UNIVERSITY OF TORONTO

Faculty of Arts and Science

April 2018 EXAMINATIONS

STA261H1S

Duration - 3 Hours

Aids Allowed: Non-programmable calculator

First Name: MARKING

Last Name: SCHEME.

Student Number: 42.

This exam booklet contains 25 pages. Write all final answers in the exam booklet. Use the 3 pages at the end, or the backs of the pages, for rough work. Fill in multiple choice answers using the bubble sheet at the end of the exam, in pencil. Answer all other questions in pen.

Questions:

| Question | Marks Achieved | Total Possible |
|---|---|---|
| MC | | 20 |
| 1 | | 18 |
| 2 | | 20 |
| 3 | | 16 |
| 4 | | 20 |
| 5 | | 16 |
| Total | | 110 |

FORMULA SHEET

You may use results on this sheet without proof.

If $Z \sim N(0,1)$ then $P(Z < -1.96) = 0.025$ and $P(Z < 1.96) = 0.975$.

If $\hat{\theta}$ is the MLE for $\theta$ and $\theta_0$ is the true value then

$$\frac{\hat{\theta} - \theta_0}{1/\sqrt{I(\theta_0)}} \xrightarrow{d} N(0,1)$$

If $\bar{X}$ is the sample mean then

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{Var(\bar{X})}} \xrightarrow{d} N(0,1)$$

If there are $d$ free parameters under $H_0$, and $p > d$ free parameters under $H_1$, then as $n \to \infty$, for a likelihood ratio test of $H_0$ against $H_1$,

$$-2 \log \Lambda \xrightarrow{d} \chi^2_{p-d}$$

For $y_{ij}$ the counts in an $R \times C$ contingency table, and $N = \sum_{i=1}^{R} \sum_{j=1}^{C} y_{ij}$, $r_i = \sum_{j=1}^{C} y_{ij}$ and $c_j = \sum_{i=1}^{R} y_{ij}$, then

$$-2 \log \Lambda = 2 \sum_{i=1}^{R} \sum_{j=1}^{C} y_{ij} \log \left( \frac{N y_{ij}}{r_i c_j} \right) \xrightarrow{d} \chi^2_D$$

where $D$ is the appropriate degrees of freedom.

If $W_d \sim \chi^2_d$ then for $d = 1, 2, 3$:

$$P(W_1 < 3.84) = 0.95$$
$$P(W_2 < 5.99) = 0.95$$
$$P(W_3 < 7.81) = 0.95$$

If $X \sim N(\mu, \sigma^2)$, then

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left( -\frac{1}{2\sigma^2}(x-\mu)^2 \right)$$
$$E(X) = \mu$$
$$Var(X) = \sigma^2$$

Each of the following multiple choice questions is worth 1 mark. FILL IN THE APPROPRIATE BUBBLE IN THE SHEET ON THE LAST PAGE OF THE EXAM. You should use pencil, and only fill in your answer when you are sure it won't be changed. Answers circled here will not be marked. This is different than the midterm.

1. Suppose we use IID random variables $X_i \sim F_\theta$ leading to $\hat{\theta}(\mathbf{X})$, an estimator for $\theta$. $\hat{\theta}(\mathbf{X})$ is
   (a) A fixed, known constant
   (b) A fixed, unknown constant
   (c) A random variable
   (d) A sufficient statistic
2. Regarding question 1: if the mean of the sampling distribution of $\hat{\theta}$ is equal to $\theta$, then $\hat{\theta}$ is ___ for $\theta$:
   (a) Efficient
   (b) Unbiased
   (c) Consistent
   (d) Sufficient
3. Regarding question 1: if as the sample size approaches infinity, $\hat{\theta}$ becomes arbitrarily close to $\theta$ with arbitrarily high probability, then $\hat{\theta}$ is ___ for $\theta$:
   (a) Efficient
   (b) Unbiased
   (c) Consistent
   (d) Sufficient
4. Regarding question 1: if $\hat{\theta}$ is unbiased, and has variance at least as low as any other unbiased estimator of $\theta$, then $\hat{\theta}$ is ___ for $\theta$:
   (a) Efficient
   (b) Unbiased
   (c) Consistent
   (d) Sufficient
5. Regarding question 1: if the conditional distribution $f(\mathbf{X}|\hat{\theta})$ doesn't depend on $\theta$, then $\hat{\theta}$ is ___ for $\theta$:
   (a) Efficient
   (b) Unbiased
   (c) Consistent
   (d) Sufficient
6. The p-value for a test of $H_0$ against $H_1$ is defined to be
   (a) The probability of observing a test statistic with equal or greater evidence against $H_0$, if $H_0$ is true
   (b) The probability that $H_0$ is true
   (c) The same thing as the significance level of the test
   (d) The highest signficance level for which we would accept $H_0$

