Qualifying Exam Name: ______________________

Department of Mathematical Sciences

Probability and Statistics

9 AM – 12 PM, 01/04/2024

Time Limit: 3 hours

Notification of this test:

1. You have 3 hours to complete the exam. You are required to show all the work for all the problems. There are two parts in the exam: (1) elementary part and (2) challenging part. Please budget your time wisely for all two parts. Problems 1 - 5 are elementary part and problems 6 - 10 are challenging part including proof questions. The suggested passing grade for the elementary part is 80% (40+), and challenging part is 70% (35+).
2. Partial credits for some questions (or some sub-questions) will be given, so please try your best and include your answer even you do not finish.

Grade Table (for grader use only)

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
10	10	

1. A random variable X has the cumulative distribution function

$$F(x) = \begin{cases} 0, & x < 1 \\ \dfrac{x^2 - 2x + 2}{2}, & 1 \le x < 2, \\ 1, & x \ge 2 \end{cases}$$

   Calculate the variance of X.

2. Claim amounts at an insurance company are independent of one another. In year one, claim amounts are modeled by a normal random variable X with mean 100 and standard deviation 25. In year two, claim amounts are modeled by the random variable $Y = 1.04X + 5$.

   Calculate the probability that a random sample of 25 claim amounts in year two average between 100 and 110.

3. A manufacturer produces computers and releases them in shipments of 100. From a shipment of 100, the probability that exactly three computers are defective is twice the probability that exactly two computers are defective. The events that different computers are defective are mutually independent.

   Calculate the probability that a randomly selected computer is defective.

4. Let X be the percentage score on a college-entrance exam for students who did not participate in an exam-preparation seminar. X is modeled by a uniform distribution on $[a, 100]$.

   Let Y be the percentage score on a college-entrance exam for students who did participate in an exam-preparation seminar. Y is modeled by a uniform distribution on $[1.25a, 100]$.

   It is given that $E(X^2) = 19600/3$. Calculate the 80th percentile of Y.

5. In a large population of patients, 20% have early stage cancer, 10% have advanced stage cancer, and the other 70% do not have cancer. Six patients from this population are randomly selected.

   Calculate the expected number of selected patients with advanced stage cancer, given that at least one of the selected patients has early stage cancer.

6. Prove that if $X_n \to X_0$ in probability, then $X_n \to X_0$ in distribution. Offer a counterexample for the converse.

7. Let A_i be a sequence of events such that $A_i \subset A_{i+1}$, $i = 1, 2, \ldots$. Prove that

$$\lim_{n\to\infty} P(A_n) = P(\cup_{i=1}^{\infty} A_i).$$

8. Let $X_1, \ldots, X_n$ be a random sample of size n from a normal distribution $N(\mu, \sigma^2)$. We would like to test the hypothesis $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. When σ is known, show that the power function of the test with type I error α under true population mean $\mu = \mu_1$ is $\Phi(-z_{\alpha/2} + \frac{|\mu_1 - \mu_0|\sqrt{n}}{\sigma})$, where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution and $\Phi(z_{\alpha/2}) = 1 - \alpha/2$.

9. Suppose that $X_1, \ldots, X_n$ is a random sample from the Rayleigh distribution with pdf

$$f(x|\theta) = \frac{2x}{\theta} e^{-x^2/\theta} I(x > 0)$$

where θ is a positive unknown parameter.

(a) Find the maximum likelihood estimator of θ.

(b) Find a sufficient statistic for θ.

(c) Find the uniformly minimum variance unbiased estimator (UMVUE) for θ.

10. Let $X_1, \ldots, X_n$ be a random sample from the Poisson distribution POI(λ). Find a uniformly most powerful (UMP) test of $H_0 : \lambda \leq \lambda_0$ versus $H_1 : \lambda > \lambda_0$. Remember to specify how to choose the constant in your rejection region.

Qualifying Exam **Name:** ______________________
Department of Mathematical Sciences
Probability and Statistics
8/17/2023
Time Limit: 3 hours

Notification of this test:

1. You have 3 hours to complete the exam. You are required to show all the work for all the problems. There are two parts in the exam: (1) Elementary part and (2) Challenging Part. Please budget your time wisely for all two parts. Problems 1 - 5 are elementary part and problems 6 - 10 are challenging part including proof questions. The suggested passing grade for the elementary part is 80% (40+), and challenging part is 70% (35+).

2. Partial credits for some questions (or some sub-questions) will be given, so please try your best and include your answer even you do not finish.

Grade Table (for teacher use only)

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
10	10	
Total:	100	

---

1. (10 points) An insurance company estimates that 40% of policyholders who have only an auto policy will renew next year and 60% of policyholders who have only a homeowners policy will renew next year. The company estimates that 80% of policyholders who have both an auto policy and a homeowners policy will renew at least one of those policies next year.

   Company records show that 65% of policyholders have an auto policy, 50% of policyholders have a homeowners policy, and 15% of policyholders have both an auto policy and a homeowners policy.

   Using the company's estimates, calculate the percentage of policyholders that will renew at least one policy next year.

2. (10 points) Claim amounts for wind damage to insured homes are mutually independent random variables with common density function

$$f(x) = \begin{cases} \frac{3}{x^4}, & x > 2 \\ 0, & \text{otherwise,} \end{cases}$$

   where x is the amount of a claim in thousands. Suppose 3 such claims will be made. Calculate the expected value of the largest of the three claims.

3. (10 points) Bowl I contains eight red balls and six blue balls. Bowl II is empty. Four balls are selected at random, without replacement, and transferred from bowl I to bowl II. One ball is then selected at random from bowl II. Calculate the conditional probability that two red balls and two blue balls were transferred from bowl I to bowl II, given that the ball selected from bowl II is blue.

4. (10 points) The annual profit of a life insurance company is normally distributed. The probability that the annual profit does not exceed 2000 is 0.7642. The probability that the annual profit does not exceed 3000 is 0.9066. Calculate the probability that the annual profit does not exceed 1000.

5. (10 points) Let X denote the loss amount sustained by an insurance company's policyholder in an auto collision. Let Z denote the portion of X that the insurance company will have to pay. An actuary determines that X and Z are independent with respective density and probability functions

$$f(x) = \begin{cases} \frac{1}{8}e^{-\frac{x}{8}}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

   and

$$P(Z = z) = \begin{cases} 0.45, & z = 1 \\ 0.55, & z = 0. \end{cases}$$

   Calculate the variance of the insurance company's claim payment ZX.

6. (10 points) For any $p > 1$, $q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, assume that $E(|X|^p) < \infty$ and $E(|Y|^q) < \infty$. Prove that

$$E(|XY|) \leq (E(|X|^p))^{1/p} \cdot (E(|Y|^q))^{1/q}.$$

7. (10 points) Let X and Y be independent Poisson random variables with parameter $\lambda > 0$ and probability mass functions given by

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \ldots$$

and

$$g(y) = P(Y = y) = \frac{e^{-\lambda}\lambda^y}{y!}, y = 0, 1, 2, \ldots$$

Prove that $Z = X + Y$ is also Poisson random variable and find its parameter.

8. (10 points) Suppose $X_n \geq 0$ for $n \geq 0$, $X_n \to X_0$ in probability, and also $E(X_n) \to E(X_0)$. Show that $E[|X_n - X_0|] \to 0$.

9. (10 points) Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with probability density function

$$f(x|\theta) = \frac{1}{2\theta}, \ -\theta < x < \theta, \ \theta > 0.$$

Find, if one exists, a best unbiased estimator of θ.

10. (10 points) A random sample is drawn from a Pareto population with probability density function

$$f(x|\theta, \nu) = \frac{\theta\nu^\theta}{x^{\theta+1}} 1_{[\nu,\infty)}(x), \ \theta > 0, \ \nu > 0.$$

   (a) (5 points) Find the maximum likelihood estimators (MLEs) of θ and ν.

   (b) (5 points) Derive the likelihood ratio test (LRT) for $H_0 : \theta = 1, \nu$ unknown v.s $H_1 : \theta \neq 1, \nu$ unknown.

Qualifying Exam **Name:** ____________________

Department of Mathematical Sciences

Probability and Statistics

1/6/2022

Time Limit: 3 hours

Notification of this test:

1. You have 3 hours to complete the exam. You are required to show all the work for all the problems. There are two parts in the exam: (1) Elementary part and (2) Challenging Part. Please budget your time wisely for all two parts. Problems 1 - 5 are elementary part and problems 6 - 10 are challenging part including proof questions. The suggested passing grade for the elementary part is 80% (40+), and challenging part is 70% (35+).

