STAT 510, Final Exam

May 8, 2020, due at 7PM

Name: TIANQI WU

NetID: twu38

Please sign the following pledge and read all instructions carefully before starting the exam.

Pledge: I have neither given nor received any unauthorized aid in completing this exam, and I have conducted myself within the guidelines of the University Student Code.

Signature: ______________________

INSTRUCTIONS:

- This is a take-home exam. However, you are not allowed to discuss with anyone else and should finish the exam by your own.
- **Show all work,** clearly and in order, if you want to receive full credit. I reserve the right to take off points if I cannot see how you arrived at your answer (even if your final answer is correct).
- **Answer all the questions in the space provided. You may attach additional sheets if necessary.**
- This test has 4 regular questions and one bonus question ($100 + 5$ points). It is your responsibility to make sure that you have all of the questions.
- **Good luck!**

Que. No.	Max Points	Earned Pts.
1	25	
2	25	
3	25	
4	25	
5 (bonus)	5	

TOTAL: __________

Question 1. (25 points) Let the distribution of $\mathbf{X} = (X_1, X_2, X_3)^T$ be $N(\boldsymbol{\mu}, \Sigma)$ (multivariate normal) with $\boldsymbol{\mu}^T = (-2, 1, 1)$ and $\Sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}$.

(a) (10 points) Find the distribution of $3X_1 - 2X_2 + X_3$.

(b) (15 points) Find a vector $\boldsymbol{a} \in \mathbb{R}^2$ such that (X_1, X_2) is independent of $X_2 - \boldsymbol{a}^T \begin{pmatrix} X_1 \\ X_3 \end{pmatrix}$.

(a) Since $X = \mu + BZ$, where $Z \overset{iid}{\sim} N(0,1)$. $\Sigma = BB'$

$DX = D(\mu + BZ) = D\mu + (DB)Z$

Hence, $DX \sim N(D\mu, DBB'D')$

$DX \sim N(D\mu, D\Sigma D')$, let $D = (3, -2, 1)$

$D\mu = (3, -2, 1)\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} = -7$

$D\Sigma D' = (3, -2, 1)\begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix} = 9.$

Hence, $3X_1 - 2X_2 + X_3 \sim \boxed{N(-7, 9)}$

(b). To let (X_1, X_2) independent of $X_2 - a^T\begin{pmatrix} X_1 \\ X_3 \end{pmatrix}$ $\Rightarrow$

$$\begin{cases} Cov(X_1, X_2 - a^T\begin{pmatrix} X_1 \\ X_3 \end{pmatrix}) = 0 \\ Cov(X_2, X_2 - a^T\begin{pmatrix} X_1 \\ X_3 \end{pmatrix}) = 0 \end{cases} \Rightarrow$$

$$\begin{cases} Cov(X_1, X_2) - a_1 Cov(X_1, X_1) - a_2 Cov(X_1, X_3) = 0 \\ Cov(X_2, X_2) - a_1 Cov(X_2, X_1) - a_2 Cov(X_2, X_3) = 0 \end{cases} \Rightarrow$$

$$\begin{cases} 1 - a_1 - a_2 = 0 \\ 3 - a_1 - 2a_2 = 0. \end{cases} \Rightarrow$$

$$\begin{cases} a_1 = -1 \\ a_2 = 2 \end{cases}$$

$$\boxed{a = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.}$$

Question 2. (25 points) Let $X_1, \ldots, X_n$ be a sample of size n from the exponential distribution

$$f_\theta(x) = \theta e^{-\theta x}, \quad x > 0,$$

where $\theta > 0$ is the parameter. In this problem, we consider a Bayesian framework for parameter estimation and hypothesis testing.

(a) (10 points) For parameter estimation, suppose we use a Gamma prior $\mathrm{Ga}(\alpha, \beta)$ for parameter θ, that is,

$$p(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \quad \text{for } \theta > 0.$$

Find the posterior distribution of θ given data $(X_1, X_2, \ldots, X_n)$. What would be a proper Bayesian estimator of θ based on this posterior?

(b) (15 points) For Bayesian hypothesis testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$, suppose we use the hierarchical prior:

$$p(H_0) = p(H_1) = \frac{1}{2},$$
$$\theta \,|\, H_0 \sim \text{point mass distribution at } \theta_0,$$
$$\theta \,|\, H_1 \sim \mathrm{Ga}(\alpha, \beta).$$

What is the posterior probability of the null? Find the reject rule in terms of $\{X_i\}_{i=1}^n$ if we want to reject the null when its posterior probability falls below 0.05.