7. Suppose we test $H_0$ against $H_1$ at significance level $\alpha$. What is the power of this test if $H_0$ is true?
   (a) $\alpha$ $P(\text{reject } H_0 \mid H_0 \text{ true}) = \alpha.$
   (b) $1-\alpha$
   (c) 0
   (d) 1
8. Which of the following random variables, coupled with their corresponding probability distributions, would be considered a pivot for parameter $\theta$?
   (a) $Z = X/\theta \sim Exp(1)$
   (b) $Z = X \sim Exp(1)$
   (c) $Z = X/\theta \sim Exp(\theta)$
   (d) $Z = \theta$ with probability 1
9. Which of the following is a Type I error?
   (a) Reject $H_0$ when it is true
   (b) Reject $H_0$ when it is false
   (c) Accept $H_0$ when it is false
   (d) Fail to reject $H_0$ when it is false
10. Which of the following is a Type II error?
   (a) Reject $H_0$ when it is true
   (b) Reject $H_0$ when it is false
   (c) Accept $H_0$ when it is false
   (d) Fail to reject $H_0$ when it is false
11. We wish to test $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ at significance level $\alpha$. We compute the corresponding $1-\alpha$ confidence interval, and find that it contains $\mu_0$. What do we conclude?
   (a) We can't make any conclusions
   (b) Reject $H_0$ in favour of $H_1$
   (c) Accept $H_0$
   (d) Fail to reject $H_0$ in favour of $H_1$
12. From IID random variables $X_i \sim F_\theta$, we construct a confidence interval for parameter $\mu$ of the form $(L(\mathbf{X}), U(\mathbf{X}))$. What is random and what is fixed (not random)?
   (a) $\mu$ random, $(L(\mathbf{X}), U(\mathbf{X}))$ random
   (b) $\mu$ random, $(L(\mathbf{X}), U(\mathbf{X}))$ fixed
   (c) $\mu$ fixed, $(L(\mathbf{X}), U(\mathbf{X}))$ random
   (d) $\mu$ fixed, $(L(\mathbf{X}), U(\mathbf{X}))$ fixed

13. We have a random sample $X_i \sim F_\theta, i = 1 \dots n$, with log-likelihood $\ell(\theta) = \sum_{i=1}^n \log f(x_i|\theta)$ and score statistic $S(\theta) = \partial\ell/\partial\theta$. Suppose the true value of $\theta$ is $\theta_0$. Which of the following is not a regularity condition for the theorem that $E(S(\theta_0)) = 0$, as described in lecture?
    (a) $\theta_0$ is an interior point of the parameter space
    (b) $\ell(\theta)$ is three-times continuously differentiable with respect to $\theta$, within the parameter space
    (c) The support of $X_i$ doesn't depend on $\theta$
    (d) $\ell(\theta)$ is a convex function of $\theta$, within the parameter space
14. We are testing $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ at the $\alpha$ level of significance, using a random sample of size $n$. Suppose the true standard deviation of the distribution of our sample is $\sigma$. On which of the following does the power of the test not depend?
    (a) $\alpha$
    (b) The effect size $d = \frac{\mu_0 - \mu_1}{\sigma}$
    (c) The p-value of the test
    (d) The sample size $n$
15. Suppose we have an IID random sample $X_i \sim F_\theta, i = 1 \dots n$ and we wish to test $H_0 : \theta = \theta_0$ using a parametric bootstrap procedure. We have in mind some estimator $\hat{\theta}$. Our first step would be
    (a) Compute $\hat{\theta}$ from our sample
    (b) Compute $\hat{\theta}$ from a random sample from $F_{\theta_0}$
    (c) Compute $\hat{\theta}$ from a random sample from $F_{\hat{\theta}}$
    (d) Compute $\hat{\theta}$ from a random sample from a $Unif(0, 1)$ distribution
16. The next step in our procedure would be to generate a boostrap sample of $\hat{\theta}$, $\left\{\hat{\theta}_b\right\}_{b=1}^B$ as follows:
    (a) Simulate $B$ values of $X_i$ from a $Unif(0, 1)$ distribution
    (b) Simulate $B$ values of $X_i$ from $F_{\theta_0}$
    (c) Simulate $B$ values of $\hat{\theta}$ by simulating $B$ samples from $F_{\theta_0}$ and calculating $\hat{\theta}$ for each
    (d) Simulate $B$ values of $\hat{\theta}$ by simulating $B$ samples from $F_{\hat{\theta}}$ and calculating $\hat{\theta}$ for each
17. The bootstrap p-value under the above method is obtained as
    (a) The frequency with which $\hat{\theta}_b$ is larger than $\hat{\theta}$
    (b) The frequency with which $\left|\hat{\theta}_b\right|$ is larger than $\hat{\theta}$
    (c) The frequency with which $\left|\hat{\theta}_b\right|$ is larger than $|\theta_0|$
    (d) The frequency with which $\left|\hat{\theta}_b - \theta_0\right|$ is larger than $\left|\hat{\theta} - \theta_0\right|$