2. Partial credits for some questions (or some sub-questions) will be given, so please try your best and include your answer even you do not finish.

Grade Table (for teacher use only)

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
10	10	
Total:	100	

---

1. (10 points) Losses due to accidents at an amusement park are exponentially distributed. An insurance company offers the park owner two different policies, with different premiums, to insure against losses due to accidents at the park. Policy A has a deductible of 1.44. For a random loss, the probability is 0.640 that under this policy, the insurer will pay some money to the park owner. Policy B has a deductible of d. For a random loss, the probability is 0.512 that under this policy, the insurer will pay some money to the park owner. What is d?

2. (10 points) A device runs until either of two components fails, at which point the device stops running. The joint density function of the lifetimes of the two components, both measured in hours, is

$$f(x, y) = \frac{x + y}{8}\ , \text{ for } 0 < x < 2 \text{ and } 0 < y < 2$$

   Calculate the probability that the device fails during its first hour of operation.

3. (10 points) The number of days an employee is sick each month is modeled by a Poisson distribution with mean 1. The numbers of sick days in different months are mutually independent. Calculate the probability that an employee is sick more than two days in a three-month period.

4. (10 points) An insurance agent offers his clients auto insurance, homeowners insurance and renters insurance. The purchase of homeowners insurance and the purchase of renters insurance are mutually exclusive. The profile of the agent's clients is as follows:

   i) 17% of the clients have none of these three products.
   ii) 64% of the clients have auto insurance.
   iii) Twice as many of the clients have homeowners insurance as have renters insurance.
   iv) 35% of the clients have two of these three products.
   v) 11% of the clients have homeowners insurance, but not auto insurance.

   Calculate the percentage of the agent's clients that have both auto and renters insurance.

5. (10 points) The number of boating accidents a policyholder experiences this year is modeled by a Poisson random variable with variance 0.10. An insurer reimburses only the first accident. Let Y be the number of unreimbursed accidents the policyholder experiences this year and let p be the probability function of Y. Determine $p(y)$.

6. (10 points) Prove the following statements:
   (a) (5 points) If X_n converges to X almost surely, then X_n converges to X in probability.
   (b) (5 points) If X_n converges to X_0 in distribution (weakly) and X_0 is a constant, then X_n converges to X_0 in probability.

7. (10 points) Let $X_1, X_2 \ldots, X_n$ are independent random variables from normal distributions $N(\mu, \sigma^2)$, with $\mu = \mu_0$ known. Find uniformly minimum variance unbiased estimate (UMVUE) of σ^2.

8. (10 points) Let $X_1, X_2 \ldots, X_n$ are independent random variables from $N(\mu, \sigma^2)$ where μ and σ^2 are both unknown. Derive the likelihood ratio test (LRT) for $H_0 : \mu = \mu_0$ v.s $H_a : \mu \neq \mu_0$?

9. (10 points) Let $X_1, X_2, \ldots, X_n$ are independent random variables from $N(0, \sigma^2)$, show that $V = \frac{\sum_{i=1}^{n} X_i^2}{\sigma^2} \sim \chi^2_n$.

10. (10 points) Let $X_1, X_2 \ldots, X_n$ are independent random variables from pdf $f(x) = \theta x^{\theta-1}$ with unknown parameter θ, $0 \leq x \leq 1$, $0 < \theta < \infty$.
    (a) (7 points) Find the maximum likelihood estimator (MLE) of θ, and show that its variance $\to 0$ as $n \to \infty$.
    (b) (3 points) Find the method of moments estimator (MME) of θ.

Qualifying Exam **Name:** ____________________
Department of Mathematical Sciences
Probability and Statistics
8/19/2021
Time Limit: 3 hours

Notification of this test:

1. You have 3 hours to complete the exam. You are required to show all the work for all the problems. There are two parts in the exam: (1) Elementary part and (2) Challenging Part. Please budget your time wisely for all two parts. Problems 1 - 5 are elementary part and problems 6 - 10 are challenging part including proof questions. The suggested passing grade for the elementary part is 80%, and challenging part is 70%.

2. Partial credits for some questions (or some sub-questions) will be given, so please try your best and include your answer even you do not finish.

Grade Table (for teacher use only)

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
10	10	
Total:	100	

---

1. (10 points) Let X and Y be identically distributed independent random variables such that the moment generating function of $X+Y$ is

$$M(t) = 0.09e^{-2t} + 0.24e^{-t} + 0.34 + 0.24e^{t} + 0.09e^{2t}\ ,\ -\infty < t < \infty$$

   Calculate $P(X \leq 0)$

2. (10 points) An insurance company's monthly claims are modeled by a continuous, positive random variable X, whose probability density function is proportional to $(1+x)^{-4}$ for $0 < x < \infty$. Calculate the company's expected monthly claims.

3. (10 points) Every day, the 30 employees at an auto plant each have probability 0.03 of having one accident and zero probability of having more than one accident. Given there was an accident, the probability of it being major is 0.01. All other accidents are minor. The numbers and severities of employee accidents are mutually independent.

   Let X and Y represent the numbers of major accidents and minor accidents, respectively, occurring in the plant today. Determine the joint moment generating function $M_{X,Y}(s,t)$.

4. (10 points) A government employee's yearly dental expense follows a uniform distribution on the interval from 200 to 1200. The government's primary dental plan reimburses an employee for up to 400 of dental expense incurred in a year, while a supplemental plan pays up to 500 of any remaining dental expense.

   Let Y represent the yearly benefit paid by the supplemental plan to a government employee. Calculate $Var(Y)$.

5. (10 points) An insurance company pays hospital claims. The number of claims that include emergency room or operating room charges is 85% of the total number of claims. The number of claims that do not include emergency room charges is 25% of the total number of claims. The occurrence of emergency room charges is independent of the occurrence of operating room charges on hospital claims. Calculate the probability that a claim submitted to the insurance company includes operating room charges.

6. (10 points) Assume that the $Y_1, Y_2, \ldots, Y_n$ are independent random variables and the distribution of Y_i is Poisson distribution with mean λx_i. What is the regression of Y on X (2 points)? Find the maximum likelihood estimator (MLE) and least square estimator (LSE) of λ (8 points).

7. (10 points) (Boole's inequality) Let A_i be a sequence of events. Show that

$$P\big(\cup_{i=1}^{\infty} A_i\big) \leq \sum_{i=1}^{\infty} P(A_i)$$

8. (10 points) (Borel Cantelli lemma). Let A_i be a sequence of event such that $\sum_{i=1}^{\infty} P(A_i) < \infty$, then

$$P(\cap_{i=1}^{\infty} \cup_{k=i}^{\infty} A_k) = 0$$

9. (10 points) Consider the following situation

x	0	1	2	3	4
$P(x\|\theta_1)$	0.00	0.05	0.05	0.80	0.10
$P(x\|\theta_2)$	0.05	0.05	0.80	0.10	0.00
$P(x\|\theta_3)$	0.90	0.08	0.02	0.00	0.00

A "single" observation is observed. Consider testing $H_0 : \theta = \theta_3$ v.s. $H_1 : \theta \in \{\theta_1, \theta_2\}$.

(a) (5 points) Define the likelihood ratio test and find the testing statistic.

(b) (3 points) Find a rejection region of likelihood ratio test for type I error $\alpha = 0.02$.

(c) (2 points) Find a power function of your test.

10. (10 points) Prove the statement: if $Z_n \to Z$ in probability them $Z_n \to Z$ in distribution (weakly).

Part 1: Sample Problems for the Elementary Section of Qualifying Exam in Probability and Statistics

https://www.soa.org/Files/Edu/edu-exam-p-sample-quest.pdf

Part 2: Sample Problems for the Advanced Section of Qualifying Exam in Probability and Statistics

Probability

1. The *Pareto* distribution, with parameters α and β, has pdf

$$f(x) = \frac{\beta\alpha^\beta}{x^{\beta+1}},\ \alpha < x < \infty,\ \alpha > 0,\ \beta > 0.$$

(a) Verify that $f(x)$ is a pdf.
(b) Derive the mean and variance of this distribution.
(c) Prove that the variance does not exist if $\beta \le 2$.

