Statistics 251 Name: ____________________________
Section 1, Autumn 2020
Final Exam Practice CNetID: ____________________________
Date: N/A
Time: N/A

Instructions: This exam contains 6 problems. Please make sure you attempt all problems.

Present your solutions in a **legible, coherent** manner. Unless otherwise specified, you should show your work; you will be evaluated on both your reasoning and your answer. Unsupported or illegible solutions may not receive full credit.

You may not print this test paper out. In that case, you should sign at the beginning of you uploaded document to affirm you have followed all the rules. **You should also draw a box at start of each question with your final answer**. An example is shown as below, *slanted* sentences are ones to be replaced with your own anwser.

Name: *(your full name in block letters)* **CNetID:** *(your CNetID)*
I am here to affirm that I have followed all the rules listed on the original midterm exam paper.
(your signature and the date)

Problem 1 (a) Answer: *(write your final answer here)*

Your solution/explanation to the final answer should follow the box.

Problem 1 (b) Answer: *(write your final answer here)*

Your solution/explanation to the final answer should follow the box.

Due to the coronavirus situation, this exam will be over Zoom. The use of outside material including books, notes, and electronic smart devices is not allowed. You are not allowed to discuss anything about this exam with anybody (in this session or in the parallel session or not in the course) except your instructor until the end of the exam. You may chat with the instructor in the Zoom room during the test. Only when you have successfully uploaded your answer, you are allowed to leave the Zoom meeting.

Please sign below to affirm that you have followed these rules.

Signature: ____________________________

- Distribution formula:

Bernoulli Distribution $X \sim \mathrm{Ber}(p)$	$\mathbb{P}(X=k) = p^k (1-p)^{1-k}$, $k = 0$ or 1
Binomial Distribution $X \sim B(n,p)$	$\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, \cdots, n$
Geometric Distribution $X \sim \mathrm{Geom}(p)$	$\mathbb{P}(X=k) = (1-p)^{k-1} p$, $k = 1,2,3,\cdots$
Poisson Distribution $X \sim \mathrm{Poisson}(\lambda)$	$\mathbb{P}(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k = 0,1,2,\cdots$
Uniform Distribution $X \sim \mathrm{Unif}([a,b])$	$f(x) = \frac{1}{b-a}$, $x \in [a,b]$
Normal Distribution $X \sim \mathcal{N}(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $x \in (-\infty, \infty)$
Exponential Distribution $X \sim \mathrm{Exp}(\lambda)$	$f(t) = \lambda e^{-\lambda t}$, $x \in [0, \infty)$
Gamma Distribution $X \sim \mathrm{Gamma}(\alpha, \lambda)$	$f(t) = \Gamma(\alpha)^{-1} \lambda^\alpha t^{\alpha-1} e^{-\lambda t}$, $x \in [0,\infty)$

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}, \quad \Gamma(r) = \int_0^\infty t^{r-1} e^{-t}\,\mathrm{d}t, \quad \text{and} \quad \Gamma(n) = (n-1)! \quad \forall n \in \mathbb{N}$$

- Change-of-Variable for Densities: Let $Y = g(X)$ where g is either strictly increasing or strictly decreasing, then

$$f_Y(y) = f_X(x) \Big/ \left|\frac{dy}{dx}\right| \quad \text{where } y = g(x).$$

If the function g is many-to-one, then

$$f_Y(y) = \sum_{\{x: g(x) = y\}} f_X(x) \Big/ |g'(x)|.$$

- Bivariate Change-of-Variables for Joint Densities: If $(U,V) = G(X,Y)$ and the Jacobian matrix of G has non-zero determinant, and (X,Y) has joint probability density function $f_{X,Y}$, then the joint probability density function $f_{U,V}$ takes the form

$$\begin{aligned} f_{U,V}(u,v) &= |\mathrm{det} J_{G^{-1}}(u,v)|\, f_{X,Y}\left(G^{-1}(u,v)\right) \\ &= \frac{1}{|\mathrm{det} J_G(x,y)|} f_{X,Y}\left(G^{-1}(u,v)\right) \end{aligned}$$

where $(u,v) = G(x,y)$ and $(x,y) = G^{-1}(u,v)$.