18. Suppose now we have the same IID random sample $X_i \sim F_\theta, i = 1 \ldots n$ and we wish to estimate the variance of $g(\mathbf{X}) = \sin(2\pi\bar{X})$. We don't know the distribution of $Y = g(\mathbf{X})$, so we turn to a non-parametric bootstrap. Our first step would be
    (a) Generate $B$ random samples from $F_\theta$ by taking $B$ random samples of size $N$ from $F_\theta$
    (b) Generate $B$ random samples from $F_\theta$ by taking $B$ random samples of size $N/B$ from $F_\theta$
    (c) Generate $B$ random samples from $F_\theta$ by taking $B$ random samples of size $N$ from our original dataset, sampling with replacement
    (d) Generate $B$ random samples from $F_\theta$ by taking $B$ random samples of size $N/B$ from our original dataset, sampling without replacement
19. Denoting the above random samples by $\{\mathbf{x}_b\}_{b=1}^{B}$, our next step would be
    (a) Compute $y_b = g(\mathbf{x}_b)$ for each sample
    (b) Compute $y_b = \bar{x}_b$ for each sample
    (c) Compute $SD(\mathbf{x}_b)$ for each sample
    (d) Combine them all and compute $y = g(\mathbf{x})$
20. The bootstrap estimate of $Var(Y) = Var(g(\mathbf{X}))$ is then
    (a) $\frac{1}{B}\sum_{b=1}^{B}(y_b - \bar{y})^2$
    (b) $g\left(\frac{1}{B}\sum_{b=1}^{B}(\bar{x}_b - \bar{\bar{x}})^2\right)$
    (c) $g\left(\frac{1}{B}\sum_{b=1}^{B}\sum_{i=1}^{n}(x_{ib} - \bar{x}_i)^2\right)$
    (d) $\frac{1}{B}\sum_{b=1}^{B}(g(y_b) - g(\bar{y}))^2$

1. (~~15~~ 16 marks) Let $X \sim N(\mu, \sigma^2)$ be a random variable following a normal distribution. Let $X_i, i = 1 \dots n$ be a random sample from this distribution. In what follows, you may assume any regularity conditions are satisfied where necessary.

(a) (4 marks) Find the Method of Moments estimators of $(\mu, \sigma^2)$.

From formula sheet: $E(X) = \mu$, $Var(X) = \sigma^2$ 2

1 $\Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$

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

OR $\frac{1}{n-1}$

(b) (6 marks) Find the Maximum Likelihood Estimators of $(\mu, \sigma^2)$

From formula sheet: $f(x_i|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1}{2\sigma^2}(x_i-\mu)^2}$

1 $\Rightarrow L(\mu,\sigma^2) = \prod_{i=1}^{n} f(x_i|\mu,\sigma^2)$

$= (2\pi\sigma^2)^{-n/2} \exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$

1 $l(\mu,\sigma^2) = \log L(\mu,\sigma^2) = \frac{-n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$

1 $S(\mu) = \partial l / \partial\mu = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu)$

1 $S(\hat{\mu}) = 0 \Rightarrow \hat{\mu} = \bar{X}$

1 $S(\sigma^2) = \partial l / \partial\sigma^2 = \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$