2. Let $U_i,\ i = 1, 2, \ldots,$ be independent uniform(0,1) random variables, and let X have distribution

$$P(X = x) = \frac{c}{x!},\ x = 1, 2, 3, \ldots,$$

where $c = 1/(e-1)$. Find the distribution of $Z = min\{U_1, \ldots, U_X\}$.

3. A point is generated at random in the plane according to the following polar scheme. A radius R is chosen, where the distribution of R^2 is χ^2 with 2 degrees of freedom. Independently, an angle θ is chosen, where $\theta \sim$ uniform$(0, 2\pi)$. Find the joint distribution of $X = R\cos\theta$ and $Y = R\sin\theta$.

4. Let X and Y be iid $N(0,1)$ random variables, and define $Z = \min(X, Y)$. Prove that $Z^2 \sim \chi^2_1$.

5. Suppose that $\mathcal{B}$ is a σ-field of subsets of Ω and suppose that $P:\ \mathcal{B} \to [0,1]$ is a set function satisfying:
(a) P is finitely additive on $\mathcal{B}$;
(b) $0 \le P(A) \le 1$ for all $A \in \mathcal{B}$ and $P(\Omega) = 1$;
(c) If $A_i \in \mathcal{B}$ are disjoint and $\cup_{i=1}^\infty A_i = \Omega$, then $\sum_{i=1}^\infty P(A_i) = 1$.
Show that P is a probability measure on $\mathcal{B}$ in Ω.

6. Suppose that $\{X_n\}_{n=1}^\infty$ is a sequence of i.i.d. random variables and c_n is an increasing sequence of positive real numbers such that for all $\alpha > 1$, we have

$$\sum_{n=1}^{\infty} P[X_n > \alpha^{-1} c_n] = \infty$$

and

$$\sum_{n=1}^{\infty} P[X_n > \alpha c_n] < \infty.$$

Prove that

$$P\left[\limsup_{n\to\infty} \frac{X_n}{c_n} = 1\right] = 1.$$

7. Suppose for $n \geq 1$ that $X_n \in L_1$ are random variables such that $sup_{n\geq 1} E(X_n) < \infty$. Show that if $X_n \uparrow X$, then $X \in L_1$ and $E(X_n) \to E(X)$.

8. Let X be a random variable with distribution function $F(x)$.
(a) Show that

$$\int_{\mathbb{R}} (F(x+a) - F(x))dx = a.$$

(b) If F is continuous, then $E[F(X)] = \frac{1}{2}$.

9. (a) Suppose that $X_n \xrightarrow{P} X$ and g is a continuous function. Prove that $g(X_n) \xrightarrow{P} g(X)$.
(b) If $X_n \xrightarrow{P} 0$, then for any $r > 0$,

$$\frac{|X_n|^r}{1 + |X_n|^r} \xrightarrow{P} 0$$

and

$$E[\frac{|X_n|^r}{1 + |X_n|^r}] \to 0.$$

10. Suppose that $\{X_n, n \geq 1\}$ are independent non-negative random variables satisfying $E(X_n) = \mu_n$, $\mathrm{Var}(X_n) = \sigma_n^2$. Define for $n \geq 1$, $S_n = \sum_{i=1}^n X_i$ and suppose that $\sum_{n=1}^{\infty} \mu_n = \infty$ and $\sigma_n^2 \leq c\mu_n$ for some $c > 0$ and all n. Show

$$\frac{S_n}{E(S_n)} \xrightarrow{P} 1.$$

11. (a) If $X_n \to X$ and $Y_n \to Y$ in probability, then $X_n + Y_n \to X + Y$ in probability.
(b) Let $\{X_i\}$ be iid, $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$. Set $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$. Show that

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \to \sigma^2$$

in probability.

12. Suppose that the sequence $\{X_n\}$ is *fundamental in probability* in the sense that for ε positive there exists an N_ε such that $P[|X_n - X_m| > \varepsilon] < \varepsilon$ for $m,\ n > N_\varepsilon$.
(a) Prove that there is a subsequence $\{X_{n_k}\}$ and a random variable X such that $\lim_k X_{n_k} = X$ with probability 1 (i.e. almost surely).
(b) Show that $f(X_n) \to f(X)$ in probability if f is a continuous function.

Statistics

1. Suppose that $X = (X_1, \cdots, X_n)$ is a sample from the probability distribution P_θ with density

$$f(x|\theta) = \begin{cases} \theta(1+x)^{-(1+\theta)}, & \text{if} \quad x > 0 \\ 0, & \text{otherwise} \end{cases}$$

for some $\theta > 0$.
(a) Is $\{f(x|\theta), \theta > 0\}$ a one-parameter exponential family? (explain your answer).
(b) Find a sufficient statistic $T(X)$ for $\theta > 0$.

2. Suppose that $X_1, \cdots, X_n$ is a sample from a population with density

$$p(x, \theta) = \theta a x^{a-1} \exp(-\theta x^a),\ x > 0,\ \theta > 0,\ a > 0.$$

(a) Find a sufficient statistic for θ with a fixed.
(b) Is the sufficient statistic in part (a) minimally sufficient? Give reasons for your answer.

3. Let $X_1, \cdots, X_n$ be a random sample from a gamma(α, β) population.
(a) Find a two-dimensional sufficient statistic for (α, β).
(b) Is the sufficient statistic in part (a) minimally sufficient? Explain your answer.
(c) Find the moment estimator of (α, β).
(d) Let α be known. Find the best unbiased estimator of β.

4. Let $X_1, \ldots, X_n$ be iid Bernoulli random variables with parameter θ (probability of a success for each Bernoulli trial), $0 < \theta < 1$. Show that $T(X) = \sum_{i=1}^n X_i$ is minimally sufficient.

5. Suppose that the random variables $Y_1, \cdots, Y_n$ satisfy

$$Y_i = \beta x_i + \varepsilon_i, i = 1, \cdots, n,$$

where $x_1, \cdots, x_n$ are fixed constants, and $\varepsilon_1, \cdots, \varepsilon_n$ are iid $N(0, \sigma^2)$, σ^2 unknown.
(a) Find a two-dimensional sufficient statistics for (β, σ^2).
(b) Find the MLE of β, and show that it is an unbiased estimator of β.
(c) Show that $[\sum(Y_i/x_i)]/n$ is also an unbiased estimator of β.

6. Let $X_1, \cdots, X_n$ be iid $N(\theta, \theta^2)$, $\theta > 0$. For this model both $\bar{X}$ and cS are unbiased estimators of θ, where

$$c = \frac{\sqrt{n-1}\Gamma((n-1)/2)}{\sqrt{2}\Gamma(n/2)}.$$

(a) Prove that for any number a the estimator $a\bar{X} + (1-a)(cS)$ is an unbiased estimator of θ.
(b) Find the value of a that produces the estimator with minimum variance.
(c) Show that $(\bar{X}, S^2)$ is a sufficient statistic for θ but it is not a complete sufficient statistic.

7. Let $X_1, \cdots, X_n$ be i.i.d. with pdf

$$f(x|\theta) = \frac{2x}{\theta} \exp\{-\frac{x^2}{\theta}\}, \ x > 0, \ \theta > 0.$$

(a) Find the Fisher information

$$I(\theta) = E_\theta \left[\left(\frac{\partial}{\partial \theta} \log f(\mathbf{X}|\theta) \right)^2 \right],$$

where $f(\mathbf{X}|\theta)$ is the joint pdf of $\mathbf{X} = (X_1, \ldots, X_n)$.
(b) Show that $\frac{1}{n}\sum_{i=1}^n X_i^2$ is an UMVUE of θ.

