Statistics 251 Name: ____________________________

Section 2, Autumn 2020

Practice Final Exam Solutions CNetID: ____________________________

Time: 2 Hours

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

Please write your **final answer** for each problem in the provided box. Please show your work in the space below the box. If you need additional space for scratchwork, you may use the blank pages stapled to the end of the exam. Please **do not write on the back side of pages**.

You will have 2 hours to complete this exam. You must take this exam between **10:30am** and **12:30pm** Chicago time on December 9, 2020. You must scan and upload the completed exam to Gradescope by **1:00pm Chicago time on December 9, 2020**.

The use of outside material including books, notes, calculators, and electronic devices is not allowed. Due to the coronavirus situation, this exam will be take-home. Please sign below to affirm that you have followed these rules.

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

Signature: ____________________________

Formulas

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

The error function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}}\, dx$$

has values given by the following table.

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 (10 points) Suppose A, B are two events such that $\mathbb{P}(A) = 0.3, \mathbb{P}(B) = 0.4$, and $\mathbb{P}(A \cup B) = 0.5$.

(a) (2.5 points) Find $\mathbb{P}(A \mid B)$.

Answer: $\boxed{0.5}$ We find that

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B)}{\mathbb{P}(B)} = \frac{0.2}{0.4} = 0.5.$$

(b) (2.5 points) Are A and B independent?

Answer: $\boxed{\text{No.}}$ We find that $\mathbb{P}(A \cap B) \neq \mathbb{P}(A) \cdot \mathbb{P}(B)$.

(c) (2.5 points) Find $\mathbb{P}(A^c \cap B)$.

Answer: $\boxed{0.2}$ We find that

$$\mathbb{P}(A^c \cap B) = \mathbb{P}(B) - \mathbb{P}(A \cap B) = 0.4 - 0.2 = 0.2.$$

(d) (2.5 points) Let $X = I_A, Y = I_B$. Find the correlation $\rho(X, Y)$.

Answer: $\boxed{\frac{2\sqrt{14}}{21}}$ We find that

$$\rho(X, Y) = \frac{\mathrm{Cov}(I_A, I_B)}{\sqrt{\mathrm{Var}(I_A)\mathrm{Var}(I_B)}} = \frac{\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)}{\sqrt{(\mathbb{P}(A) - \mathbb{P}(A)^2)(\mathbb{P}(B) - \mathbb{P}(B)^2)}} = \frac{0.08}{\sqrt{0.21 \cdot 0.24}} = \frac{2\sqrt{14}}{21}.$$

Problem 2 (10 points) If U is uniform on $(0, 2\pi)$ and Z is independent of X and exponential with rate 1, show that the random variables X and Y defined by

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

are independent standard normal random variables.

Answer:
The joint density of U and Z is

$$f_{U,Z}(u, z) = \mathbf{1}_{(0,2\pi)}(u) \cdot \frac{1}{2\pi} e^{-z}.$$

The Jacobian of the map from U, Z to X, Y is given by

$$J = \begin{vmatrix} \frac{\cos U}{\sqrt{2Z}} & -\sqrt{2Z}\sin U \\ \frac{\sin U}{\sqrt{2Z}} & \sqrt{2Z}\cos U \end{vmatrix} = 1.$$

Finally, solving for U, Z in terms of X, Y, we find that

$$Z = \frac{1}{2}(X^2 + Y^2) \qquad U = \arctan(Y/X).$$

We conclude that the joint density of X, Y is given by

$$f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} = \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\right)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\right),$$

which is the density of i.i.d. standard normal random variables.

Problem 3 (10 points) Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables.

(a) (5 points) Calculate $\mathbb{E}[X_1 \mid X_1 + \cdots + X_n = x]$

Answer: $\boxed{\frac{x}{n}}$

By symmetry, we have that $\mathbb{E}[X_i \mid X_1 + \cdots + X_n = x]$ is independent of i. We thus find that

$$n\mathbb{E}[X_1 \mid X_1 + \cdots + X_n = x] = \sum_{i=1}^{n} \mathbb{E}[X_i \mid X_1 + \cdots + X_n = x] = x,$$

which implies that $\mathbb{E}[X_1 \mid X_1 + \cdots + X_n = x] = \frac{x}{n}$.