1 $S(\hat{\sigma}^2) = 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2$

(c) (4 marks) Give two separate arguments as to why $\hat{\mu} = \bar{X}$ and $s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ are consistent for $(\mu, \sigma^2)$

2 each, any 2 of 3:

I) They are MoM, always consistent

II) They are MLE, always consistent (note: we assumed regularity)

III) Prove directly using LLN

(d) (~~4~~ 2 marks) Give ~~two separate~~ One arguments as to why $\hat{\mu} = \bar{X}$ and $s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ are sufficient for $(\mu, \sigma^2)$

2 ~~4~~

I) MLE always sufficient

OR: II) $f(\underline{x}|\mu,\sigma^2) = L(\mu,\sigma^2) = \ldots$

Prove by factorization. Much harder.

2. (20 marks) A particular coffee chain runs a particular campaign every winter, where each coffee you buy has a chance to win something. In particular, they claim that 1 out of every 6 coffees is a winner. This year, I purchased $n = 20$ coffees from this chain during this campaign. Let $X_i = 1$ be the event that I win on the $i^{th}$ coffee and $X_i = 0$ if I lose. Then $X_i \sim Bern(\theta)$ with

$$P(X_i = x) = \theta^x(1-\theta)^{1-x}$$

for $x = 0, 1$ and unknown parameter $\theta \in (0, 1)$. My results (they are sad) were as follows:

$$\mathbf{x} = (0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0)$$

(a) (2 marks) Write the likelihood function appropriate for making inferences about $\theta$, assuming independence between the $X_i$.

2 $L(\theta) = \prod_{i=1}^{n} P(X_i = x) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}$

0 marks if they don't actually write it out.

(b) (2 marks) State a sufficient statistic for $\theta$- you can just look at the likelihood and write down the answer, you don't have to prove it.

2 $\sum_{i=1}^{n} X_i$.

(c) (2 marks) Given that I only won once, give a brief explanation as to why it doesn't matter which of the $n = 20$ coffees was the single winner, in terms of making inferences about $\theta$ via the likelihood function.

2 All datasets $\underline{x}$ that give the same $\sum x_i$ give the same likelihood function for $\theta$ (for fixed $n$)

(d) (2 marks) Find the maximim likelihood estimator of $\theta$ and state the corresponding maximum likelihood estimate for these observed data. You can assume the regularity conditions are satisfied.

$$\ell(\theta) = \Sigma x_i \log\theta + (n - \Sigma x_i)\log(1-\theta)$$

$$S(\theta) = \partial \ell / \partial \theta = \Sigma x_i / \theta - (n - \Sigma x_i)/(1-\theta)$$

1 $S(\hat{\theta}) = 0 \Rightarrow \hat{\theta} = \bar{X}$

1 $\hat{\theta} = 1/20$

(e) (2 marks) Find the Fisher Information for $\theta$ at the true value $\theta = \theta_0$. Note $E(X) = \theta_0$.

1 $J(\theta) = -\partial S / \partial \theta = \dfrac{\Sigma x_i}{\theta^2} + \dfrac{n - \Sigma x_i}{(1-\theta)^2}$

$$I(\theta_0) = E J(\theta_0) = \frac{n}{\theta_0} + \frac{n}{1-\theta_0}$$

1 $$= \frac{n}{\theta_0(1-\theta_0)}$$

(f) (2 marks) I'm mad that I barely won anything, so I want to know: is the claim about the chances of winning supported by the observed data? State the null hypothesis and corresponding alternative hypothesis we would use to test the company's claim.

(1) $H_0: \theta = 1/6$

(1) $H_1: \theta \neq 1/6$.

(g) (1 mark) Is $H_0$ simple or composite?

(1) Simple.

(h) (2 marks) State a test statistic that has an approximate $N(0,1)$ distribution if $H_0$ is true. Make sure to clearly state the theorem you use, and to simplify your answer as far as possible.

From formula sheet: $\dfrac{\hat{\theta} - \theta_0}{1/\sqrt{I(\theta_0)}} \xrightarrow{d} N(0,1)$

Under $H_0$, $\theta_0 = 1/6$.

(2) $$\frac{\hat{\theta} - \theta_0}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} \approx N(0,1)$$

Full marks if they plug in the numbers, as long as they get it right.

(i) (1 mark) Evaluate this test statistic for the observed data and our $H_0$.