- If random variables X and Y follow a continuous distribution on the plane with probability density function $f_{X,Y}(x,y)$, then $X+Y$ has probability density function

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x, z-x)\,\mathrm{d}x.$$

- Conditional probability density function: If (X,Y) has joint density $f_{X,Y}$ and X, Y has probability density functions f_X, f_Y, respectively, then

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \qquad f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}.$$

- Markov Inequality: If X is a random variable that takes only non-negative values, then, for any value $a > 0$
$$\mathbb{P}\{X \geq a\} \leq \frac{\mathbb{E}[X]}{a}.$$
- Normal Approximation (Central Limit Theorem): Let $S_n = X_1 + \cdots + X_n$ be the sum of n i.i.d. discrete random variables. For large n, the distribution of S_n is approximately normal, with mean $\mathbb{E}(S_n) = n\mu$ and standard deviation $\mathrm{SD}(S_n) = \sigma\sqrt{n}$, where $\mu = \mathbb{E}X_1$ and $\sigma = \mathrm{SD}(X_1)$, i.e., for all $a \leq b$,
$$\mathbb{P}\left(a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \approx \Phi(b) - \Phi(a)$$
where Φ is the standard normal c.d.f. $\Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}}\,\mathrm{d}x$, and the error in this approximation tends to zero as $n \to \infty$. **In other words, S_n/n is asymptotically distributed as $\mathcal{N}(\mu, \sigma^2/n)$ as $n \to \infty$.**

a	-2.58	-1.96	-1.65	-1.28	0	1.28	1.65	1.96	2.58
$\Phi(a)$	0.005	0.025	0.05	0.1	0.5	0.9	0.95	0.975	0.995

.

Problem 1 Suppose A, B are two events such that $P(A) = 0.3, P(B) = 0.4$, and $P(A \cup B) = 0.5$

(1) Find $P(A \mid B)$
(2) Are A and B independent?
(3) Find $P\left(A^c B\right)$.
(4) Let $X = I_A, Y = I_B$. Find the correlation $\rho(X, Y)$.

Problem 2 If U is uniform on $(0, 2\pi)$ and Z, independent of U, is exponential with rate 1, show that X and Y defined by

$$X = \sqrt{2Z}\cos U$$
$$Y = \sqrt{2Z}\sin U$$

are independent standard normal random variables.

Problem 3 Let $X_1, X_2, X_3, \cdots, X_n$ are independent and identically distributed random variables.

(1) Calculate $E[X_1 | X_1 + \cdots + X_n = x]$
(2) If X_1, X_2 follow exponential distribution with parameter λ, find the variance of X_1 when $X_1 + X_2 = x$.

Problem 4 Let U_1 and U_2 be two independent uniform [0,1] random variables. Let

$$X = \min\left(U_1, U_2\right)$$
$$Y = \max\left(U_1, U_2\right)$$

where $\min\left(u_1, u_2\right)$ is the smaller and $\max\left(u_1, u_2\right)$ the larger of two numbers u_1 and u_2. Find:

(1) the probability density function f_X of X;
(2) the joint density function $f_{X,Y}$ of (X, Y)
(3) $P(X \leq 1/2 \mid Y \geq 1/2)$

Problem 5 Let T_1 and T_2 be two independent exponential variables, with rates λ_1 and λ_2. Think of T_i as the lifetime of component i, $i = 1, 2$. Let $T_{\min}$ represent the lifetime of a system which fails whenever the first of the two components fails, so $T_{\min} = \min\left(T_1, T_2\right)$. Let $X_{\min}$ designate which component failed first, so $X_{\min}$ has value 1 if $T_1 < T_2$ and value 2 if $T_2 < T_1$. Show:

(1) that the distribution of $T_{\min}$ is exponential $\left(\lambda_1 + \lambda_2\right)$
(2) that the distribution of $X_{\min}$ is given by the formula $P\left(X_{\min} = i\right) = \frac{\lambda_i}{\lambda_1 + \lambda_2}$ for $i = 1, 2$
(3) that the random variables $T_{\min}$ and $X_{\min}$ are independent;

Problem 6 Let X and Y be independent variables with gamma (r, λ) and gamma (s, λ) distribution, respectively.

(1) Find distribution of $X + Y$
(2) Show that $X/(X + Y)$ is independently of $X + Y$.

Problem 7 (For this problem you may need calculator, but in final, the computation will be easier. So there is no need for any calculator.) A seed manufacturer sells seeds in packets of 50. Assume that each seed germinates with a chance of 99%, independently of all others. The manufacturer promises to replace, at no cost to the buyer, any packet that has 3 or more seeds that do not germinate. What is the chance that the manufacturer has to replace more than 161 of the next 10000 packets sold?