(a) $f(X|\theta) = \prod_{i=1}^{n} (\theta \cdot e^{-\theta x}) = \theta^n e^{-\theta \Sigma x_i}$

$P(\theta|X) \propto f(X|\theta) \cdot P(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \theta^{n+\alpha-1} \cdot e^{-(\beta+\Sigma x_i)\theta}$

Hence, $P(\theta|X) \sim \boxed{Ga(n+\alpha, \beta+\Sigma x_i)}$, $\hat{\theta} = \boxed{\frac{n+\alpha}{\beta+\Sigma x_i}}$

(b). $P(X|H_0) = \theta_0^n e^{-\theta_0 \Sigma x_i}$

$P(X|H_1) = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \theta^{n+\alpha-1} e^{-(\beta+\Sigma x_i)\theta} d\theta$

$= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(n+\alpha)}{(\beta+\Sigma x_i)^{n+\alpha}}$ · pdf of $Ga(n+\alpha, \beta+\Sigma x_i)$

$BA_0 = P(X|H_1) / P(X|H_0)$

$P(H_0|X) = \frac{1}{1+BA_0} = \frac{1}{1 + \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(n+\alpha)}{(\beta+\Sigma x_i)^{n+\alpha}} \cdot \frac{1}{\theta_0^n e^{-\theta_0 \Sigma x_i}}}$

$P(H_0|X) < 0.05 \Rightarrow$.

$\Gamma(\alpha) \cdot (\beta+\Sigma x_i)^{n+\alpha} \cdot \theta_0^n e^{-\theta_0 \Sigma x_i} < \frac{\beta^\alpha \cdot \Gamma(n+\alpha)}{19} \Rightarrow$

$$\boxed{(\beta+\Sigma x_i)^{n+\alpha} \cdot e^{-\theta_0 \Sigma x_i} < \frac{\beta^\alpha \cdot \Gamma(n+\alpha)}{19 \cdot \Gamma(\alpha) \cdot \theta_0^n}}$$

Question 3. (25 points) Consider the same sample $\{X_i\}_{i=1}^n$ of size n from the exponential distribution in Question 2. Now we will focus on the frequentist procedure for parameter estimation.

(a) (5 points) Find the maximum likelihood estimator (MLE) $\widehat{\theta}_n$.

(b) (5 points) We know that as the MLE, $\widehat{\theta}_n$ is a consistent estimator of θ, and is asymptotically normal. Find expressions for the asymptotic bias $b(\theta)$ and the asymptotic variance $V(\theta)$ of $\widehat{\theta}_n$, so that

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{d} N(b(\theta), V(\theta)) \quad \text{as } n \to \infty.$$

(c) (5 points) Construct a confidence interval for θ with asymptotic level 0.95 based on the asymptotic distribution of $\widehat{\theta}_n$ you derived in (b). (The 97.5% quantile of the standard normal is 1.96.)

(d) (5 points) Use the Delta method to find a function $g : (0, \infty) \to \mathbb{R}$ such that the asymptotic variance of $g(\widehat{\theta}_n)$ is a constant independent of θ, or

$$\sqrt{n}\left(g(\widehat{\theta}_n) - g(\theta)\right) \xrightarrow{d} N(0, \sigma^2) \quad \text{as } n \to \infty,$$

for some constant σ^2 that is independent of θ. Such a function g is called a variance stabilizing transformation of the MLE $\widehat{\theta}_n$.

(e) (5 points) Construct a confidence interval for θ with asymptotic level 0.95 based on the result in (d). Which confidence interval is better, the one constructed in (c) or the one here?

(a) $L(\theta \mid X) = \prod_{i=1}^{n} \theta e^{-\theta x} \Rightarrow l(\theta, x) = n \log\theta - \theta \Sigma x_i \Rightarrow$

$l'(\theta, x) = \frac{n}{\theta} - \Sigma x_i = 0 \Rightarrow \boxed{\hat{\theta} = \frac{n}{\Sigma x_i}}$

(b). $I_X(\theta) = -E_\theta[\frac{d^2}{d\theta^2} l(\theta \mid X)] = E_\theta[\frac{1}{\theta^2}] = \frac{1}{\theta^2}$

Since, $\hat{\theta}$ is consistent and asymptotically normal.

$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, I_X^{-1}(\theta)) = N(0, \theta^2) \Rightarrow \boxed{b(\theta) = 0,\ V(\theta) = \theta^2}$

(c). $[\hat{\theta}_n - \frac{1.96}{\sqrt{n}}\hat{\theta}_n,\ \hat{\theta}_n + \frac{1.96}{\sqrt{n}}\hat{\theta}_n]$, where $\hat{\theta}_n = \frac{n}{\Sigma x_i}$

(d). Since $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \theta^2)$, let $\boxed{g(\theta) = \ln\theta}$ $g'(\theta) = \frac{1}{\theta}$.

Then, $\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{d} \frac{1}{\theta} \cdot N(0, \theta^2) = N(0, 1) \Rightarrow \sigma^2 = 1$.