① $$\frac{\hat{\theta} - \theta_0}{\sqrt{\frac{\theta_0(1-\theta_0)}{n}}} = \frac{1/20 - 1/6}{\sqrt{\frac{1/6(5/6)}{20}}} = -1.4.$$

(j) (4 marks) State your conclusion at the 5% significance level. An answer worth full marks will include both a mathematical conclusion, and a conclusion in words. Convince me whether or not the observed data suggests that the company's claim about the chances of winning is reasonable.

② From formula sheet, critical region is $t(x) < -1.96$ or $t(x) > 1.96$. $t(x) = -1.4$, so fail to reject $H_0: \theta = 1/6$ at the 5% level.

② The observed data does not provide evidence to suggest that the chances of winning on one cup of coffee are different than one in 6.

3. (16 marks) For the campaign in the previous question, the particular coffee chain also claims that each of their cup sizes has an equal chance of winning. Supposing this chain has sizes of {small, medium, large}, I go out and buy {9, 41, 17} coffees of each respective size, and I win {1, 14, 2} times.

(a) (3 marks) Write these data in an appropriate contingency table. Specify all entries in the table, as well as all row and column totals, and the grand total. Clearly label which variable is in the rows, and which is in the columns.

3

| | small | med | large | |
|---|---|---|---|---|
| win | 1 | 14 | 2 | 17 |
| lose | 8 | 27 | 15 | 50 |
| | 9 | 41 | 17 | 67 |

OR:

| | win | lose | |
|---|---|---|---|
| small | 1 | 8 | 9 |
| med | 14 | 27 | 41 |
| large | 2 | 15 | 17 |
| | 17 | 50 | 67 |

(b) (2 marks) State the appropriate null and alternative hypotheses for using a likelihood ratio test to test the hypothesis that the chances of winning are independent of the cup size.

1 $H_0$: $P_{ij} = P_{i\cdot} \times P_{\cdot j}$ for all $i,j$; where $P_{ij}$ = Probability of cell $i,j$ in above table.

1 $H_1$: at least 1 $P_{ij} \neq P_{i\cdot} \cdot P_{\cdot j}$

(c) (2 marks) What is the number of free parameters under $H_0$, and under $H_1$?

1 $H_0: (R-1) + (C-1) = 1 + 2 = 3$

1 $H_1: RC - 1 = (2)(3) - 1 = 5$

(d) (2 marks) State the asymptotic distribution of the likelihood ratio statistic $-2\log\Lambda$, including the appropriate degrees of freedom.

① for $\chi^2$

$$-2\log\Lambda \xrightarrow{d} \chi^2_{2}$$

(oops!)

① for df.

(e) (3 marks) Compute the likelihood ratio statistic $-2\log\Lambda$ for these data

① $-2\log\Lambda = 2\sum_{i=1}^{R}\sum_{j=1}^{C} Y_{ij}\log\left(\frac{NY_{ij}}{r_i c_j}\right)$

① $= 2\Big(1\times\log\left(\frac{67\times 1}{9\times 17}\right) + 14\times\log\left(\frac{67\times 14}{41\times 17}\right) + 2\times\log\left(\frac{67\times 2}{17\times 17}\right)$

$+ 8\times\log\left(\frac{67\times 8}{9\times 50}\right) + 27\times\log\left(\frac{67\times 27}{41\times 50}\right) + 15\log\left(\frac{67\times 15}{17\times 50}\right)\Big)$

$= 2\big(-0.8257 + 4.1575 + (-1.5372)$

$+ 1.399 + (-3.3768) + 2.5126\big)$

$= 2(2.3294)$

① $= 4.6588.$

(f) (4 marks) State your conclusions at the 5% signficance level, mathematically and in words. You don't need to provide a p-value, but the formula sheet gives enough information to make an explicit decision about $H_0$.

(1) $\chi^2_2$ 95% crit. val is 5.99 from formula sheet

(1) Fail to reject $H_0$ at 95% sig. level ($-2\log\Lambda < 5.99$)

(2) The observed data do not provide evidence at this significance level to conclude that the size of the cup is related to whether or not it is a winner.

4. (20 marks) Suppose $X_i \sim N(\mu, \sigma^2)$ is an IID sample, where $\sigma^2$ is known, and we wish to test $H_0 : \mu = \mu_0$ against $H_0 : \mu \neq \mu_0$ at the $\alpha$ significance level using the test statistic

$$T(\mathbf{X}) = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

(a) (2 marks) State what it means for $T(\mathbf{X})$ to be a pivot for $\mu$.