8. Let $X_1, \cdots, X_n$ be a random sample from a $n(\theta, \sigma^2)$ population, σ^2 known. Consider estimating θ using squared error loss. Let $\pi(\theta)$ be a $n(\mu, \tau^2)$ prior distribution on θ and let δ^π be the Bayes estimator of θ. Verify the following formulas for the risk

function, Bayes estimator and Bayes risk.
(a) For any consatnts a and b, the estimator $\delta(X) = a\bar{X} + b$ has risk function

$$R(\theta, \delta) = a^2 \frac{\sigma^2}{n} + (b - (1-a)\theta)^2.$$

(b) Show that the Bayes estimator of θ is given by

$$\delta^\pi(X) = E(\theta|\bar{X}) = \frac{\tau^2}{\tau^2 + \sigma^2/n}\bar{X} + \frac{\sigma^2/n}{\tau^2 + \sigma^2/n}\mu.$$

(c)Let $\eta = \sigma^2/(n\tau^2 + \sigma^2)$. The risk function for the Bayes estimator is

$$R(\theta, \delta^\pi) = (1-\eta)^2 \frac{\sigma^2}{n} + \eta^2(\theta - \mu)^2.$$

(d) The Bayes risk for the Bayes estimator is

$$B(\pi, \delta^\pi) = \tau^2 \eta.$$

9. Suppose that $X = (X_1, \cdots, X_n)$ is a sample from normal distribution $N(\mu, \sigma^2)$ with $\mu = \mu_0$ known.
(a) Show that $\hat{\sigma_0}^2 = n^{-1} \sum_{i=1}^n (X_i - \mu_0)^2$ is a uniformly minimum variance unbiased estimate (UMVUE) of σ^2.
(b) Show that $\hat{\sigma_0}^2$ converges to σ^2 in probability as $n \to \infty$.
(c) If μ_0 is not known and the true distribution of X_i is $N(\mu, \sigma^2)$, $\mu \neq \mu_0$, find the bias of $\hat{\sigma_0}^2$.

10. Let $X_1, \cdots, X_n$ be i.i.d. as $X = (Z, Y)^T$, where $Y = Z + \sqrt{\lambda}W$, $\lambda > 0$, Z and W are independent $N(0, 1)$.
(a)Find the conditional density of Y given $Z = z$.
(b)Find the best predictor of Y given Z and calculate its mean squared prediction error (MSPE).
(c)Find the maximum likelihood estimate (MLE) of λ.
(d)Find the mean and variance of the MLE.

11. Let $X_1, \cdots, X_n$ be a sample from distribution with density

$$p(x, \theta) = \theta x^{\theta - 1} 1\{x \in (0, 1)\}, \ \theta > 0.$$

(a) Find the most powerful (MP) test for testing $H : \theta = 1$ versus $K : \theta = 2$ with $\alpha = 0.10$ when $n = 1$.

(b) Find the MP test for testing $H : \theta = 1$ versus $K : \theta = 2$ with $\alpha = 0.05$ when $n \geq 2$.

12. Let $X_1, \cdots, X_n$ be a random sample from a $N(\mu_1, \sigma_1^2)$, and let $Y_1, \cdots, Y_m$ be an independent random sample from a $N(\mu_2, \sigma_2^2)$. We would like to test

$$H : \mu_1 = \mu_2 \quad \text{versus} \quad K : \mu_1 \neq \mu_2$$

with the assumption that $\sigma_1^2 = \sigma_2^2$.
(a) Derive the likelihood ratio test (LRT) for these hypotheses. Show that the LRT can be based on the statistic

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{S_p^2 \left(\frac{1}{n} + \frac{1}{m}\right)}},$$

where

$$S_p^2 = \frac{1}{n+m-2} \left(\sum_{i=1}^{n} (X_i - \bar{X})^2 + \sum_{j=1}^{m} (Y_j - \bar{Y})^2 \right).$$

(b)Show that, under H, T has a t_{n+m-2} distribution.

Part 3: Required Proofs for Probability and Statistics Qualifying Exam

In what follows X_i's are always i.i.d. real random variables (unless otherwise specified).

You are allowed to use some well known theorems (like Lebesgue Dominant Convergence Theorem or Chebyshev inequality), but you must state them and explain how and where do you use them.

Warning: If X and Y have the same moment generating function it does not mean that their distributions are the same.

1. Prove that

$$\text{if } X_n \to X_0 \text{ in probability, then } X_n \to X_0 \text{ in distribution.}$$

Offer a connterexample for the converse.

2. Prove that

$$\text{if } E|X_n - X_0| \to 0. \quad \text{then } X_n \to X_0 \text{ in probability.}$$

Offer a connterexample for the converse.

3. We define $d_{BL}(X_n, X_0) = Sup_{H \in BL}|EH(X_n) - EH(X_0)|$, where BL is a set of all real functions that are Lipshitz and bounded by 1. Prove that

$$\text{if } d_{BL}(X_n, X_0) \to 0, \quad \text{then } P(X_n \leq t) \to P(X_0 \leq t)$$

for every t for which function $F(t) = P(X_0 \leq t)$ is continuous.

4. Prove that

$$\text{if } X_n \to X_0 \text{ in probability and } Y_n \to Y_0 \text{ in distribution,}$$

then

$$X_n + Y_n \to X_0 + Y_0 \text{ in distribution.}$$

5. Prove that if $EX_i^2 < \infty$, then

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to E(X_1) \text{ in probability.}$$

6. (Count as two) Prove that if $E(|X_i|)$ exists, then

$$n^{-1}\sum_{i=1}^{n} X_i \to EX_1 \text{ in probability.}$$

7. Prove that if $EX_i^4 < \infty$, then

$$n^{-1}\sum_{i=1}^{n} X_i \to EX_1 \text{ a.s.}$$

Hint: Work with: $P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} |n^{-1}\sum_{i=1}^{n} X_i - EX_1| > \varepsilon)$.

8. (Count as two) Prove that if $E|X_i|^3 < \infty$, then

$$n^{-1/2}\sum_{i=1}^{n}(X_i - EX_1) \to Z \text{ in distribution,}$$

where Z is a centered normal random variable with $E(Z^2) = Var(X_i) = \sigma^2$.

9. Prove: For any $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$

$$E|XY| \le (E|X|^p)^{1/p} (E|X|^q)^{1/q}.$$

10. Prove that if

$$X_n \to X_0 \text{ in probability and } |X_i| \le M < \infty,$$

then

$$E|X_n - X_0| \to 0.$$

11. (Count as two) Let $F_n(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \le t\}}$ and $F(t) = P(X_i \le t)$ be a continuous function. Then

$$\sup_t |F_n(t) - F(t)| \to 0 \text{ in probability.}$$

12. Let X and Y be independent Poisson random variables with their parameters equal λ. Prove that $Z = X + Y$ is also Poisson and find its parameter.

13. Let X and Y be independent normal random variables with $E(X) = \mu_1, E(Y) = \mu_2, Var(X) = \sigma_1^2, Var(Y) = \sigma_2^2$. Show that $Z = X + Y$ is also normal and find $E(Z)$ and $Var(Z)$.

14. Let X_n converge in distribution to X_0 and let $f : R \to R$ be a continuous function. Show that $f(X_n)$ converges in distribution to $f(X_0)$.

15. Using only the Axioms of probability and set theory, prove that
a)
$$A \subset B \Rightarrow P(A) \leq P(B).$$

b)
$$P(X + Y > \varepsilon) \leq P(X > \varepsilon/2) + P(Y > \varepsilon/2).$$

c) If A and B are independent events, then A^c and B^c are independent as well.
d) If A and B are mutually exclusive and $P(A) + P(B) > 0$, show that
$$P(A|A \cup B) = \frac{P(A)}{P(A) + P(B)}.$$

16. Let A_i be a sequence of events. Show that
$$P(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

17. Let A_i be a sequence of events such that $A_i \subset A_{i+1}$, $i = 1, 2, ...$ Prove that
$$\lim_{n\to\infty} P(A_n) = P(\cup_{i=1}^{\infty} A_i).$$

18. Formal definition of weak convergence states that $X_n \to X_0$weakly if for every continuous and bounded function f :$R \to R$, $Ef(X_n) \to Ef(X_0)$. Show that:
$$X_n \to X_0 \text{ weakly } \Rightarrow P(X_n \leq t) \to P(X \leq t)$$
for every t for which the function $F(t) = P(X \leq t)$ is continuous.

19. (Borel-Cantelli lemma). Let A_i be a sequence of events such that $\sum_{i=1}^{\infty} P(A_i) < \infty$, then
$$P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) = 0.$$

20. Consider the linear regression model $Y = X\beta + e$, where Y is an $n \times 1$ vector of the observations, X is the $n \times p$ design matrix of the levels of the regression variables, β is an $p \times 1$ vector of the regression coefficients, and e is an $n \times 1$ vector of random errors. Prove that the least squares estimator for β is $\hat{\beta} = (X'X)^{-1}X'Y$.