(b) (5 points) If X_1, X_2 are exponential with parameter λ, find the conditional variance $\mathrm{Var}(X_1 \mid X_1 + X_2 = x)$.

Answer: $\boxed{\frac{1}{12}x^2}$

Let $Y = X_1$ and $X = X_1 + X_2$ so that the Jacobian of the map from X_1, X_2 to X, Y is 1. The joint density of X_1, X_2 is

$$f(x_1, x_2) = \lambda^2 e^{-\lambda(x_1+x_2)},$$

which means that the joint density of X, Y is

$$f_{X,Y}(x, y) = \mathbf{1}_{[0,x]}(y)\lambda^2 e^{-\lambda x}.$$

The conditional density of Y given X is therefore

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \mathbf{1}_{[0,x]}(y)\frac{1}{x}$$

We have that

$$\mathrm{Var}(X_1 \mid X_1 + X_2 = x) = \mathbb{E}[Y^2 \mid X = x] - \mathbb{E}[Y \mid X = x]^2 = \int_0^x y^2 \frac{1}{x} dy - \frac{1}{4}x^2 = \frac{1}{12}x^2.$$

Problem 4 (15 points) Let U_1 and U_2 be two independent uniform random variables on $[0,1]$. Define

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

Find

(a) (5 points) the probability density function f_X of X

Answer: $\boxed{\mathbf{1}_{[0,1]}(x)2(1-x)}$

We have

$$F_X(x) = 1 - \mathbb{P}(X \geq x) = 1 - \mathbb{P}(U_1 \geq x)\mathbb{P}(U_2 \geq x) = \mathbf{1}_{[0,1]}(x)(1-(1-x)^2),$$

which means that

$$f_X(x) = F_X'(x) = \mathbf{1}_{[0,1]}(x)2(1-x).$$

(b) (5 points) the joint density function $f_{X,Y}$ of (X,Y)

Answer: $\boxed{2 \cdot \mathbf{1}_{0\leq x\leq y\leq 1}(x,y)}$

Notice that if $x \leq y$, then

$$F_{X,Y}(x,y) = \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(U_1 \leq x, U_2 \leq y) + \mathbb{P}(U_2 \leq x, U_1 \leq y) - \mathbb{P}(U_1 \leq x, U_2 \leq x) = 2xy - x^2.$$

This implies that

$$f_{X,Y}(x,y) = \partial_x \partial_y F_{X,Y}(x,y) = 2\mathbf{1}_{0\leq x\leq y\leq 1}(x,y).$$

(c) (5 points) $\mathbb{P}(X \leq 1/2 \mid Y \geq 1/2)$

Answer: $\boxed{2/3}$

We have that

$$\mathbb{P}(X \leq 1/2 \mid Y \geq 1/2) = \frac{\int_0^{1/2}\int_{1/2}^1 f_{X,Y}(x,y)dydx}{\int_{1/2}^1\int_0^y f_{X,Y}(x,y)dxdy} = \frac{1/2}{3/4} = \frac{2}{3}.$$

Problem 5 (20 points) Let T_1 and T_2 be two independent exponential variables with rates λ_1 and λ_2, respectively. Define $T_{\min} = \min(T_1, T_2)$, and let $X_{\min}$ be a random variable which equals 1 if $T_1 < T_2$ and 2 if $T_2 < T_1$.

(a) (5 points) The distribution of $T_{\min}$ is exponential with some parameter λ. Find λ.

Answer: $\boxed{\lambda_1 + \lambda_2}$
We find that

$$F_{T_{\min}}(t) = 1 - \mathbb{P}(T_{\min} \geq t) = 1 - \mathbb{P}(T_1 \geq t)\mathbb{P}(T_2 \geq t) = 1 - e^{-(\lambda_1+\lambda_2)t},$$

which means that $T_{\min}$ is exponential with parameter $\lambda_1 + \lambda_2$.

(b) (5 points) Find $\mathbb{P}(X_{\min} = 1)$.

