STA 413F2008 Greatest Hits

The following questions taken from earlier assignments. Review them to practice for the Final Examination. They are not to be handed in.

1. Based on lecture notes (or on the text if you want to look in the index),
   (a) State and prove Markov's inequality for a continuous random variable X.
   (b) Use Markov's inequality to prove Chebyshev's inequality.
   (c) Use Chebyshev's inequality to prove the Weak Law of Large Numbers.

2. Let $T_1, T_2, \ldots$ be a sequence of real-valued random variables. If $\lim_{n\rightarrow\infty} E(T_n) = \theta$ and $\lim_{n\rightarrow\infty} Var(T_n) = 0$, show that $T_n \stackrel{P}{\rightarrow} \theta$. We will call this the *Variance Rule*.

3. Let $P(T_n = 0) = \frac{n-1}{n}$ and $P(T_n = n) = \frac{1}{n}$.
   (a) Show $T_n \stackrel{P}{\rightarrow} 0$.
   (b) Does $E(T_n) \rightarrow 0$?
   (c) Does $Var(T_n) \rightarrow 0$?

   This shows you that the Variance Rule is a sufficient but not a necessary condition for convergence in probability.

4. Let $X_1, \ldots, X_n$ be a random sample from a distribution with expected value μ and variance σ^2. The Law of Large Numbers says $\overline{X}_n \stackrel{P}{\rightarrow} \mu$. Show that $\overline{X}_n + a \stackrel{P}{\rightarrow} \mu + a$, where a is a constant. Cite facts from the formula sheet as you use them.

5. As before, let $X_1, \ldots, X_n$ be a random sample from a distribution with expected value μ and variance σ^2. Prove that $3\overline{X}_n \stackrel{P}{\rightarrow} 3\mu$, citing facts from the formula sheet as you use them.

6. Let $X_1, \ldots, X_n$ be a random sample from a distribution with expected value μ and variance σ^2_x. Independently of $X_1, \ldots, X_n$, let $Y_1, \ldots, Y_n$ be a random sample from a distribution with the same expected value μ and variance σ^2_y. Let Let $T_n = \alpha\overline{X}_n + (1-\alpha)\overline{Y}_n$, where α is a constant between zero and one.
   (a) How do you know $\overline{X}_n$ and $\overline{Y}_n$ are independent?
   (b) Is T_n an unbiased estimator of μ? Answer Yes or No and show your work.
   (c) Is T_n a consistent estimator of μ? Answer Yes or No and show your work.
   (d) What value of α would make the estimator T_n most accurate in terms of having the smallest possible variance? Show your work.

7. Let $X_1, \ldots, X_n$ be independent random variables with a continuous uniform distribution on $[0, \theta]$, and let $Y_n = \max(X_1, \ldots, X_n)$.

(a) Using the definition of convergence in probability, show that Y_n is a consistent estimator of θ.

(b) Show that $2\overline{X}_n$ is also consistent for θ.

8. Let X be a random variable with expected value μ and variance σ^2. Show that $\frac{X}{n} \stackrel{P}{\rightarrow} 0$.

9. Let $X_1, \ldots, X_n$ be a random sample from an exponential distribution with parameter θ, and let $T_n = nY_1$, where $Y_1 = \min(X_1, \ldots, X_n)$.

   (a) Find the probability density function of T_n. Don't forget the support.

   (b) Is T_n unbiased? Answer Yes or No. You can use the formula sheet.

   (c) Is T_n consistent? Justify your answer by giving an explicit formula for $P(|T_n - \theta| < \epsilon)$.

10. A model for simple regression through the origin is

$$Y_i = \beta x_i + \epsilon_i$$

for $i = 1, \ldots, n$, where $\epsilon_1, \ldots, \epsilon_n$ are a random sample from a distribution with expected value zero and variance σ^2, and β and σ^2 are unknown constants.

   (a) What is $E(Y_i)$?

   (b) What is $Var(Y_i)$?

   (c) Find the Least Squares estimate of β by minimizing $Q = \sum_{i=1}^{n}(Y_i - \beta x_i)^2$ over all values of β. Let $\widehat{\beta}_n$ denote the point at which Q is minimal.