21. Prove that if X follows a F distribution $F(n_1, n_2)$, then X^{-1} follows $F(n_2, n_1)$.

22. Let $X_1, \cdots, X_n$ be a random sample of size n from a normal distribution $N(\mu, \sigma^2)$. We would like to test the hypothesis $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. When σ is known, show that the power function of the test with type I error α under true population mean $\mu = \mu_1$ is $\Phi(-z_{\alpha/2} + \frac{|\mu_1 - \mu_0|\sqrt{n}}{\sigma})$, where $\Phi(.)$ is the cumulative distribution function of a standard normal distribution and $\Phi(z_{\alpha/2}) = 1 - \alpha/2$.

23. Let $X_1, \cdots, X_n$ be a random sample of size n from a normal distribution $N(\mu, \sigma^2)$. Prove that (a) the sample mean $\bar{X}$ and the sample variance S^2 are independent; (b) $\frac{(n-1)S^2}{\sigma^2}$ follows a Chi-squared distribution $\chi^2(n-1)$.

Qualifying Exam **Name:** ____________________

Department of Mathematical Sciences

Probability and Statistics

8/20/2020

Time Limit: 3 hours

Notification of this test:

1. You have 3 hours to complete the exam. You are required to show all the work for all the problems. There are two parts in the exam: (1) Elementary part and (2) Challenging Part. Please budget your time wisely for all two parts. Problems 1 - 5 are elementary part and problems 6 - 10 are challenging part including proof questions. The suggested passing grade for the elementary part is 80%, and challenging part is 70%.

2. Partial credits for some questions (or some sub-questions) will be given, so please try your best and include your answer even you do not finish.

3. The Normal distribution function is provided in Appendix I.

Grade Table (for teacher use only)

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
10	10	
Total:	100	

---

1. (10 points) In a group of 15 health insurance policyholders diagnosed with cancer, each policyholder has probability 0.90 of receiving radiation and probability 0.40 of receiving chemotherapy. Radiation and chemotherapy treatments are independent events for each policyholder, and the treatments of different policyholders are mutually independent.

   The policyholders in this group all have the same health insurance that pays 2 for radiation treatment and 3 for chemotherapy treatment.

   Calculate the variance of the total amount the insurance company pays for the radiation and chemotherapy treatments for these 15 policyholders.

2. (10 points) An actuary studied the likelihood that different types of drivers would be involved in at least one collision during any one-year period. The results of the study are:

Type of driver	Percentage of all drivers	Probability of at least one collision
Teen	8%	0.15
Young adult	16%	0.08
Midlife	45%	0.04
Senior	31%	0.05
Total	100%	

   Given that a driver has been involved in at least one collision in the past year, calculate the probability that the driver is a young adult driver.

3. (10 points) An insurance company sells two types of auto insurance policies: Basic and Deluxe. The time until the next Basic Policy claim is an exponential random variable with mean two days. The time until the next Deluxe Policy claim is an independent exponential random variable with mean three days.

   Calculate the probability that the next claim will be a Deluxe Policy claim.

4. (10 points) The joint probability density function of X and Y is given by

$$f(x,y) = \begin{cases} \dfrac{x+y}{8} & 0 < x < 2, 0 < y < 2 \\ 0 & \text{otherwise.} \end{cases}$$

   Calculate the variance of $(X+Y)/2$.

5. (10 points) The annual profit of a life insurance company is normally distributed.

   The probability that the annual profit does not exceed 2000 is 0.7642. The probability that the annual profit does not exceed 3000 is 0.9066.

   Calculate the probability that the annual profit does not exceed 1000.

6. (10 points) State (2pts) and prove (8pts) the Chebyshev inequality. (*hint:* prove the moment inequality first)

7. (10 points) The *Pareto* distribution, with parameters α and β, has p.d.f

$$f(x) = \frac{\beta\alpha^\beta}{x^{\beta+1}} \ , \alpha < x < \infty, \alpha > 0, \beta > 0$$

   (a) (3 points) Verify that $f(x)$ is a probability density function (pdf)

   (b) (4 points) Derive the mean and variance of this distribution.

   (c) (3 points) Prove that the variance does not exist if $\beta \le 2$.

8. (10 points) Prove if Z_n converges to z_0 in probability, and function $g(\cdot)$ is continuous at z_0, then $g(Z_n)$ converges to $g(z_0)$ in probability.

9. (10 points) Let $X_1, X_2 \ldots, X_n$ be independent exponential random variables with expectation $i\theta$, so

$$f(x_i|i\theta) = \frac{1}{i\theta} \exp\big(-\frac{x_i}{i\theta}\big)$$

   for $i = 1, 2, \ldots, n$.

   (a) (2 points) Find the maximum likelihood estimator (MLE) of θ.

   (b) (2 points) From the MLE $\hat{\theta}$ in (a), show that $2n\hat{\theta}/\theta$ has a χ^2-distribution with $2n$ degrees of freedom.

   (c) (2 points) Find a complete sufficient statistic for θ.

   (d) (2 points) Find uniformly minimum-variance unbiased estimator (UMVUE) of θ.

   (e) (2 points) Find the Cramer-Rao lower bound for θ.

10. (10 points) Let X and Y be independent random variables with pdfs $f(x) = e^{-x}$, $x > 0; f(y) = e^{-y}$, $y > 0$.
   Find the pdf of $Z = \ln\big(\frac{Y}{X}\big)$.

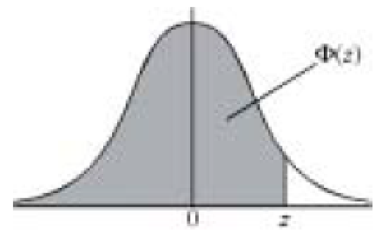
Appendix I: Standard Normal distribution.

THE NORMAL DISTRIBUTION FUNCTION

If Z has a normal distribution with mean 0 and variance 1 then, for each value of z, the table gives the value of $\Phi(z)$, where

$$\Phi(z) = P(Z \le z).$$

For negative values of z use $\Phi(-z) = 1 - \Phi(z)$.



z	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
															ADD				
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	4	8	12	16	20	24	28	32	36
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	4	8	12	16	20	24	28	32	36
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	4	8	12	15	19	23	27	31	35
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	4	7	11	15	19	22	26	30	34
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	4	7	11	14	18	22	25	29	32
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	3	7	10	14	17	20	24	27	31
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549	3	7	10	13	16	19	23	26	29
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	3	6	9	12	15	18	21	24	27
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	3	5	8	11	14	16	19	22	25
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389	3	5	8	10	13	15	18	20	23
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621	2	5	7	9	12	14	16	19	21
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830	2	4	6	8	10	12	14	16	18
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015	2	4	6	7	9	11	13	15	17
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177	2	3	5	6	8	10	11	13	14
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319	1	3	4	6	7	8	10	11	13
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441	1	2	4	5	6	7	8	10	11
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545	1	2	3	4	5	6	7	8	9
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633	1	2	3	4	4	5	6	7	8
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706	1	1	2	3	4	4	5	6	6
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767	1	1	2	2	3	4	4	5	5
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817	0	1	1	2	2	3	3	4	4
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857	0	1	1	2	2	2	3	3	4
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890	0	1	1	1	2	2	2	3	3
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916	0	1	1	1	1	2	2	2	2
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936	0	0	1	1	1	1	1	2	2
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952	0	0	0	1	1	1	1	1	1
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964	0	0	0	0	1	1	1	1	1
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974	0	0	0	0	0	1	1	1	1
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981	0	0	0	0	0	0	0	1	1
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986	0	0	0	0	0	0	0	0	0

Qualifying Exam on Probability and Statistics
January 15, 2020

Name:

Instruction: There are ten problems at two levels: 5 problems at elementary level and 5 proof problems at graduate level. Therefore, please make sure to budget your time to complete problems at both levels. Show your detailed steps in order to receive credits. The suggested passing grade for the elementary part is 80%, and the proof part is 70%. You have 3 hours to complete the exam.

Level 1: Elementary Problems

1. A hospital receives $\frac{1}{5}$ of its flu vaccine shipments from Company X and the remainder of its shipments from other companies. Each shipment contains a very large number of vaccine vials. For Company X's shipments, 10% of the vials are ineffective. For every other company, 2% of the vials are ineffective. The hospital tests 30 randomly selected vials from a shipment and finds that one vial is ineffective. Calculate the probability that this shipment came from Company X.

2. An urn contains 10 balls: 4 red and 6 blue. A second urn contains 16 red balls and an unknown number of blue balls. A single ball is drawn from each urn. The probability that both balls are the same color is 0.44. Determine the number of blue balls in the second urn.

