Prob 140 Fall 2018 Final Exam
A. Adhikari

1. Let X and Y have joint density given by

$$f(x,y) = \begin{cases} 3e^{-(2x+y)} & \text{for } 0 < x < y \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the density of X. Identify it as one of the famous ones and provide the parameters.
(b) Are X and Y independent? Explain your answer.
(c) For $c > 0$, find $P(Y - X > c)$, and hence identify the distribution of $Y - X$ as one of the famous ones. Don't forget the parameters.

2. A random coin lands heads with probability R picked according to the beta $(4, 6)$ distribution. The coin is tossed 20 times. Let H be the number of heads in the 20 tosses.
(a) What is the conditional distribution of H given $R = 0.35$?
(b) What is the posterior distribution of R given $H = 11$?
(c) The same coin will be tossed one more time. Given $H = 11$, what is the chance that it lands heads?
(d) Which is bigger, $P(R > 0.5)$ or $P(R > 0.5 \mid H = 11)$? Explain.

3. Dibya and Jason are UGSIs for a class that has 50 students labeled 1 through 50. Students 1 through 20 are in Jason's section. The remaining 30 students, labeled 21 through 50, are in Dibya's section.
All 50 students take the midterm and professor stacks the 50 graded exams on top of each other. Unfortunately for the UGSIs, she stacks the exams randomly so that all permutations of the 50 exams are equally likely.
Jason picks up the 20 exams at the top of the stack. Dibya picks up the remaining 30 exams.

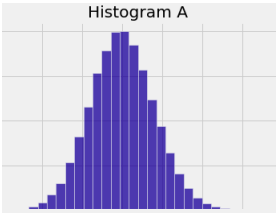
- Let J be the number of Jason's students whose exams are among the 20 that he picks up.
- Let D be the number of Dibya's students whose exams are among the 30 that he picks up.
- Let T be the total number of students whose exams are picked up by their own UGSI.

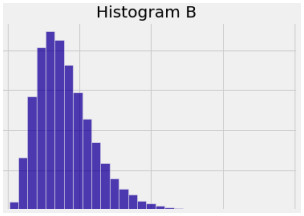
(a) What is the distribution of J?
(b) Are J and D independent? Explain your answer. If they are not independent, what is the relation between them?
(c) Find $E(T)$.
(d) Find $Var(T)$.

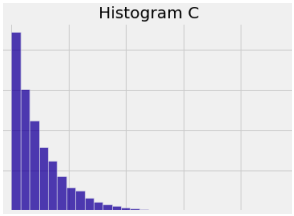
4. The code below was run after importing all the necessary libraries.

```
def exponential_sum(n):
    x =  stats.expon.rvs(scale = 1, size = n)  # n simulated i.i.d. exponential (1) variables
    return sum(x)

def simulate_sum(n, repetitions):
    sums = make_array()
    for i in np.arange(repetitions):
        sums = np.append(sums, exponential_sum(n))
    return sums

w = simulate_sum(400, 100000)

Table().with_column('w', w).hist(bins = 30)  # histogram has 30 bins of equal width
```

(a) The last line of code resulted in the display of one of the histograms A, B, and C. Which do you think it was, and why?







Each of the expressions in Parts (b)-(d) evaluated to a number that is approximately equal to one of the options below. In each part, choose the correct option and provide a brief explanation.

(i) 1/400 (ii) 1/20 (iii) 1/2 (iv) 1 (v) 2 (vi) 20 (vii) 400

(b) `np.mean(w)`

(c) `np.std(w)`

(d) `np.count_nonzero(w > 400) / 100000`

5. Let X and Y have a bivariate normal distribution with parameters $(\mu, \mu, \sigma^2, \sigma^2, \rho)$, where the correlation ρ is in the interval $(-1, 1)$.

(a) What is the joint distribution of X and $X - Y$?

(b) What is the conditional distribution of X given $X - Y = w$?

(c) Find $E(|X - Y|)$.

(d) Find $Var(|X - Y|)$.

6. Let U have uniform distribution on the interval $(0, 1)$. For a fixed $r > 0$, let $X, X_1, X_2, \ldots, X_{10}$ be independent and identically distributed beta $(r, 1)$ random variables.

Find the distribution of each of the random variables below. In each case, recognize it as one of the famous ones and provide its name and parameters.

(a) $-\log(U)$

(b) X^r

(c) $-\log\left((X_1 X_2 \cdots X_5)^r\right)$ [Strongly recommended: Use Parts (a) and (b).]

(d) $-2\log\left((X_1 X_2 \cdots X_5)^r\right)$

(e) $-\log\left((X_1 X_2 \cdots X_5)^r\right) - \log\left((X_6 X_7 \cdots X_{10})^r\right)$

7. According to a Hardy-Weinberg genetics model, each individual in a population belongs to one of three classes independently of all other indviduals. The model says that there is a parameter $\theta \in (0, 1)$ such that for every individual the chance of belonging to the three classes is as specified in the table below.

Class	A	B	C
Probability	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

Suppose there are n people. Let N_A be the number of people in Class A, N_B the number in class B, and N_C the number in class C, so that $N_A + N_B + N_C = n$.

(a) Suppose $n = 10$. Find $P(N_A = 2, N_B = 5, N_C = 3)$ in terms of θ.

(b) Now let n be any positive integer, and suppose your data are N_A, N_B, and N_C with $N_A + N_B + N_C = n$ as defined above. For $\theta \in (0, 1)$, find the likelihood $lik(\theta)$.

(c) Find the maximum likelihood estimate (MLE) of θ.

(d) Is the MLE an unbiased estimate of θ? Prove your answer.

8. Let $U_1, U_2, U_3, \ldots$ be independent uniform $(0,1)$ random variables. Let N have the Poisson (μ) distribution, independently of $U_1, U_2, U_3, \ldots$.
(a) Find $E\big(\frac{1}{N+1}\big)$. Be careful, and show all your calculations.
(b) Define the random variable M by

$$M \;=\; \begin{cases} \min\{U_1, U_2, \ldots, U_n\} & \text{if } N = n \geq 1 \\ 1 & \text{if } N = 0 \end{cases}$$

Find $E(M)$.

9. Consider three independent random variables X, Y, and Z such that:

- X has the exponential (α) distribution
- Y has the gamma (r, λ) distribution
- Z is non-negative with density f_Z and moment generating function M_Z that is finite on $(-\infty, \infty)$

(a) Use the survival function of X to find $P(X > Y)$.
(b) Find $P(X > Z)$ in terms of M_Z.