   (d) Is $\widehat{\beta}_n$ unbiased? Answer Yes or No and show your work.

   (e) Give a sufficient condition for $\widehat{\beta}_n$ to be consistent. Show your work. Remember, in this model the x_i are fixed constants, not random variables.

11. Let $T_1, T_2, \ldots$ be a sequence of random variables. Let T be another random variable, and let θ be a constant.

   (a) Prove $T_n \stackrel{a.s.}{\rightarrow} T \Rightarrow T_n \stackrel{P}{\rightarrow} T$. Recall the approach from lecture. $T_n \stackrel{a.s.}{\rightarrow} T$ means there is a set A contained in the sample space $\mathcal{C}$ with $\lim_{n\rightarrow\infty} T_n(c) = T(c)$ for every $c \in A$. Let $\epsilon > 0$ be given, and define $A_n = \{c \in A : |T_k(c) - T(c)| < \epsilon \text{ for all } k \geq n\}$, and so on.

   (b) Prove $T_n \stackrel{P}{\rightarrow} \theta \Rightarrow T_n \stackrel{d}{\rightarrow} \theta$. The trick is to make sure ϵ is small enough so that the point you are considering is outside the interval from $\theta - \epsilon$ to $\theta + \epsilon$.

   (c) Prove $T_n \stackrel{d}{\rightarrow} \theta \Rightarrow T_n \stackrel{P}{\rightarrow} \theta$. This holds only when the target is a constant, not a (non-degenerate) random variable.

12. Let $X_1, \ldots, X_n$ be a random sample from a distribution with a density that is uniform on the interval from zero to θ, and let Y_n denote the maximum. We will investigate the limiting behaviour of $T_n = n(1 - \frac{Y_n}{\theta})$.

(a) What is the support of Y_n?

(b) What is the support of T_n?

(c) If T_n has a limiting distribution, what does the support of that limiting distribution have to be?

(d) Now write the cumulative distribution function of T_n as $Pr\{n(1-\frac{Y_n}{\theta}) \leq t\}$, simplify, and take the limit as $n \rightarrow \infty$. Use Problem **??**.

(e) If you recognize the result, great; what is the limiting distribution called? otherwise, differentiate it to obtain a familar density.

13. Do Exercise 4.3.3. Hints: Assume the cumulative distribution function F is strictly increasing and therefore has a unique inverse. Remember, Z_n must always be non-negative.

14. Let $X : \mathcal{C} \rightarrow \mathbb{R}$ be a random variable. Prove that

$$\frac{X}{n} \stackrel{a.s.}{\rightarrow} 0.$$

This is surprisingly fast.

15. Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with mean μ and variance σ^2. Prove that S_n^2 is consistent for σ^2. When you use something from the formula sheet, say so.

16. Let $X_1, \ldots, X_n$ be a random sample from a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$. Is the estimator

$$\widehat{\alpha}_n = \frac{\overline{X}_n^2}{S_n^2}$$

consistent for α? Answer Yes or No and prove your answer. When you use something from the formula sheet, say so.

17. Let $f_{X_n}(x) = \frac{1}{4}I(x=0) + \frac{1}{2}I(x=1) + \frac{1}{4}I(x=\frac{n+1}{n})$.

(a) Is X_n discrete, or is it continuous?

(b) What is $F_{X_n}(x)$? You may write the answer as a case function, or you may write it using indicators.

(c) Let $g(x) = \lim_{n\rightarrow\infty} f_{X_n}(x)$. Consider the cases $x = 0$, $x = 1$ and x equals something else separately. Is $g(x)$ a probability distribution?

(d) Let $G(x) = \lim_{n\rightarrow\infty} F_{X_n}(x)$. Your answer should apply to all x. Is $G(x)$ a cumulative distribution function?

(e) Let X be a Bernoulli random variable with $\theta = \frac{3}{4}$. Denote the cumulative distribution function of X by $F_X(x)$. At what points is $F_X(x)$ discontinuous? Does $G(x)$ equal $F_X(x)$ except possibly at those points? Does this mean $X_n \stackrel{d}{\rightarrow} X$? (Check the definition.)

(f) Do we have $\lim_{n\to\infty} E(X_n) = E(X)$? Answer Yes or No. Show your work.

