Statistics 251 Name: ____________________________

Section 2, Autumn 2020

Final Exam CNetID: ____________________________

Date: December 9, 2020

Time: 10:30am to 12:30pm

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 $X \sim \text{Geom}(p)$	$\mathbb{P}(X=k) = (1-p)^{k-1} p$, $k = 1, 2, 3, \cdots$
Poisson $X \sim \text{Poisson}(\lambda)$	$\mathbb{P}(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}$, $k = 0, 1, 2, \cdots$
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 $X \sim \text{Exp}(\lambda)$	$f(t) = \lambda e^{-\lambda t}$, $x \in [0, \infty)$
Gamma $X \sim \text{Gamma}(\alpha, \lambda)$	$f(t) = \Gamma(\alpha)^{-1} \lambda^\alpha t^{\alpha-1} e^{-\lambda t}$, $x \in [0, \infty)$

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 (20 points) You are designing a spam classifier. In a large body of spam, you find that 10% of spam messages contain a dollar sign "$", while 1% of normal messages contain a dollar sign.

(a) (10 points) If 90% of all messages are spam, what is the probability that a message is spam given that it contains a dollar sign?

Answer:

(b) (10 points) Your email's spam classifier flags 90% of all spam messages and 1% of non-spam messages. If being flagged and having a dollar sign are independent for both spam and non-spam messages, what is the probability that a flagged message without a dollar sign is spam?

Answer:

Problem 2 (15 points) Let X and Y be independent exponential random variables with parameter λ.

(a) (5 points) Compute $\mathbb{E}[(X+Y)^2]$.

Answer:

(b) (10 points) Show that $\frac{X}{X+Y}$ is uniform on $[0,1]$.

Answer:

Problem 3 (20 points) Let X and Y be independent standard Gaussian random variables. Define $Z = 2X + Y$ and $W = 2X - Y$.

(a) (5 points) Compute the correlation $\rho(Z, W)$.

Answer:

(b) (5 points) Find $\mathbb{E}[Z \mid Y = 1]$.

Answer:

(c) (10 points) Find $\mathbb{E}[Z \mid W = 1]$.

Answer:

Problem 4 (15 points) You are trying to estimate an unknown quantity x^*. You have a random variable X which is an estimator for x^* in the sense that $\mathbb{E}[X] = x^*$ as well as another random variable Y with known expectation y. You would like to produce a new estimator for x^* which has lower variance than X. Suppose $\mathrm{Var}(X) = \sigma_X^2$, $\mathrm{Var}(Y) = \sigma_Y^2$, and $\rho(X, Y) = \rho$.

(a) (5 points) Compute the variance of $Z = X + c(Y - y)$

Answer:

(b) (5 points) Find c for which the variance of Z is minimized.

Answer:

(c) (5 points) What is this minimum variance?

Answer:

Problem 5 (15 points) You are running an ad to promote your donation drive. Every time someone watches your ad, they donate with probability p. If they do donate, they donate an amount Y which is a non-negative random variable with mean 1 and variance 1.

(a) (5 points) Let X be the amount donated by a viewer who sees your ad. Find $\mathbb{E}[X]$ and $\text{Var}(X)$.

Answer:

(b) (5 points) Suppose all viewers behave independently, and let X_i be the amount donated by viewer i. Using Markov's inequality, give an upper bound for $\mathbb{P}(X_1 + X_2 > 3)$.

Answer:

(c) (5 points) Using the central limit theorem, estimate $\mathbb{P}(X_1 + \cdots + X_{100} > 100)$.

Answer:

Problem 6 (15 points) Let X be a uniform random variable on $[0, 1]$. Give an example of a continuous random variable Y so that X and Y are uncorrelated but not independent.

Answer: