Midterm Exam

Last Name	First Name	SID

Rules.

- You have 80 minutes (12:40pm - 2:00pm) to complete this exam.
- The maximum you can score is 120.
- The exam is not open book, but you are allowed one side of a sheet of handwritten notes; calculators will be allowed. No phones.
- No form of collaboration between the students is allowed. If you are caught cheating, you may fail the course and face disciplinary consequences.

Please read the following remarks carefully.

- Show all work to get any partial credit.
- Take into account the points that may be earned for each problem when splitting your time between the problems.

Problem	points earned	out of
Problem 1		45
Problem 2		35
Problem 3		20
Problem 4		20
Total		120

Problem 1: Answer these questions briefly but clearly.

(a) [10] X is a uniform random variable over $[0,1]$. Calculate the coefficient of correlation, $\rho(X,Y)$, for X and $Y=X^2$.

$\mathbb{E}[X^k]=(k+1)^{-1}$ for $k=1,2,3,4$ so $\operatorname{var} X = 1/12$, $\operatorname{var} X^2 = 1/5-1/9=4/45$, $\operatorname{cov}(X,Y) = \mathbb{E}[X^3]-\mathbb{E}[X]\,\mathbb{E}[X^2]=1/4-1/6=1/12$, so $\rho(X,Y)=(1/12)/\sqrt{(1/12)(4/45)}=\sqrt{15}/4$.

(b) [10] X is uniformly distributed in the interval $[a,b]$, $a\geq 0$. Find $f_Y(y)$ for $Y=X^2$.

For $x\in[a^2,b^2]$, $\mathbb{P}(X^2\leq x)=\mathbb{P}(X\leq\sqrt{x})=(\sqrt{x}-a)/(b-a)$, so $f_Y(y)=(2\sqrt{x}(b-a))^{-1}$ for $y\in[a^2,b^2]$.

(c) [10] Let X be uniformly distributed in the interval $[0,1]$. For what function g is $Y=g(X)$ an exponential random variable with parameter 1?

We need the inverse function of the CDF $F(x)=1-\exp(-x)$ for $x>0$, which is $g(x)=\ln(1/(1-x))$. Alternatively since X has the same distribution as $1-X$, we could also choose $g(x)=-\ln x$.

(d) [15] Bob plays the following game: He is given a real number number X drawn from the uniform distribution over $[0,1]$. Then Bob must keep drawing numbers from the same distribution until he has drawn a number $>X$. If Bob has drawn N numbers, he wins $N-1$ dollars. If W represents the amount Bob wins, what is $\mathbb{E}[W]$?

Since $N\mid X\sim\text{Geometric}(1-X)$, $\mathbb{E}(N\mid X)=(1-X)^{-1}$ and $\mathbb{E}[N]=\mathbb{E}[\mathbb{E}(N\mid X)]=\mathbb{E}[(1-X)^{-1}]$ which does not exist. So, $\mathbb{E}[W]$ does not exist either.

Problem 2

(a) [10] Alice and Bob are both taking EE 126 in a classroom with r rows of m seats each. If there are exactly $r\ m$ students attending each class (so there are no empty seats) and the students take their seats at random, what is the probability that Alice and Bob are sitting on the same row in adjacent seats?

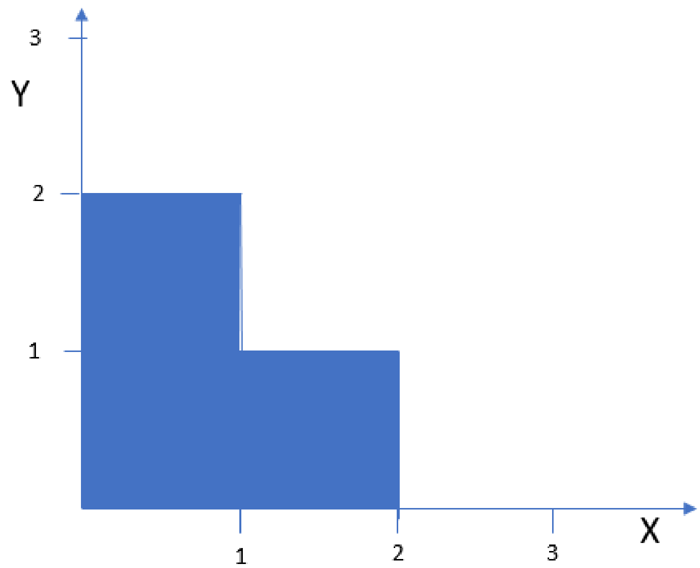
Out of $\binom{rm}{2}$ total unordered choices of their seats, $r(m-1)$ choices are such that they seat in adjacent seats so the probability is $r(m-1)/\binom{rm}{2}$.

(b) [10] A deck of 52 cards is dealt to four players. What is the probability that one of the players is dealt all four aces?