(g) Do we have $\lim_{n\to\infty} Var(X_n) = Var(X)$? Answer Yes or No. Show your work.

18. Let the discrete random variable X_n have probability mass function $p_{X_n}(x) = \frac{1}{3}I(x=0) + \frac{2}{3}\left(\frac{n-1}{n}\right)I(x=1) + \frac{2}{3n}I(x=n)$.

    (a) What is $F_{X_n}(x)$? You may write the answer as a case function, or you may write it using indicators.

    (b) Let $p(x) = \lim_{n\to\infty} p_{X_n}(x)$. Consider the cases $x = 0$, $x = 1$ and x equals something else separately. Is $p(x)$ a probability distribution?

    (c) Let $F(x) = \lim_{n\to\infty} F_{X_n}(x)$. Your answer should apply to all x.

    i. What is $F(x)$? You may write the answer as a case function, or you may write it using indicators.

    ii. Is $F(x)$ a cumulative distribution function? Does it correspond to $p(x)$?

    iii. We seem to have $X_n \stackrel{d}{\rightarrow} X$, where X is Bernoulli again. What is the parameter θ?

    (d) Do we have $\lim_{n\to\infty} E(X_n) = E(X)$? Answer Yes or No. Show your work.

    (e) Do we have $\lim_{n\to\infty} Var(X_n) = Var(X)$? Answer Yes or No. Show your work.

19. Let $X_1, \ldots, X_n$ be a random sample from a distribution with expected value μ and variance σ^2. The "Modified" Central Limit Theorem (see lecture notes and Formula Sheet) says that the usual Central Limit Theorem still holds if σ is replaced by any consistent estimator – call it $\widehat{\sigma}_n$. Using this result, *derive* an approximate $(1-\alpha)100\%$ confidence interval for μ. Show your work.

20. A random sample of size $n = 150$ yields a sample mean of $\overline{X}_n = 8.2$. Give a point estimate and an approximate 95% confidence interval

    (a) For λ, if $X_1, \ldots, X_n$ are from a Poisson distribution with parameter λ.

    (b) For θ, if $X_1, \ldots, X_n$ are from an Exponential distribution with parameter θ.

    (c) For μ, if $X_1, \ldots, X_n$ are from a Normal distribution with mean μ and variance one. (This confidence interval is exact, not an approximation.)

    (d) For θ, if $X_1, \ldots, X_n$ are from a Uniform distribution on $[0, \theta]$.

    (e) For θ, if $X_1, \ldots, X_n$ are from a Uniform distribution on $[\theta, \theta + 1]$.

    (f) For θ, if $X_1, \ldots, X_n$ are from a Geometric distribution with parameter θ.

    (g) For θ, if $X_1, \ldots, X_n$ are from a Binomial distribution with parameters 10 and θ.

    In each case, your answer is three numbers: the point estimate, the lower confidence limit and the upper confidence limit.

21. Let $T_n \stackrel{P}{\rightarrow} \theta$, and let Y_n that is between T_n and θ. Prove that $Y_n \stackrel{P}{\rightarrow} \theta$. We will call this the Squeeze Theorem for Convergence in Probability.

22. The various parts of this question will lead you through the proof of the Delta Method for the special but important case of a function of the sample mean. That is, let $X_1, \ldots, X_n$ be a random sample from a distribution with expected value μ and variance σ^2, and let $g(x)$ be a function with $g'(\mu) \neq 0$ and $g''(x)$ continuous at $x = \mu$. Then

$$\sqrt{n}(g(\overline{X}_n) - g(\mu)) \stackrel{d}{\rightarrow} Y \sim N(0, g'(\mu)^2\sigma^2)$$

(a) First, use Taylor's Theorem (on the formula sheet) to re-write

$$\sqrt{n}(g(\overline{X}_n) - g(\mu)), \tag{1}$$

distributing $\sqrt{n}$. You now have two terms.

(b) How do you know $\sqrt{n}(\overline{X}_n - \mu)$ converges in probability to a normal random variable? Cite something from the formula sheet.

(c) How do you know $g''(\mu^*) \stackrel{P}{\rightarrow} g''(\mu)$? Use the formula sheet as well as another problem from this assignment.