3. An auto insurance company insures an automobile worth $15,000$ for one year under a policy with a $1,000$ deductible. During the policy year there is a 0.04 chance of partial damage to the car and a 0.02 chance of a total loss of the car. If there is partial damage to the car, the amount X of damage (in thousands) follows a distribution with density function

$$f(x) = \begin{cases} 0.5003e^{-\frac{x}{2}}, & \text{if } 0 < x < 15 \\ 0, & \text{otherwise.} \end{cases}$$

Calculate the expected claim payment.

4. The joint probability density function of X and Y is

$$f(x, y) = \begin{cases} 10x^2y, & \text{if } 0 < y < x < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Calculate $P(Y > \frac{1}{3} | X > \frac{2}{3})$.

5. Claim amounts at an insurance company are independent of one another. In year one, claim amounts are modeled by a normal random variable X with mean 100 and standard deviation 25. In year two, claim amounts are modeled by the random variable $Y = 1.04X + 5$. Calculate the probability that a random sample of 25 claim amounts in year two average between 100 and 110.

Level 2: Proof Problems

6. Let X and Y be independent normal random variables with means μ_1 and μ_2 and standard deviations σ_1 and σ_2, respectively. Show that $W = X + Y$ is also normally distributed with mean $\mu = \mu_1 + \mu_2$ and standard deviation $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$.

7. Prove that if $X_n \to X_0$ in probability, then $X_n \to X_0$ in distribution.

8. Suppose $X_n \to X$ in probability and f is a continuous function. Show that $f(X_n) \to f(X)$ in probability.

9. Let $\{X_1, \ldots, X_n\}$ be a random sample from a population with probability density function

$$f(x|\theta) = e^{-(x-\theta)}, \theta < x < \infty, -\infty < \theta < \infty.$$

(a). Find a minimal sufficient statistic for θ.

(b). Is the minimal sufficient statistic found in part (a) a complete sufficient statistic for θ? Justify your answer.

10. Suppose that we have two independent random samples: $X_1, \ldots, X_n$ are exponential(θ) with probability density function $f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, 0 \leq x < \infty,\ \theta > 0$, and $Y_1, \ldots, Y_m$ are exponential(μ) with probability density function $g(y) = \frac{1}{\mu} e^{-\frac{y}{\mu}}, 0 \leq y < \infty,\ \mu > 0$.

(a). Find the likelihood ratio test of $H_0 : \theta = \mu$ versus $H_1 : \theta \neq \mu$.

(b). Show that the test in part (a) can be based on the statistic

$$T = \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j}.$$

Qualifying Exam on Probability and Statistics
August 21, 2019

Name:

Instruction: There are ten problems at two levels: 5 problems at elementary level and 5 proof problems at graduate level. Therefore, please make sure to budget your time to complete problems at both levels. Show your detailed steps in order to receive credits. The suggested passing grade for the elementary part is 80%, and the proof part is 70%. You have 3 hours to complete the exam.

Level 1: Elementary Problems

1. An insurance company estimates that 40% of policyholders who have only an auto policy will renew next year and 60% of policyholders who have only a homeowners policy will renew next year. The company estimates that 80% of policyholders who have both an auto policy and a homeowners policy will renew at least one of those policies next year. Company records show that 65% of policyholders have an auto policy, 50% of policyholders have a homeowners policy, and 15% of policyholders have both an auto policy and a homeowners policy. Using the company's estimates, calculate the percentage of policyholders that will renew at least one policy next year.

2. The loss due to a fire in a commercial building is modeled by a random variable X with probability density function

$$f(x) = \begin{cases} 0.005(20 - x), & \text{if } 0 < x < 20 \\ 0, & \text{otherwise.} \end{cases}$$

Given that a fire loss exceeds 8, calculate the probability that it exceeds 16.

3. An insurance policy on an electrical device pays a benefit of 4000 if the device fails during the first year. The amount of the benefit decreases by 1000 each successive year until it reaches 0. If the device has not failed by the beginning of any given year, the probability of failure during that year is 0.4. Calculate the expected benefit under this policy.

4. An insurance company insures a large number of drivers. Let X be the random variable representing the company's losses under collision insurance, and let Y represent the company's losses under liability insurance. X and Y have joint probability density function

$$f(x,y) = \begin{cases} \frac{2x+2-y}{4}, & 0 < x < 1 \quad \text{and } 0 < y < 2 \\ 0, & \text{otherwise.} \end{cases}$$

Calculate the probability that the total company loss is at least 1.

5. An investor invests 100 dollars in a stock. Each month, the investment has probability 0.5 of increasing by 1.10 dollars and probability 0.5 of decreasing by 0.90 dollars. The changes in price in different months are mutually independent. Calculate the approximate probability that the investment has a value greater than 91 dollars at the end of month 100.

Level 2: Proof Problems

6. Let $X_n \to X$ in distribution and $Y_n \to c$ (a constant) in probability. Show that $X_n + Y_n \to X + c$ in distribution.

7. Let $\{X_k\}_{k=1}^{\infty}$ be i.i.d. with $E|X_1| < \infty$. Show that $\frac{1}{n}\sum_{k=1}^{n} X_k \to E(X_1)$ in probability as $n \to \infty$.

8. Consider the linear regression model $Y = X\beta + e$, where Y is an $n \times 1$ vector of the observations, X is the $n \times p$ design matrix of the levels of the regression variables, β is an $p \times 1$ vector of the regression coefficients, and e is an $n \times 1$ vector of random errors. Assume that $X'X$ is invertible, i.e., $(X'X)^{-1}$ exists. Prove that the least squares estimator for β is
$$\hat{\beta} = (X'X)^{-1}X'Y.$$

9. Suppose that $\{X_1, \ldots, X_n\}$ is a random sample from the normal distribution $N(\mu, \sigma^2)$ with μ fixed.

   (a). Find a sufficient statistic for σ^2.

   (b). Is the sufficient statistic in part (a) minimally sufficient? Justify your answer.

10. Suppose that $X_1, \cdots, X_n$ are i.i.d. with a beta(μ,1) probability density function given by $f(x) = \mu x^{\mu-1}$, $0 < x \leq 1$, $\mu > 0$, and $Y_1, \cdots, Y_m$ are i.i.d. with a beta(θ, 1) probability density function given by $g(y) = \theta y^{\theta-1}$, $0 < y \leq 1$, $\theta > 0$. Also, assume that X_i's are independent of Y_j's.

(a). Find the likelihood ratio test of $H_0 : \theta = \mu$ versus $H_1 : \theta \neq \mu$.

(b). Show that the test in part (a) can be based on the statistic

$$T = \frac{\sum_{i=1}^{n} \log X_i}{\sum_{i=1}^{n} \log X_i + \sum_{j=1}^{m} \log Y_j}.$$

Qualifying Exam **Name:** ______________________
Department of Mathematical Sciences
Probability and Statistics
8/28/2018
Time Limit: 180 Minutes

Notification of this test:

1. You have 3 hours to complete the exam. You are required to show all the work for all the problems. There are two parts in the exam: (1) Elementary part and (2) Challenging Part. Please budget your time wisely for all two parts. Problems 1 - 5 are elementary part and problems 6 - 10 are challenging part including proof questions. The suggested passing grade for the elementary part is 80%, and challenging part is 70%.

2. Partial credits for some questions (or some sub-questions) will be given, so please try your best and include your answer even you do not finish.

3. The Normal distribution function is provided in Appendix I.

Grade Table (for teacher use only)

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
10	10	
Total:	100	

---

1. (10 points) For Company A there is a 60% chance that no claim is made during the coming year. If one or more claims are made, the total claim amount is normally distributed with mean 10,000 and standard deviation 2,000. For Company B there is a 70% chance that no claim is made during the coming year. If one or more claims are made, the total claim amount is normally distributed with mean 9,000 and standard deviation 2,000. Here we assume the total claim amounts of the two companies are independent. What is the the probability that, in the coming year, Company B's total claim amount will exceed Company A's total claim amount?

2. (10 points) In a given region, the number of tornadoes in a one-week period is modeled by a Poisson distribution with mean 2. The numbers of tornadoes in different weeks are mutually independent. What is the probability that fewer than four tornadoes occur in a three-week period?

3. (10 points) Random variables X and Y are uniformly distributed on the region bounded by the x and y axes, and the curve $y = 1 - x^2$. What is $E(XY)$?

