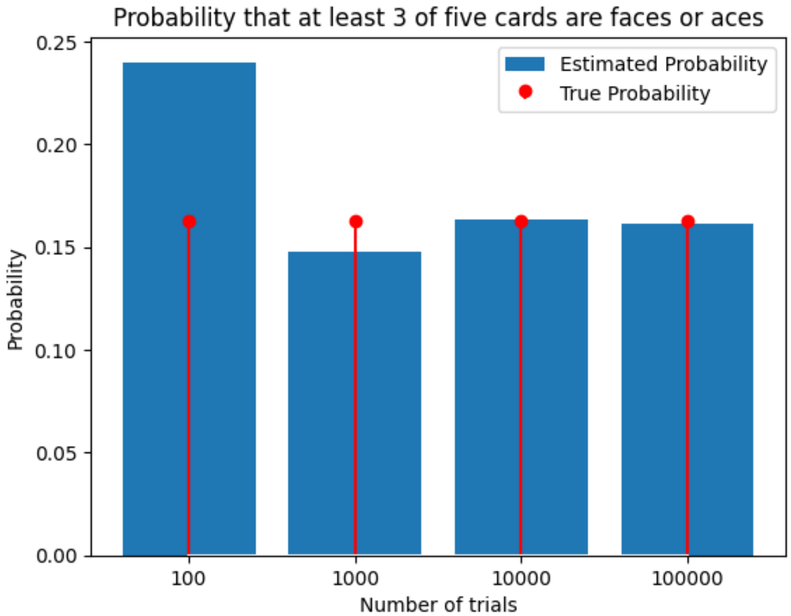
Part b. Write a program to estimate the probability that the first five cards of a randomly shuffled deck have no more than two clubs. Remember that clubs are cards numbered 1-13. Do it for numtrials = 100, 1000, 10000, 100000. Plot the bar chart of probability estimate versus number of trials. Plot as a stem plot the true probability, which is (nchoosek(39,5) + 13*nchoosek(39,4) + nchoosek(13,2)*nchoosek(39,3))/nchoosek(52,5).

In [ ]:
```
N = 52 # number of cards
trials = [100,1000,10000,100000]  # number of cases
probability = np.zeros(4,);
cards = list(range(1,53)) # a list with elements 1, 2, ..., 52

for i in range(4):
    numtrials = trials[i]
    successes = 0

    for k in range(numtrials):
        p = np.random.permutation(cards) # generates random permutation of numbers from 1 to 52: a shuffle

        #%%% Write your code here: You have to add a success if
        #%%% the number of clubs in the first five cards is less than or
        #%%% equal to 2.  See how it was done in the previous part.

        count_clubs = 0
        for l in range(5):
            if p[l] <= 13:
                count_clubs += 1

        if count_clubs <= 2:
            successes += 1

    probability[i] = successes/numtrials

trueProb = (math.comb(39,5)+ 13*math.comb(39,4) + math.comb(13,2)*math.comb(39,3))/math.comb(52,5)


fig = plt.figure()
plt.bar(np.arange(4),probability)
plt.xticks(np.arange(4), trials)

plt.xlabel('Number of trials')
plt.ylabel('Probability')
plt.stem(np.arange(4),trueProb*np.ones(4,),linefmt='r',markerfmt='ro',basefmt='w')
plt.title("Probability that no more than 2 of five cards are clubs")
plt.legend(['Estimated Probability','True Probability'])
plt.show()
fig.savefig("P3_5b.png",bbox_inches='tight')
```



Part c. Write a program to estimate the probability that you get 3 or more face cards or aces (jacks, queens, kings, aces) in the first five cards of a randomly shuffled deck. Remember that, if a card is numbered n, then (n-1) mod 13 = 0 for aces and (n-1) mod 13 >= 10 for jacks, queens and kings. Do it for numtrials = 100,1000, 10000, 100000. Plot the bar chart of probability estimate versus number of trials. Plot as a stem plot the true probability, which is (nchoosek(16,5)+ nchoosek(16,4)*nchoosek(36,1) + nchoosek(16,3)*nchoosek(36,2))/nchoosek(52,5).

In [ ]:
```
N = 52 # number of cards
trials = [100,1000,10000,100000]  # number of cases
probability = np.zeros(4,);
cards = list(range(1,53)) # a list with elements 1, 2, ..., 52

for i in range(4):
    numtrials = trials[i]
    successes = 0

    for k  in range(numtrials):

        p = np.random.permutation(cards) # generates random permutation of numbers from 1 to 52: a shuffle

        # %%% Write your code here: You have to add a success if
        # %%% the number of face cards or aces in the first in the first
        # %%% five cards is at least 3.
        # %%% See how it was done in the part (a).

        count_faces_or_ace = 0
        for l in range(5):
            if (p[l] - 1) % 13 == 0 or (p[l] - 1) % 13 >= 10:
                count_faces_or_ace += 1

        if count_faces_or_ace >= 3:
            successes += 1

    probability[i] = successes/numtrials

trueProb = (math.comb(16,5) + math.comb(16,4)*36 + math.comb(16,3)*math.comb(36,2))/math.comb(52,5)


fig = plt.figure()
plt.bar(np.arange(4),probability)
plt.xticks(np.arange(4), trials)

plt.xlabel('Number of trials')
plt.ylabel('Probability')
plt.stem(np.arange(4),trueProb*np.ones(4,),linefmt='r',markerfmt='ro',basefmt='w')
plt.title(f"Probability that at least 3 of five cards are faces or aces")
plt.legend(['Estimated Probability','True Probability'])
plt.show()
fig.savefig("P3_5c.png",bbox_inches='tight')
```



Part d. Part c is more difficult to predict than parts a or b because the event in part c is intrinsically more complex and more variable than the events in parts a and b. There are more, complex combinations of cards that satisfy the event in part c.