2 $T(X)$ depends on $\mu_0$ and on $X$, but has a known dist$^n$ that does not depend on $\mu_0$.

(b) (2 marks) Define the significance level of the test, $\alpha$, in words. There is more than one possible correct answer, so just write down the one you know.

2 Probability of committing a type I error (rejecting $H_0$ when true)

OR

Highest p-value for which we would reject $H_0$.

(c) (4 marks) Give an explicit formula for the p-value of this test, simplified as much as possible. Denote the observed value of the test statistic as $t(\mathbf{x})$.

1 $$P_0 = P(|T(X)| \geq |t(x)|)$$

(can give part marks here).

$$= 1 - P(-|t(x)| \leq T(X) < |t(x)|)$$

$$= 1 - \Phi(|t(x)|) + 1 - \Phi(|t(x)|)$$

3 $$= 2(1 - \Phi(|t(x)|))$$

(d) (7 marks) Define the power of the test, $\eta$, in words. Clearly indicate the three variables on which $\eta$ depends, and briefly explain how power analysis is used to determine the sample size in a scientific study.

$\eta = P(\text{reject } H_0 \mid H_0 \text{ false})$

(1) The power of the test is the ability/sensitivity of the test to detect a false $H_0$. It depends on

(1) $d = \frac{\mu_0 - \mu_1}{\sigma}$, the effect size (how wrong $H_0$ is)

(1) $n$; the sample size.

(1) $\alpha$, the significance level of the test.

To determine the sample size for a scientific study,

- decide on an effect size of interest
- decide on the power with which you would like to be able to detect this effect
- choose the significance level of the test.

(3)

- Plug these in ~~th~~ to the power function, and determine the smallest $n$ that gives the requisite power.

(e) (5 marks) Derive the power function for this test to detect $\mu = \mu_1 \neq \mu_0$. State your answer in terms of the effect size:

$$d = \frac{\mu_0 - \mu_1}{\sigma}$$

1 $\eta = P(\text{reject } H_0 \mid H_0 \text{ false}, \mu = \mu_1)$

1 $= 1 - P\left(-z_{1-\alpha/2} < \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < z_{1-\alpha/2}\right)$

$= 1 - P\left(\frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} - z_{1-\alpha/2} < \frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} < \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + z_{1-\alpha/2}\right)$

1 $= 1 - P(d\sqrt{n} - z_{1-\alpha/2} < Z < d\sqrt{n} + z_{1-\alpha/2})$

1 Because if $\mu = \mu_1$, $Z = \frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} \sim N(0,1)$

1 $= 1 - \left(\Phi(d\sqrt{n} + z_{1-\alpha/2}) - \Phi(d\sqrt{n} - z_{1-\alpha/2})\right)$

5. (16 marks) Let $X_i \sim Exp(\theta), i = 1 \ldots n$ be a random sample from an exponential distribution with the parametrization

$$f(x|\theta_0) = \theta_0 e^{-x\theta_0}$$
$$E(X) = 1/\theta_0$$
$$Var(X) = 1/\theta_0^2$$
$$\text{MLE: } \hat{\theta} = 1/\bar{X}$$
$$\text{Fisher Info: } I(\theta_0) = n/\theta_0^2$$

where $\theta_0$ is the true value of $\theta$.

(a) (2 marks) Find $E(\bar{X})$ and $Var(\bar{X})$, where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

1 $E(\bar{X}) = \frac{1}{n}\sum E(X_i) = \frac{n\,1/\theta_0}{n} = 1/\theta_0 = E(X)$

1 $Var(\bar{X}) = \frac{1}{n^2}\sum Var(X_i) = \frac{n\,1/\theta_0^2}{n^2} = \frac{1}{n\theta_0^2}$

(b) (3 marks) Use the Central Limit Theorem for the sample mean to find an approximate 95% confidence interval for $1/\theta$. $Var(\bar{X})$ depends on $\theta_0$, so replace it with an appropriate consistent estimator.