4. (10 points) In a large population of patients, 20% have early stage cancer, 10% have advanced stage cancer, and the other 70% do not have cancer. Six patients from this population are randomly selected. What is the expected number of selected patients with advanced stage cancer, given that at least one of the selected patients has early stage cancer?

5. (10 points) A tour operator has a bus that can accommodate 20 tourists. The operator knows that tourists may not show up, so he sells 21 tickets. The probability that an individual tourist will not show up is 0.02, independent of all other tourists. Each ticket costs 50, and is non-refundable if a tourist fails to show up. If a tourist shows up and a seat is not available, the tour operator has to pay 100, which is the ticket cost plus the penalty cost 50 to the tourist. What is the expected revenue of the tour operator?

6. (10 points) Consider the following situation

x	0	1	2	3	4
$P(x\|\theta_1)$	0.00	0.05	0.05	0.80	0.10
$P(x\|\theta_2)$	0.05	0.05	0.80	0.10	0.00
$P(x\|\theta_3)$	0.90	0.08	0.02	0.00	0.00

A "single" observation is observed. Consider testing $H_0 : \theta = \theta_3$ v.s. $H_1 : \theta \in \{\theta_1, \theta_2\}$.

(a) (5 points) Define the likelihood ratio test and find the test statistic.

(b) (3 points) Find a rejection region of likelihood ratio test (LRT) for $\alpha = 0.02$.

(c) (2 points) Find a power function of your test.

7. (10 points) Let $(X_1, X_2, \ldots, X_n)$ be a simple random sample from a Poisson $P(\lambda)$ distribution and let $S_m = \sum\limits_{i=1}^{m} X_i, m \leq n$.

(a) (5 points) Show that the conditional distribution of $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ given $S_n = k$ is multinomial distribution $M(k, 1/n, \ldots, 1/n)$.

(b) (5 points) Show that $E(S_m|S_n) = (m/n)S_n$.

8. (10 points) Suppose that $X_1, X_2, \ldots, X_n$ is a simple random sample from normal distribution $N(\mu, \sigma^2)$.

(a) (5 points) Show that $\bar{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ and $S_n^2 = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i - \bar{X})^2$ are uniformly minimum variance unbiased estimators (UMVUE) of μ and σ^2, respectively.

(b) (5 points) Prove that S_n^2 converges to σ^2 in probability as $n \to \infty$.

9. (10 points) Prove the following two statements:

(a) (5 points) Using only the Axioms of probability and set theory, prove that

$$A \subset B \Rightarrow P(A) \leq P(B)$$

(b) (5 points) Let $\{A_i\}$ be a sequence of events. Show that

$$P\Big(\bigcup_{i=1}^{\infty} A_i\Big) \leq \sum_{i=1}^{\infty} P(A_i)$$

10. (10 points) Let Z_n converge in probability to z_0 and let $g : R \to R$ be a continuous function. Show that $g(Z_n)$ converges in probability to $g(z_0)$.

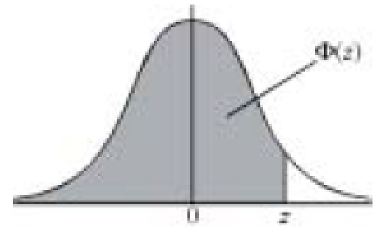
Appendix I: Standard Normal distribution.

THE NORMAL DISTRIBUTION FUNCTION

If Z has a normal distribution with mean 0 and variance 1 then, for each value of z, the table gives the value of $\Phi(z)$, where

$$\Phi(z) = \mathrm{P}(Z \le z).$$

For negative values of z use $\Phi(-z) = 1 - \Phi(z)$.



z	0	1	2	3	4	5	6	7	8	9	ADD 1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	4	8	12	16	20	24	28	32	36
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753	4	8	12	16	20	24	28	32	36
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141	4	8	12	15	19	23	27	31	35
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517	4	7	11	15	19	22	26	30	34
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879	4	7	11	14	18	22	25	29	32
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224	3	7	10	14	17	20	24	27	31
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549	3	7	10	13	16	19	23	26	29
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852	3	6	9	12	15	18	21	24	27
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133	3	5	8	11	14	16	19	22	25
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389	3	5	8	10	13	15	18	20	23
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621	2	5	7	9	12	14	16	19	21
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830	2	4	6	8	10	12	14	16	18
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015	2	4	6	7	9	11	13	15	17
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177	2	3	5	6	8	10	11	13	14
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319	1	3	4	6	7	8	10	11	13
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441	1	2	4	5	6	7	8	10	11
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545	1	2	3	4	5	6	7	8	9
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633	1	2	3	4	4	5	6	7	8
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706	1	1	2	3	4	4	5	6	6
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767	1	1	2	2	3	4	4	5	5
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817	0	1	1	2	2	3	3	4	4
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857	0	1	1	2	2	2	3	3	4
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890	0	1	1	1	2	2	2	3	3
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916	0	1	1	1	1	2	2	2	2
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936	0	0	1	1	1	1	1	2	2
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952	0	0	0	1	1	1	1	1	1
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964	0	0	0	0	1	1	1	1	1
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974	0	0	0	0	0	1	1	1	1
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981	0	0	0	0	0	0	0	1	1
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986	0	0	0	0	0	0	0	0	0

Qualifying Exam on Probability and Statistics
August 24, 2017

Name:

Instruction: There are ten problems at two levels: 5 problems at elementary level and 5 proof problems at graduate level. Therefore, please make sure to budget your time to complete problems at both levels. Show your detailed steps in order to receive credits. You have 3 hours to complete the exam. GOOD LUCK!

Level 1: Elementary Problems

1. A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence of the disease 0.5% of the time when the disease is not actually present. One percent of the population actually has the disease.

   (a). What is the probability that the test indicates the presence of the disease?

   (b). Calculate the probability that a person actually has the disease given that the test indicates the presence of the disease.

2. Let X and Y be independent and identically distributed random variables such that the moment generating function of $X+Y$ is

$$M(t) = 0.09e^{-2t} + 0.24e^{-t} + 0.34 + 0.24e^{t} + 0.09e^{2t}, -\infty < t < \infty.$$

(a). Find $P(X \leq 0)$.

(b). Find $\mu = E(X)$ and $\sigma^2 = Var(X)$.

3. On May 5, in a certain city, temperatures have been found to be normally distributed with mean $\mu = 24°\text{C}$ and variance $\sigma^2 = 9$. The record temperature on that day is $27°\text{C}$.

   (a). What is the probability that the record of $27°\text{C}$ will be broken on next May 5 ?

   (b). What is the probability that the record of $27°\text{C}$ will be broken at least 3 times during the next 5 years on May 5 ? (Assume that the temperatures during the next 5 years on May 5 are independent.)

   (c). How high must the temperature be to place it among the top 5% of all temperatures recorded on May 5 ?

4. A company offers earthquake insurance. Annual premiums are modeled by an exponential random variable with mean 2. Annual claims are modeled by an exponential random variable with mean 1. Premiums and claims are independent. Let X denote the ratio of claims to premiums. Determine the probability density function $f(x)$ of X.

5. Ten cards from a deck of playing cards are in a box: two diamonds, three spades, and five hearts. Two cards are randomly selected without replacement. Find the conditional variance of the number of diamonds selected, given that no spade is selected.

Level 2: Proof Problems

1. Let $\{A_i\}$ be a sequence of events satisfying $\sum_{i=1}^{\infty} P(A_i) < \infty$. Show that

$$P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) = 0.$$

2. Let $\{X_i\}_{i=1}^n$ be i.i.d. and normally distributed with mean μ and variance σ^2. Show that

   (a). the sample mean $\bar{X}_n$ and sample variance S_n^2 are independent.

   (b). $\frac{(n-1)S_n^2}{\sigma^2}$ follows a chi-squared distribution χ_{n-1}^2.

3. Show that the sequence of random variables X_1, $X_2, \ldots$ converges in probability to a constant a if and only if the sequence also converges in distribution to a. That is the statement

$$P(|X_n - a| > \varepsilon) \to 0 \text{ for every } \varepsilon > 0$$

is equivalent to

$$P(X_n \le x) \to \begin{cases} 0 & \text{if } x < a \\ 1 & \text{if } x > a. \end{cases}$$

Questions 4-5: Let $X_1, \ldots, X_n$ be a simple random sample of size n from a population X, and $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ be the sample mean and $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ be the sample variance, respectively.