Answer: $\boxed{\frac{\lambda_1}{\lambda_1+\lambda_2}}$
We find that

$$\mathbb{P}(X_{\min} = 1) = \mathbb{P}(T_1 < T_2) = \int_0^\infty \int_0^{t_2} \lambda_1 \lambda_2 e^{-\lambda_1 t_1 - \lambda_2 t_2} dt_1 dt_2$$

$$= \lambda_2 \int_0^\infty e^{-\lambda_2 t_2}(1 - e^{-\lambda_1 t_2}) dt_2 = 1 - \frac{\lambda_2}{\lambda_1 + \lambda_2} = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

(c) (10 points) Show that $T_{\min}$ and $X_{\min}$ are independent.

Answer:
We compute that

$$\mathbb{P}(T_{\min} \geq t, X_{\min} = 1) = \mathbb{P}(t \leq T_1 < T_2) = \int_t^\infty \int_t^{t_2} \lambda_1 \lambda_2 e^{-\lambda_1 t_1 - \lambda_2 t_2} dt_1 dt_2$$

$$= \lambda_2 \int_t^\infty e^{-\lambda_2 t_2}(e^{-\lambda_1 t} - e^{-\lambda_1 t_2}) dt_2 = e^{-(\lambda_1+\lambda_2)t} - \frac{\lambda_2 e^{-(\lambda_1+\lambda_2)t}}{\lambda_1 + \lambda_2} = \frac{\lambda_1}{\lambda_1 + \lambda_2} e^{-(\lambda_1+\lambda_2)t}$$

$$= \mathbb{P}(T_{\min} \geq t)\mathbb{P}(X_{\min} = 1),$$

which gives independence.

Problem 6 (20 points) A box contains three coins, of which two are fair and one is unfair, meaning it lands heads with proability 1.

(a) (5 points) If you choose a coin at random and toss it, what is the probability it lands heads?

Answer: $\boxed{\frac{2}{3}}$

The probability of getting heads is

$$\mathbb{P}(H) = \mathbb{P}(H \mid \text{fair})\mathbb{P}(\text{fair}) + \mathbb{P}(H \mid \text{unfair})\mathbb{P}(\text{unfair}) = \frac{1}{2} \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}.$$

(b) (15 points) If you choose a coin at random and get heads when tossing it, what is the probability it is the unfair coin?

Answer: $\boxed{\frac{1}{2}}$

We find that

$$\mathbb{P}(\text{unfair} \mid H) = \frac{\mathbb{P}(H \mid \text{unfair})\mathbb{P}(\text{unfair})}{\mathbb{P}(H)} = \frac{1 \cdot \frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}.$$

Problem 7 (15 points) Consider the sample average $\overline{X}_n = (X_1 + X_2 + \cdots + X_n)/n$ of n i.i.d. random variables $X_1, \ldots, X_n$ which are uniformly distributed on $[0, 1]$.

(a) (5 points) Use Markov's inequality to upper bound $\mathbb{P}(\overline{X}_n \geq 0.99)$.

Answer: $\boxed{\frac{50}{99}}$

Notice that $\mathbb{E}[\overline{X}_n] = \frac{1}{2}$. We find that

$$\mathbb{P}(\overline{X}_n \geq 0.99) \leq \frac{\mathbb{E}[\overline{X}_n]}{0.99} = \frac{50}{99}.$$

(b) (10 points) Use the Central Limit Theorem to find n so that $\mathbb{P}(\overline{X}_n < 0.51)$ is approximately 90%.

Answer: $\boxed{64}$

Notice that $\mathrm{Var}(\overline{X}_n) = \frac{1}{4n}$. The CLT shows that

$$\mathbb{P}(\overline{X}_n - 0.5 < \frac{a}{2n}) \to \Phi(a),$$

so setting $a = 0.02 \cdot n$ we find that

$$\mathbb{P}(\overline{X}_n < 0.51) \to \Phi(0.02 \cdot n).$$

Since $\Phi(1.28) = 0.9$, we need $0.02 \cdot n = 1.28$, meaning that $n = 64$.