STAT 415 — Final (Spring 2020)

Name: ______________________________

Exam rules:

- You have 2hrs 15 minutes to to complete the exam, scan it and upload onto Gradebook.
- There are **5** Questions (The exam is worth 40pts in total).
- This exam is open book. You are free to use your class notes and HW solutions to solve the problems.

  Do not do a brain dump. Answers which are irrelevant will be penalized.

  Do not blindly copy answers from the HW solutions.
- State precisely in the derivations all the results that you use.
- Explain your arguments carefully, show all your work, and write out all your steps. Credit will be given for clarity. No credit will be given for a solution without working.
- If you are caught cheating or helping someone to cheat on this exam, you will both receive a grade of zero on the exam and you will be turned in.
- If a problem requires a numerical answer, you may express answer in terms of elementary functions or cdfs (e.g. $\Gamma(r+1)$ or $r!$).
- Presentation and scanning rules (20% will be knocked off if these rules are not adhered too).
  1. No light pencils. Black is easiest to read.
  2. Write big; scrawls are painful to read.
  3. Name at top of exam paper (this makes it easier to identify them).
  4. Name the submitted file LASTNAME_FIRSTNAME.pdf
  5. Submit only **one** pdf file.
  6. The solutions should be on separate pages.
  7. Check clarity of scan before submission.

(1) Suppose the random variables $\underline{X} = (X_1, X_2, X_3, X_4, X_5)$ have the covariance matrix

$$\begin{pmatrix} 1 & 1/2 & 0 & 0 & 0 \\ 1/2 & 1 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1 \end{pmatrix}$$

Obtain its covariance network. [3]

(2) Suppose $\{(X_t, Y_t)\}_{t=1}^{n}$ are independent bivariate random variables (meaning that over t they are independent), where

$$X_t = 2t + \varepsilon_t \qquad Y_t = \cos\left(\frac{2\pi t}{3}\right) + e_t$$

and $\{\varepsilon_t, e_t\}$ are random with $\mathrm{E}[\varepsilon_t] = \mathrm{E}[e_t] = 0$, $\mathrm{var}[\varepsilon_t] = \mathrm{var}[e_t] = \sigma^2$ and $\delta = \mathrm{cov}(\varepsilon_t, e_t)$. Let

$$S_n = \frac{1}{n}\sum_{t=1}^{n}(X_t - Y_t)\,.$$

Obtain $\mathrm{E}[S_n]$ and $\mathrm{var}[S_n]$. [5]

(3) Suppose that $\{X_i\}$ are iid random variables with mean μ ($\mu \neq 0$) and variance σ^2. Let $\bar{X}$ denote the sample mean.

(i) Obtain the asymptotic distribution of $\log \bar{X}$.

(ii) Based on (ii) construct an approximate 95% confidence interval for $\log \mu$. The confidence interval should be computable, where unknown parameters are replaced with their estimators.

[5]

4 Suppose that $\{X_i\}_{i=1}^n$ are iid Binomial random variables where $X_i \sim Bin(m,p)$.

(i) If $n = 3$ and $m = 10$, obtain the rejection region for the log-likelihood ratio test at the 5% level for

$$H_0 : p = 0.6 \text{ vs } H_A : p = 0.8.$$

Give all working (you may use R).

Hint: Recall that the Binomial distribution is the sum of Bernoulli random variables. I.e. If if $S_n = \sum_{i=1}^K \delta_i$ (where $\{\delta_i\}$ are iid, $\delta_i \in \{0,1\}$ and $P(\delta_i = 1) = p$), then $S_n \sim Bin(K,p)$.

(ii) Calculate the power of the test for $p = 0.8$. Briefly explain why this test should be used for detecting the alternative $p = 0.8$?

(iii) You observe the data $(x_1, x_2, x_3) = (8, 6, 7)$. Based on the data what are the conclusions of the test at the 5% level.

[12]

- Bonus question (worth nothing). Submit (if you want) after the exam, through email.

  Obtain an alternative rejection region for (X_1, X_2, X_3), where the probability of a Type I error is 5% or less. Obtain the power for this new test when $p = 0.8$.

5 Suppose that $\{X_i\}_{i=1}^n$ are iid normal random variables with mean μ and variance σ^2. We will assume in this question that μ is known.

You may use that if $Z \sim N(0, \sigma^2)$, then $\text{var}[Z^2] = 2\sigma^4$.

(i) Obtain the maximum likelihood estimator of σ^2. Denote this as $\widehat{\sigma}^2_{MLE}$

(ii) Derive the exact distribution of $\widehat{\sigma}^2_{MLE}$.

(iii) Derive the asymptotic distribution of $\widehat{\sigma}^2_{MLE}$.

(iv) Derive the generalized log-likelihood ratio statistic for

$$H_0 : \sigma^2 = \sigma_0^2 \text{ vs } H_A : \sigma^2 \neq \sigma_0^2.$$

(v) Derive the (asymptotic) rejection region for the test statistic at the 5% level.

[15]

- Bonus Question (worth nothing), Submit (if you want) after the exam, through email.

  Make a sketch of the distribution of $2 \log \Delta(\underline{x})$, under the alternative $\sigma^2 = \sigma_1^2$. Explaining if and how the test has power.