4. If $E(X) = \mu$ and $0 < Var(X) = \sigma^2 < \infty$, show that

   (a). $S_n^2 \to \sigma^2$ in probability as $n \to \infty$.

   (b). $\frac{S_n}{\sigma} \to 1$ in probability as $n \to \infty$.

5. Assume that X is normally distributed with both mean and standard deviation being θ, $\theta > 0$.

(a). Show that both $\bar{X}_n$ and cS_n are unbiased estimators of θ, where

$$c = \frac{\sqrt{n-1}\Gamma((n-1)/2)}{\sqrt{2}\Gamma(n/2)}.$$

(b). Find the value of t so that the linear interpolation estimator $t\bar{X}_n + (1-t)(cS_n)$ has minimum variance.

(c). Show that $(\bar{X}_n, S_n^2)$ is a sufficient statistic for θ but it is not a complete sufficient statistic.

CHALLENGING PART

1) Let X_n and S_n be two sequences of real random variables such that S_n converges to Z in distribution, where Z is a continuous random variable. If X_n converges to a constant K in probability, show that the following is true

$$X_n S_n \to KZ \text{ in distribution as } n \to \infty$$

Hint: First show that one can assume (without loss of generality) that $|X_n| < M$ for sufficiently large M.

2) Let X_n be a sequence of random variables such that $EX_n = \mu$ and $|Cov(X_n, X_m)| \leq \frac{1}{1+|n-m|}$ for all $n, m \in N$.

a) Show that

$$n^{-1} \sum_{i=1}^{n} X_i \to \mu \text{ in probability as } n \to \infty$$

Hint: Chebishev.

b) Show that

$$n^{-1/2} \sum_{i=1}^{n} (X_i - \mu) \text{ does not converge in distribution (as } n \to \infty)$$

Hint: $Var(n^{-1/2} \sum_{i=1}^{n} X_i) \to \infty$.

3) Let (X, Y) be a random vector and let $H(s,t) = P(X \leq s, Y \leq t)$, $F(s) = P(X \leq s)$, $G(t) = P(Y \leq t)$. We also assume that F and G are continuous functions and H is not necessarily continuous. Define $C(a,b) = H(F^{-1}(a), G^{-1}(b))$, for $(a,b) \in [0,1]^2$.

a) Show that

$$H(s,t) = C(F(s), G(t))$$

b) Show that for every $s \leq s_o$ and $t \leq t_o$ the following is true

$$C(s,t) \leq C(s_o, t_o)$$

PROOF PART

Prove the following:

1) Using only the Axioms of probability and set theory, prove that

a)

$$A \subset B \Rightarrow P(A) \le P(B)$$

b)

$$P(X+Y > \varepsilon) \le P(X > \varepsilon/2) + P(Y > \varepsilon/2)$$

c) If A and B are independent events than A^c and B^c are independent as well

d) If A and B are mutually exclusive and $P(A) + P(B) > 0$ than show that

$$P(A/A \cup B) = \frac{P(A)}{P(A) + P(B)}$$

2) If

$$X_n \to X_0 \text{ in probability and } |X_i| \le M < \infty$$

then

$$E|X_n - X_0| \to 0$$

3) This problem count as two. If EX_i exists then

$$n^{-1} \sum_{i=1}^{n} X_i \to EX_1 \text{ in probability}$$

Qualifying Exam on Probability and Statistics
Spring, January 19, 2016

Instruction: You have 3 hours to complete the exam. You are required to show all the work for all the problems. There are three parts in the exam. Please budget your time wisely for all three parts. There are 10 problems in Elementary part, 3 problems in Challenging part and 3 proof problems. The suggested passing grade for the three parts are: Elementary part 80%, Challenging part 50% and Proofs 80%.

1 Elementary part

(1). The number of injury claims per month is modeled by a random variable N with $P(N = n) = \frac{1}{(n+1)(n+2)}$ for non negative integral $n's$. Calculate the probability of at least one claim during a particular month, given that there have been at most four claims during that month.

(2). Let X be a continuous random variable with density function

$$f(x) = \frac{|x|}{10} \text{ for } x \in [-1, 4] \text{ and } \; f(x) = 0 \text{ otherwise.}$$

Calculate $E(X)$.

(3). A device that continuously measures and records sesmic activity is placed in a remote region. The time to failure of this device, T, is exponentialy distributed with mean 3 years. Since the device will not be monitored during its first two years of service, the time to discovery of its failure is $X = \max(T, 2)$. Calculate $E(X)$.

(4). The time until failure, T, of a product is modeled by uniform distribution on $[0, 10]$. An extended warranty pays a benefit of 100 if failure occurs between $t = 1.5$ and $t = 8$. The present value, W of this benefit is

$$W = 100e^{-0.04T} \text{ for } T \in [1.5, 8] \text{ and zero otherwise.}$$

Calculate $P(W < 79)$.

(5). On any given day, a certain machine has either no malfunctions or exactly one malfunction. The probability of malfunction on any given day is 0.4. Machine malfunctions on different days are mutually independent. Calculate the probability that the machine has its third malfunction on the fifth day, given that the machine has not had three malfunctions in the first three days.

(6). Two fair dice are rolled. Let X be the absolute value of the difference between the two numbers on the dice. Calculate $P(X < 3)$.

(7). A driver and a passenger are in a car accident. Each of them independently has probability 0.3 of being hospitalized. When a hospilatization occurs, the loss is uniformly distributed on $[0, 1]$. When two hospitalization occur, the losses are independent. Calculate the expected number of people in the car who are hospitalized, given that the total loss due to hospitalization is less than 1.

(8). Let X and Y be independent and identically distributed random variables such that the moment generating function for $X + Y$ is

$$M(t) = 0.09e^{-2t} + 0.24e^{-t} + 0.34 + 0.24e^{t} + 0.09e^{2t} \text{ for } \; t \in (-\infty, \infty)$$

Calculate $P(X \leq 0)$.

(9). The number of workplace injuries, N, occuring in a factory on any given day is Poisson distributed with mean λ. The parameter λ itself is a random variable that is determined by the level of activity in the factory and is uniformly distributed on inteval $[0, 3]$. Calculate $Var(N)$.

(10). Let X and Y be continuous random variables with joint density function

$$f(x, y) = \begin{cases} 24xy, & \text{for } \; 0 < y < 1 - x, \; x \in (0, 1); \\ 0, & \text{otherwise} \end{cases}$$

Calculate $P(Y < X | X = \frac{1}{3})$.

2 Challenging Part

(1). Let Y be a non negative random variable. Show that

$$EY \leq \sum_{k=0}^{\infty} P(Y > k) \leq EY + 1.$$

(2). Let X_n be a sequence of random variables such that $\sqrt{n}(X_n - \mu) \to N(0, \sigma^2)$ in distribution. For any given function g and a specific μ, suppose that $g'(\mu)$ exists and $g'(\mu) \neq 0$. Then prove that

$$\sqrt{n}(g(X_n) - g(\mu)) \to N(0, \sigma^2[g'(\mu)]^2) \text{ in distribution.}$$

(3). Let $\{X_n\}$ be a sequence of random variables with $E(X_n) = 0$, and $Var(X_n) \leq C$ (C is a constant), $E(X_i X_j) \leq \rho\big(i-j\big)$ for any $i > j$ and $\rho(n) \to 0$ as $n \to \infty$. Show that

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to 0 \text{ in probability.}$$

3 Proofs

(1). Let $\{X_n\}$ be a sequence of independent and identically distributed random variables with $E|X_n| < \infty$. Prove or disprove the following statement

$$\frac{1}{n}\sum_{k=1}^{n} X_k \to EX_1 \quad \text{in probability as } n \to \infty.$$

(2). Let $X_n : \Omega \to R^d$ and such that X_n converges weakly (in distribution) to random vector Z. Let $F : R^d \to R$ be a continuous function and let $Y_n = F(X_n)$. Then prove or disprove the following statement:

$$Y_n \to F(Z) \text{ weakly (in distribution) as } \quad n \to \infty.$$

(3). Consider the linear regression model $Y = X\beta + e$, where Y is an $n \times 1$ vector of the observations, X is the $n \times p$ design matrix of the levels of the regression variables, β is a $p \times 1$ vector of regression coefficients and e is an $n \times 1$ vector of random errors. Show that the least square estimator for β is $\widehat{\beta} = (X'X)^{-1}X'Y$.