(d) Now show that the second of the two terms from Problem 22a converges in probability to zero. When you use something from the formula sheet, cite it.

(e) Now show that the first term from Problem 22a converges in distribution to something. What is its target distribution? Again, when you use something from the formula sheet, cite it.

(f) Now apply Slutsky (c) for convergence in distribution.

(g) Now do the last step. You are using a particular continuous function $g(a, b)$. Write it in terms of a and b.

23. Let $X_1, \ldots, X_n$ be a random sample from a Bernoulli distribution with parameter θ. Find the limiting distribution of

$$Z_n = 2\sqrt{n}\left(\sin^{-1}\sqrt{\overline{X}_n} - \sin^{-1}\sqrt{\theta}\right).$$

Hint: $\frac{d}{dx}\sin^{-1}(x) = \frac{1}{\sqrt{1-x^2}}$.

24. Let $X_1, \ldots, X_n$ be a random sample from an exponential distribution with parameter θ.

(a) Find a variance-stabilizing transformation. That is, find a function $g(x)$ such that the limiting distribution of

$$Y_n = \sqrt{n}[g(\overline{X}_n) - g(\theta)]$$

does not depend on θ.

(b) To check, find the limiting distribution of Y_n.

25. Let $X_1, \ldots, X_n$ be a random sample from a uniform distribution on $(0, \theta)$.

(a) Find a variance-stabilizing transformation. That is, find a function $g(x)$ such that the limiting distribution of

$$Y_n = \sqrt{n}[g(2\overline{X}_n) - g(\theta)]$$

does not depend on θ.

(b) To check, find the limiting distribution of Y_n.

26. Let $X_1, \ldots, X_n$ be a random sample from a chi-square distribution with parameter ν.

    (a) Find a variance-stabilizing transformation. That is, find a function $g(x)$ such that the limiting distribution of

    $$Y_n = \sqrt{n}[g(\overline{X}_n) - g(\nu)]$$

    does not depend on ν.

    (b) To check, find the limiting distribution of Y_n.

27. Let $X_1, \ldots, X_n$ be a random sample from a binomial distribution with parameters m and θ. The constant m is fixed and known.

    (a) Find a variance-stabilizing transformation. That is, find a function $g(x)$ such that the limiting distribution of

    $$Y_n = \sqrt{n}[g(\overline{X}_n/m) - g(\theta)]$$

    does not depend on θ.

    (b) To check, find the limiting distribution of Y_n.

28. Let $X_1, \ldots, X_n$ be a random sample from a distribution (not necessarily normal) with expected value μ and variance σ^2. We will test $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ using the critical region

    $$C = \{(x_1, \ldots, x_n) : \left| \frac{\overline{x} - \mu_0}{s/\sqrt{n}} \right| > z_{\alpha/2}\}$$

    (a) What is the approximate size of the test?

    (b) Show that H_0 is rejected if and only if the $(1-\alpha)100\%$ confidence interval for μ does *not* include μ_0. It is easiest to start by writing the set of $(x_1, \ldots, x_n)$ such that μ_0 is in the confidence interval, and then work on it until it becomes C^c.

29. Let $X_1, \ldots, X_{n_1}$ be a random sample from an exponential distribution with parameter θ. The null hypothesis is $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$.

    (a) Give the critical region C_1 for an approximate size α test based on the modified Central Limit Theorem. Your answer will use critical values from the standard normal distribution.

(b) Write the power function of this test explicitly as a function of the true parameter θ. You answer will involve Φ, the cumulative distribution function of a standard normal.

(c) Suppose $\theta_0 = 2$, $\alpha = 0.05$ and $n = 30$. What is the approximate power for $\theta = 2.5$? The answer is a single number. I get a power of 0.4129.

(d) What is the smallest sample size that will guarantee an approximate power of 0.80 for $\theta = 2.5$? The answer is a single number (an integer, of course). I get $n = 117$.

(e) Again for $\theta_0 = 2$, $\alpha = 0.05$ and $n = 30$, suppose we observe $\overline{X}_n = 3$.

i. What is the value of the test statistic? The answer is a number.

ii. Is $\mathbf{X} \in C$? Answer Yes or No.

iii. Do you reject H_0? Answer Yes or No.

iv. Calculate the p-value. The answer is a number. I get 0.0031.