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{Var(\bar{X})}} \overset{\text{approx}}{\sim} N(0,1)$$

$\Rightarrow$ Approx 95% CI for $E(\bar{X}) = \frac{1}{\theta}$ is

$$\left(\bar{X} - \frac{1}{\theta_0\sqrt{n}} z_{1-\alpha/2},\ \bar{X} + \frac{1}{\theta_0\sqrt{n}} z_{1-\alpha/2}\right)$$

$$\approx \left(\bar{X} - \frac{\bar{X}}{\sqrt{n}} z_{1-\alpha/2},\ \bar{X} + \frac{\bar{X}}{\sqrt{n}} z_{1-\alpha/2}\right)$$

$\alpha = 0.05$
OR they can plug in $z_{1-\alpha/2} = 1.96$.

(c) (3 marks) Use your answer to the previous question to find an approximate 95% confidence interval for $\theta$. Call this interval $V_n$

-1 mark if they don't flip the +, - in the bounds.

Invert the bounds of the previous interval to find an approx interval for $\theta = (1/\theta)^{-1}$:

3

$$\left(\left(\bar{X} + \frac{\bar{X}}{\sqrt{n}} z_{1-\alpha/2}\right)^{-1}, \left(\bar{X} - \frac{\bar{X}}{\sqrt{n}} z_{1-\alpha/2}\right)^{-1}\right) = V_n$$

(d) (3 marks) Use the Central Limit Theorem for the MLE to find an approximate 95% confidence interval for $\theta$. Call this interval $W_n$.

1

$$\frac{\hat{\theta} - \theta_0}{\sqrt{1/I(\theta_0)}} \overset{\text{approx}}{\sim} N(0,1)$$

$\Rightarrow$ Approx interval for $\theta$ is $\left(\hat{\theta} - \frac{z_{1-\alpha/2}}{\sqrt{I(\theta_0)}}, \hat{\theta} + \frac{z_{1-\alpha/2}}{\sqrt{I(\theta_0)}}\right)$

1

$$= \left(\frac{1}{\bar{X}} - \frac{z_{1-\alpha/2}}{\sqrt{n/\theta_0^2}}, \frac{1}{\bar{X}} + \frac{z_{1-\alpha/2}}{\sqrt{n/\theta_0^2}}\right)$$

Replace $\theta_0$ by $\hat{\theta} = 1/\bar{X}$:

1

$$W_n = \left(\frac{1}{\bar{X}} - \frac{\bar{X} z_{1-\alpha/2}}{\sqrt{n}}, \frac{1}{\bar{X}} + \frac{\bar{X} z_{1-\alpha/2}}{\sqrt{n}}\right)$$

(e) (5 marks) These intervals for $\theta$ are different for any finite $n$. Show that as $n \to \infty$, both intervals converge in probability to the singleton set $\{\theta_0\}$:

$$V_n \xrightarrow{p} \{\theta_0\}$$
$$W_n \xrightarrow{p} \{\theta_0\}$$

If you use a familiar theorem, be sure to state it.

$$V_n = \left( \left( \bar{X} + \frac{\bar{X} z_{1-\alpha/2}}{\sqrt{n}} \right)^{-1}, \left( \bar{X} - \frac{\bar{X} z_{1-\alpha/2}}{\sqrt{n}} \right)^{-1} \right)$$

2 $\xrightarrow{p} \left( \frac{1}{\frac{1}{\theta_0} + 0}, \frac{1}{\frac{1}{\theta_0} - 0} \right)$ because by LLN, $\bar{X} \xrightarrow{p} E(X) = \frac{1}{\theta_0}$ and $\frac{\bar{X}}{\sqrt{n}} \xrightarrow{p} 0$, and $g(x) = x^{-1}$ is continuous.

0.5 $= (\theta_0, \theta_0) = \{\theta_0\}$.

$$W_n = \left( \frac{1}{\bar{X}} - \frac{\bar{X} z_{1-\alpha/2}}{\sqrt{n}}, \frac{1}{\bar{X}} + \frac{\bar{X} z_{1-\alpha/2}}{\sqrt{n}} \right)$$

2 $\xrightarrow{p} \left( \frac{1}{1/\theta_0} - 0, \frac{1}{1/\theta_0} + 0 \right)$ Because (same as $V_n$)

0.5 $= (\theta_0, \theta_0) = \{\theta_0\}$.

THIS PAGE IS FOR ROUGH WORK. NOTHING ON THIS PAGE WILL BE MARKED

THIS PAGE IS FOR ROUGH WORK. NOTHING ON THIS PAGE WILL BE MARKED

THIS PAGE IS FOR ROUGH WORK. NOTHING ON THIS PAGE WILL BE MARKED