Split into the disjoint union of four events, where each event is the event that a particular player receives all four aces. This occurs with probability $\binom{48}{9}/\binom{52}{13}$ so the overall probability is $4\binom{48}{9}/\binom{52}{13}$.

(c) [15] A coin shows Heads with probability p. Let X_n be the number of tosses required to get n heads in a row. Find $\mathbb{E}[X_n]$. (Hint: Use conditioning.)

For $k = 1, \ldots, n$, let Y_k be the number of tosses required to get n heads in a row, conditioned on starting from $n-k$ heads already tossed. Thus, $\mathbb{E}[X_n] = \mathbb{E}[Y_n]$. Also, $\mathbb{E}[Y_k] = 1+p\,\mathbb{E}[Y_{k-1}]+(1-p)\,\mathbb{E}[Y_n]$. So, $\mathbb{E}[Y_n] = 1+p\,\mathbb{E}[Y_{n-1}]+(1-p)\,\mathbb{E}[Y_n] = 1+p+p^2\,\mathbb{E}[Y_{n-2}]+(1-p)(1+p)\,\mathbb{E}[Y_n] = \cdots = 1+p+\cdots+p^{n-1}+(1-p)(1+p^2+\cdots+p^{n-1})\,\mathbb{E}[Y_n] = (1-p^n)/(1-p)+(1-p^n)\,\mathbb{E}[Y_n]$, so $p^n\,\mathbb{E}[Y_n] = (1-p^n)/(1-p)$ and $\mathbb{E}[Y_n] = (p^{-n}-1)/(1-p)$.



Problem 3

(a) **[10]** The joint density function $f_{X,Y}(x, y)$ in the Figure. It has the same value in the shaded region and is zero outside the shaded region. What is $f_Z(z)$ for $Z = X + Y$?

The density increases linearly, starting from 0, in $[0, 2]$, and then decreases linearly to 0 in $[2, 3]$. Enforcing the condition that the density has area 1, the density is $z/3$ in $[0, 2]$ and $2 - 2z/3$ in $[2, 3]$.

(b) **[10]** X and Y are independent and identically distributed exponential random variables with parameter λ. Let $Z = X - Y$. Find $f_Z(z)$, $\operatorname{var}(Z)$, and $M_Z(s)$.

First note that $X - Y$ has the same distribution as $Y - X = -(X - Y)$ so the distribution is necessarily symmetric around the origin. Next compute for $x \geq 0$, $\mathbb{P}(X - Y \geq x) = \mathbb{E}[\mathbb{P}(X \geq x + Y \mid Y)] = \mathbb{E}[\exp(-\lambda(x + Y))] = \exp(-\lambda x)M_Y(-\lambda) = \exp(-\lambda x)\lambda/(\lambda + \lambda) = \exp(-\lambda x)/2$. So $f_Z(z) = \lambda \exp(-\lambda|z|)/2$. By independence, $\operatorname{var} Z = \operatorname{var} X + \operatorname{var} Y = 2/\lambda^2$. Finally $M_Z(s) = M_X(s)M_Y(-s) = \lambda^2/((\lambda - s)(\lambda + s)) = \lambda^2/(\lambda^2 - s^2)$.

Problem 4

(a) [10] Let $X_1, \dots, X_N$ be independent exponential random variables with parameter λ. Let $X_{\max} = \max\{X_1, \dots, X_N\}$ and $X_{\min} = \min\{X_1, \dots, X_N\}$. Use the conditioning and the memoryless property of the exponential distribution to find $\mathbb{P}(X_{\max} - X_{\min} \le r)$ for $r \ge 0$.

Conditioned on $X_{\min}$, $X_{\max} - X_{\min}$ has the distribution of the maximum of $n-1$ exponential random variables with parameter λ due to memorylessness. So $\mathbb{P}(X_{\max} - X_{\min} \le r) = \mathbb{E}[\mathbb{P}(X_{\max} - X_{\min} \le r \mid X_{\min})] = \mathbb{E}[(1 - \exp(-\lambda r))^{n-1}] = (1 - \exp(-\lambda r))^{n-1}$.

(b) [10] Bob has n pairs of different colored socks in the dryer. He pulls out k socks at random. Of the k socks he has X matching pairs. What is $\mathbb{E}[X]$?

For $i = 1, \dots, n$, X_i is the indicator that the ith pair is drawn by Bob. So $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n\binom{2n-2}{k-2}/\binom{2n}{k} = k(k-1)/(2(2n-1)) = \binom{k}{2}/(2n-1)$. The last expression gives another interpretation: define an indicator Y_i, $i = 1, \dots, \binom{k}{2}$ for each subset of size 2 in the chosen k socks for the event that two socks form a pair. The probability that any two given socks forms a pair is $1/(2n-1)$, so the expected number of pairs is $\binom{k}{2}/(2n-1)$.