30. Let $X_1, \ldots, X_n$ be a random sample from a Poisson distribution with parameter λ. Consider the following critical region for testing $H_0 : \lambda \geq \lambda_0$ versus $H_1 : \lambda < \lambda_0$:

$$C = \left\{ (x_1, \ldots, x_n) : \frac{\sqrt{n}(\overline{x}_n - \lambda_0)}{\sqrt{\lambda_0}} < -z_\alpha \right\}$$

(a) Derive a formula for the power function $P_\lambda(\mathbf{X} \in C)$. "Derive" means show all your work.

(b) For $\alpha = 0.05$, $n = 50$ and $\lambda_0 = 4$, what is the approximate power of the test at $\lambda = 3$? I get 0.539.

(c) What is ω_0 for this problem?

(d) What is ω_1 for this problem?

(e) What is the approximate distribution of the test statistic when $\lambda = \lambda_0$? Cite the formula sheet to support your answer.

(f) If $\alpha = 0.05$, and $\lambda_0 = 4$, what is the minimum sample size required so that the approximate power of the test will be at least 0.80 at $\lambda = 3$?

31. Let $X_1, \ldots, X_n$ be a random sample from an exponential distribution with parameter θ. Consider the following critical region for testing $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$:

$$C_1 = \left\{ (x_1, \ldots, x_n) : \sqrt{n} \left(\frac{\overline{x}_n}{\theta_0} - 1 \right) > z_\alpha \right\}$$

(a) Derive a formula for the power function $P_\theta(\mathbf{X} \in C_1)$. "Derive" means show all your work.

(b) For $\alpha = 0.05$, $n = 50$ and $\theta_0 = 4$, what is the approximate power of the test at $\theta = 5$?

(c) What is ω_0 for this problem?

(d) What is ω_1 for this problem?

(e) For $\theta \in \omega_0$, does the power function attain its maximum at $\theta = \theta_0$? Answer Yes or No and prove your answer. A picture may help.

(f) What is the approximate distribution of the test statistic when $\theta = \theta_0$? Show some calculations and cite something from the formula sheet to support your answer.

(g) Does this mean the test is of (approximate) size α? Answer Yes or No.

(h) If $\alpha = 0.05$, and $\theta_0 = 4$, what is the minimum sample size required so that the approximate power of the test will be at least 0.80 at $\theta = 5$? Your answer is a single number. My answer is 117.

32. Let $X_1, \ldots, X_n$ be a random sample from a Gamma distribution with parameters α (unknown) and $\beta = 1$ (known). Using just the Modified Central Limit Theorem (no variance-stabilizing transformations, please!), give an approximate critical region of size 0.05 for the following null and alternative hypotheses. Please use specific numbers for your critical values, not symbols. You don't have to prove anything or justify your answers. But you do have to be aware of the justification in order to get the right answer.

    (a) $H_0 : \alpha = \alpha_0$ versus $H_1 : \alpha \neq \alpha_0$.

    (b) $H_0 : \alpha = \alpha_0$ versus $H_1 : \alpha > \alpha_0$.

    (c) $H_0 : \alpha \leq \alpha_0$ versus $H_1 : \alpha > \alpha_0$.

    (d) $H_0 : \alpha = \alpha_0$ versus $H_1 : \alpha < \alpha_0$.

    (e) $H_0 : \alpha \geq \alpha_0$ versus $H_1 : \alpha < \alpha_0$.

33. For Question 32, suppose $n = 150$, $\alpha_0 = 7.5$ and you observe a sample mean of $\overline{X}_n = 8.2$. Test $H_0 : \alpha \leq \alpha_0$ versus $H_1 : \alpha > \alpha_0$.

    (a) What is the value of the test statistic? The answer is a single number.

    (b) What is the p-value? The answer is a single number.

    (c) Do you reject H_0? Answer Yes or No.

    (d) Is $p < \alpha$? Answer Yes or No.