(e). CI for $\ln(\hat{\theta}_n)$: $[\ln(\hat{\theta}_n) - \frac{1.96}{\sqrt{n}}\sqrt{1},\ \ln(\hat{\theta}_n) + \frac{1.96}{\sqrt{n}}]$

CI for $\hat{\theta}_n$: $[\hat{\theta}_n \cdot e^{-\frac{1.96}{\sqrt{n}}},\ \hat{\theta}_n \cdot e^{\frac{1.96}{\sqrt{n}}}]$

$R_{(c)}$: range of CI in (c) : $(\hat{\theta}_n \frac{2 \cdot 1.96}{\sqrt{n}})$

$R_{(e)}$: range of CI in (e) : $(\hat{\theta}_n(e^{\frac{1.96}{\sqrt{n}}} - e^{-\frac{1.96}{\sqrt{n}}}))$

let $a = \frac{1.96}{\sqrt{n}}$, then, $R_{(e)} - R_{(c)} = \hat{\theta}_n(e^a - e^{-a} - 2a)$

Since $e^a - e^{-a} - 2a > 0$ when $a > 0$, $R_{(e)} > R_{(c)}$

Hence, $\boxed{\text{CI in (c) is better}}$ since its range is smaller.

Question 4. (25 points) Suppose $X_1, X_2, \ldots, X_n$ are independently and identically distributed (iid) with a Beta$(\alpha, 1)$ distribution and $Y_1, Y_2, \ldots, Y_m$ are iid with a Beta$(\beta, 1)$ distribution. Assume $\{X_i\}_{i=1}^n$ are independent of $\{Y_i\}_{i=1}^m$.

(a) (10 points) Find likelihood ratio test of $H_0 : \alpha = \beta$ versus $H_1 : \alpha \neq \beta$.

(b) (5 points) Show that the test in part (a) can be based on the statistic

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

(c) (10 points) Find the distribution of T when H_0 is true, and then show how to get a test of size $\alpha = 0.05$.

(a) under H_1: $L(\alpha, \beta \mid x, y) = \alpha^n \left(\prod_i x_i\right)^{\alpha-1} \beta^m \left(\prod_{i=1}^m y_i\right)^{\beta-1}$

$l' = \frac{n}{\alpha} + \sum_{i=1}^n \log x_i = 0 \Rightarrow \hat{\alpha} = -\frac{n}{\sum \log x_i}$, similarly, $\hat{\beta} = -\frac{m}{\sum \log y_i}$

under H_0: $L(\alpha_0 \mid x, y) = \alpha_0^{n+m} \left(\prod_{i=1}^n x_i \prod_{i=1}^m y_i\right)^{\alpha_0 - 1}$

$l' = \frac{n+m}{\alpha_0} + \sum_{i=1}^n \log x_i + \sum_{i=1}^m \log y_i = 0 \Rightarrow \hat{\alpha}_0 = -\frac{n+m}{\sum \log x_i + \sum \log y_i}$

$$\Lambda(x, y) = \frac{L(\alpha, \beta \mid x, y)}{L(\alpha_0 \mid x, y)} = \frac{\hat{\alpha}^n \hat{\beta}^m \left(\prod^n x_i\right)^{\hat{\alpha}-1} \cdot \left(\prod^m y_i\right)^{\hat{\beta}-1}}{\hat{\alpha}_0^{n+m} \left(\prod^n x_i \prod^m y_i\right)^{\hat{\alpha}_0 - 1}}$$

$$\boxed{= \frac{\hat{\alpha}^n \hat{\beta}^m}{\hat{\alpha}_0^{n+m}} \left(\prod^n x_i\right)^{\hat{\alpha} - \hat{\alpha}_0} \left(\prod^m y_i\right)^{\hat{\beta} - \hat{\alpha}_0}}$$

(b). $\ln(\Lambda(x, y)) = n \ln \hat{\alpha} + m \ln \hat{\beta} - (n+m) \ln \hat{\alpha}_0 + \underbrace{(\hat{\alpha} - \hat{\alpha}_0) \sum^n \log x_i + (\hat{\beta} - \hat{\alpha}_0) \sum^n \log y_i}_{\to 0}$

$= n \ln \hat{\alpha} + m \ln \hat{\beta} - (n+m) \ln \hat{\alpha}_0$

$$\lambda(x, y) = \left(-\frac{n}{\sum \log x_i}\right)^n \left(-\frac{m}{\sum \log y_i}\right)^m \left(-\frac{\sum \log x_i + \sum \log y_i}{n+m}\right)^{n+m}$$

$$\boxed{= \frac{n^n m^n}{(n+m)^{n+m}} \left(\frac{1}{1-T}\right)^m \left(\frac{1}{T}\right)^n}$$

(c). $\log x_i \sim \exp(\alpha)$, $\log y_i \sim \exp(\beta)$.

$\sum_{i=1}^n \log x_i \sim Ga(n, \alpha)$, $\sum_{i=1}^m \log y_i \sim Ga(m, \beta)$

under H_0, $\alpha = \beta$, $\sum_{i=1}^n \log x_i + \sum_{i=1}^m \log y_i = Ga(n+m, \alpha)$

Hence, $\boxed{T \sim \text{Beta}(n, m)}$. To get a test of size α, we need

$$\boxed{P(C_1 \leq T \leq C_2) = 0.05, \left(\frac{1}{1-C_1}\right)^m \left(\frac{1}{C_1}\right)^n = \left(\frac{1}{1-C_2}\right)^m \left(\frac{1}{C_2}\right)^n}$$