34. Let $X_1, \ldots, X_n$ be a random sample from a Binomial distribution with parameters $m = 10$ (known) and θ (unknown). Using just the Modified Central Limit Theorem (no variance-stabilizing transformations, please!), give an approximate critical region of size 0.05 for the following null and alternative hypotheses. Please use specific numbers for your critical values, not symbols. You don't have to prove anything or justify your answers. But you do have to be aware of the justification in order to get the right answer.

    (a) $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.

    (b) $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$.

(c) $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$.

(d) $H_0 : \theta = \theta_0$ versus $H_1 : \theta < \theta_0$.

(e) $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$.

35. For Question 34, suppose $n = 150$, $\theta_0 = 0.75$ and you observe a sample mean of $\overline{X}_n = 8.2$. Test $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$.

    (a) What is the value of the test statistic? The answer is a single number.

    (b) What is the p-value? The answer is a single number.

    (c) Do you reject H_0? Answer Yes or No.

    (d) Is $p < \alpha$? Answer Yes or No.

36. Let $X_1, \ldots, X_n$ be a random sample from a Geometric distribution with parameter θ. Using just the Modified Central Limit Theorem (no variance-stabilizing transformations, please!), give an approximate critical region of size 0.05 for the following null and alternative hypotheses. Please use specific numbers for your critical values, not symbols. You don't have to prove anything or justify your answers. But you do have to be aware of the justification in order to get the right answer.

    (a) $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.

    (b) $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$.

    (c) $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$.

    (d) $H_0 : \theta = \theta_0$ versus $H_1 : \theta < \theta_0$.

    (e) $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$.

37. For Question 36, suppose $n = 150$, $\theta_0 = 0.10$ and you observe a sample mean of $\overline{X}_n = 8.2$. Test $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.

    (a) What is the value of the test statistic? The answer is a single number.

    (b) What is the p-value? The answer is a single number.

    (c) Do you reject H_0? Answer Yes or No.

    (d) Is $p < \alpha$? Answer Yes or No.

38. Let the discrete random variable X have probability mass function $p(x;\theta)$, and assume the regularity condition

$$\frac{\partial^2}{\partial\theta^2} \sum_x p(x;\theta) = \sum_x \frac{\partial^2}{\partial\theta^2} p(x;\theta).$$

    (a) Is the regularity condition satisfied by the Bernoulli distribution? Answer Yes or No and justify your answer.

    (b) Let the *score* random variable $S = \frac{\partial}{\partial\theta} \ln p(X;\theta)$. Show $E(S) = 0$.

(c) Show $Var(S) = I(\theta)$, where

$$I(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2}\ln p(X;\theta)\right).$$

(d) Let $X_1, \ldots, X_n$ be a random sample from a discrete distribution with probability mass function $p(x;\theta)$, let $Y = u(X_1, \ldots, X_n)$ have expected value $E(Y) = k(\theta)$, and let $Z = \sum_{i=1}^{n} S_i$. Prove $k'(\theta) = E(YZ)$.

(e) What is $E(Z)$?

(f) What is $Cov(Y, Z)$?

(g) What is $Corr(Y, Z)$?

(h) Using the fact that the absolute value of a correlation cannot be greater than one, establish the Cramér-Rao inequality of Theorem 6.2.1.

(i) Give a lower bound for the variance of any unbiased estimator of θ, based on a rndom sample from this distribution.

39. Let $X_1, \ldots, X_n$ be a random sample from a Poisson distribution with parameter λ. Find the MLE of λ. Is it efficient? Answer Yes or No and prove your answer.

40. Let $X_1, \ldots, X_n$ be a random sample from a Binomial distribution with parameters 4 and θ. Find the MLE of θ. Is it efficient? Answer Yes or No and prove your answer.

41. Let $X_1, \ldots, X_n$ be a random sample from a distribution with density $f(x;\theta) = \theta x^{\theta-1} I(0 < x < 1)$, where $\theta > 0$.

   (a) Calculate the Fisher information $I(\theta)$.

   (b) What is the lowest possible variance for an unbiased estimator of θ, based on a random sample of size n? Your answer is a function of n and θ.

   (c) Find the Maximum Likelihood Estimator of θ.

42. Let $X_1, \ldots, X_n$ be a random sample from a distribution with density $f(x;\theta_0)$.

   (a) The random variable
   $$\frac{1}{n}\sum_{i=1}^{n} \frac{\partial}{\partial\theta} \ln f(X_i;\theta)$$
   converges in probability to something. What is the target? How do you know?

   (b) The random variable
   $$-\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta^2} \ln f(X_i;\theta)$$
   converges in probability to something. What is the target? How do you know?

   (c) Write down a Taylor expansion (two terms plus remainder) of the function $\ell'(\theta)$; expand about the true parameter θ_0.

   (d) Assuming the remainder term may be disregarded, find the limiting distribution of $\sqrt{n}(\widehat{\theta}_n - \theta_0)$. Cite facts from the formula sheet and Homework Assignment 10 as you use them.

   (e) Prove $\sqrt{nI(\widehat{\theta}_n)}\,(\widehat{\theta}_n - \theta_0) \stackrel{d}{\rightarrow} Z \sim N(0,1)$, citing facts from the formula sheet as you use them.

   (f) Derive a $(1-\alpha)100\%$ confidence interval for θ. Show your work.

43. Let $X_1, \ldots, X_n$ be a random sample from a distribution with density
   $$f(x;\theta) = (\theta+1)x^{\theta} I(0 < x < 1),$$
   where $\theta > 0$.

   (a) Verify
   $$\mu = \frac{\theta+1}{\theta+2} \quad \text{and} \quad \sigma^2 = \frac{\theta+1}{(\theta+3)(\theta+2)^2}.$$

   (b) Let
   $$T_n = \frac{2\overline{X}_n - 1}{1 - \overline{X}_n}.$$
   Is T_n a consistent estimator of θ? Answer Yes or No and prove your answer, citing facts from the formula sheet as you use them.

(c) Use the delta method to show $\sqrt{n}(T_n - \theta_0) \xrightarrow{d} X$. What is the distribution of X? Give its mean and variance.

(d) Obtain a formula for the MLE of θ. Show your work.

(e) Calculate the Fisher Information $I(\theta)$. Show your work.

(f) $\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{d} Y$. What is the distribution of Y? Give its mean and variance.

(g) Compare $Var(X)$ from Question 43c to $Var(Y)$ from Question 43f. Which is smaller? Prove your answer. Of course smaller is better.

(h) A random sample of size $n = 50$ yields $\sum_{i=1}^{n} X_i = 39.557$ and $\sum_{i=1}^{n} \ln(X_i) = -12.947$.

i. What is T_n? Your answer is a single number.

ii. What is $\widehat{\theta}_n$? Your answer is a single number.

iii. Give an approximate 95% confidence interval for θ based on T_n. Your answer is a pair of numbers.

iv. Give an approximate 95% confidence interval for θ based on $\widehat{\theta}_n$. Your answer is a pair of numbers.

44. Let $X_1, \ldots, X_n$ be a random sample from a geometric distribution with parameter θ.

(a) Let $T_n = 1/\overline{X}_n$. Is T_n a consistent estimator of θ? Answer Yes or No and prove your answer, citing facts from the formula sheet as you use them.

(b) Use the delta method to show $\sqrt{n}(T_n - \theta) \xrightarrow{d} X$. What is the distribution of X? Give its mean and variance.

(c) Obtain a formula for the MLE of θ. Show your work.

(d) Calculate the Fisher Information $I(\theta)$. Show your work.

(e) $\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{d} Y$. What is the distribution of Y? Give its mean and variance.

(f) Compare $Var(X)$ from Question 44b to $Var(Y)$ from Question 44e.

(g) A random sample of size $n = 100$ yields $\overline{X}_n = 1.44$.

i. What is T_n? Your answer is a single number.

ii. What is $\widehat{\theta}_n$? Your answer is a single number.

iii. Give an approximate 95% confidence interval for θ based on T_n. Your answer is a pair of numbers.

iv. Give an approximate 95% confidence interval for θ based on $\widehat{\theta}_n$. Your answer is a pair of numbers.

45. From the text, do Exercises 4.2.4, 4.2.5, 4.3.1, 4.3.2, 4.3.3, 4.3.8, 4.4.11, 4.4.12, 6.1.1, 6.1.2, 6.1.3, 6.1.5, 6.1.6, 6.2.1, 6.2.7, 6.2.8.